А.И.Орлов

ЭКОНОМЕТРИКА

Учебник

Москва
"Экзамен"
2002

Содержание

Предисловие

Эконометрика исследует конкретные количественные и качественные взаимосвязи экономических объектов и процессов с помощью математических и статистических методов и моделей.

Как учебная дисциплина эконометрика изучается после прохождения курсов теории вероятностей и математической статистики и общей теории статистики (иногда - экономической статистики). Эти дисциплины обязательны для подготовки экономистов и менеджеров, особенно в технических вузах.

Целью изучения учебной дисциплины "Эконометрика" является овладение современными эконометрическими методами анализа конкретных экономических данных на уровне, достаточном для использования в практической деятельности менеджера и менеджера-экономиста, инженера.

Основные задачи курса - изучение современных эконометрических методов и моделей, в том числе методов прикладной статистики (статистики случайных величин, многомерного статистического анализа, временных рядов, статистики нечисловых и интервальных данных), экспертного оценивания, эконометрических моделей инфляции, инвестиций, качества, прогнозирования и риска.

Теоретическую базу эконометрики составляют математические дисциплины - общий курс (математический анализ, линейная алгебра), теория вероятностей и математическая статистика, дискретная математика, исследование операций; а также основы экономической теории и статистика (общая теория статистики, экономическая статистика).

Настоящий учебник соответствует Государственным образовательным стандартам по экономическим дисциплинам. Кроме того, материалы учебника можно использовать при изучении курсов "Математические методы прогнозирования", "Экономика отрасли", "Прогнозирование и технико-экономическое планирование", "Экология и инвестиционная активность предприятия", "Экономика предприятия" и др.

Учебник адресован в первую очередь студентам дневных отделений экономических специальностей. Они найдут весь необходимый материал для изучения различных вариантов эконометрических курсов. Особенно хочется порекомендовать учебник тем, кто получает наиболее ценимое в настоящее время образование - на экономических факультетах в технических вузах. Слушатели вечерних отделений, в том числе получающие второе образование по экономике и менеджменту, смогут изучить основы эконометрики и познакомиться с основными вопросами ее практического использования. Менеджерам, экономистам и инженерам, изучающим эконометрику самостоятельно или в институтах повышения квалификации, учебник позволит познакомиться с ее ключевыми идеями и выйти на мировой уровень образования. Специалистам по теории вероятностей и математической статистике эта книга также может быть интересна и полезна, в ней описан современный взгляд на прикладную математическую статистику, основные подходы и результаты в этой области, открывающие большой простор для дальнейших математических исследований.

В отличие от учебной литературы по математическим дисциплинам, в настоящей книге практически отсутствуют доказательства. В нескольких случаях мы сочли целесообразным их привести. При первом чтении доказательства теорем можно пропустить.

Особо надо сказать о роли ссылок на литературу. Чтобы усвоить материал, представленный в книге, необходимо знать указанные выше стандартные учебные курсы. Доказательства же всех приведенных в учебнике теорем читатель найдет в публикациях, указанных в списках литературы, которые даны в каждой главе. Каждая глава учебника - это только введение в большую область эконометрики, и может появиться вполне естественное желание выйти за пределы учебника. Приведенные литературные списки могут этому помочь.

Автор настоящего учебника - ученый и педагог с тридцатилетним опытом работы в области эконометрики и прикладной статистики. Автор пользуется возможностью выразить признательность за совместную работу своим 170 соавторам по различным публикациям. Познакомиться с современной научной информацией по эконометрике можно на сайтах www.antorlov.chat.ru, www.newtech.ru/~orlov, www.antorlov.euro.ru, www.antorlov.euro.ru, входящих в Интернет. Достаточно большой объем информации содержит еженедельная рассылка "Эконометрика", выпускаемая с июля 2000 г. (автор благодарен А.А.Орлову за компьютерную поддержку настоящего проекта).

По ряду причин исторического характера основное место публикаций научных работ по статистическим методам и прикладной статистике в нашей стране - раздел "Математические методы исследования" журнала "Заводская лаборатория". В разделе публикуются статьи по статистическим методам анализа технических и технико-экономических данных. Автор благодарит главного редактора академику РАН Н.П.Лякишева, зам. главного редактора М.Г.Плотницкую, редактору отдела М.Е.Носову. Автору приятно выразить радость от возможности работать вместе с коллегами по секции "Математические методы исследования" редколлегии журнала, прежде всего с заслуженным деятелем науки РФ проф. В.Г.Горским. Автор искренне благодарен своим учителям - академику АН УССР Б.Г. Гнеденко, проф. В.В. Налимову.

Автор выражает признательность заведующему кафедрой "Экономика и организация производства" факультета "Инженерный бизнес и менеджмент" Московского государственного технического университета им. Н.Э. Баумана проф., докт. эконом. наук С.Г. Фалько за постоянную поддержку проекта по разработке и внедрению эконометрических курсов. Хотелось бы сказать «спасибо» всему коллектива кафедры и факультета в целом, декану В.К.Селюкову и членам Ученого Совета, поддержавшим инициативу о введении эконометрики в учебный процесс МГТУ им. Н.Э.Баумана.

Автор благодарен научному редактору Е.Е.Узловой за большую работу при подготовке рукописи к печати.

В учебнике дано представление об эконометрике, соответствующее общепринятому в мире. Сделана попытка изложить материал на современном уровне научных исследований в области эконометрики. Конечно, возможны различные точки зрения по тем или иным частным вопросам. Автор будет благодарен читателям, если они сообщат свои вопросы, замечания по адресу издательства или по электронной почте E-mail: orlov@professor.ru.

Глава 1. Структура современной эконометрики

Эконометрика – это наука, изучающая конкретные количественные и качественные взаимосвязи экономических объектов и процессов с помощью математических и статистических методов и моделей (Энциклопедический Словарь). Эконометрические методы - это прежде всего методы статистического анализа конкретных экономических данных, естественно, с помощью компьютеров. В нашей стране они пока сравнительно мало известны, хотя именно у нас наиболее мощная научная школа в области основы эконометрики – теории вероятностей. В настоящей главе дается общее представление о структуре и возможностях эконометрики, включая ее последние достижения.

Что дает эконометрика для формирования мышления менеджера и экономиста? Почему необходимо учить будущих экономистов и менеджеров эконометрике? Эти вопросы - центральные для нашего обсуждения.

1.1.Эконометрика сегодня

Статистические (эконометрические) методы используются в зарубежных и отечественных экономических и технико-экономических исследованиях, работах по управлению (менеджменту). Применение прикладной статистики и других статистических методов дает заметный экономический эффект. Например, в США - не менее 20 миллиардов долларов ежегодно только в области статистического контроля качества. В 1988 г. затраты на статистический анализ данных в нашей стране оценивались в 2 миллиарда рублей ежегодно [1]. Согласно расчетам сравнительной стоимости валют на основе потребительских паритетов (см. главу 7), эту величину можно сопоставить с 2 миллиардами долларов США. Следовательно, объем отечественного "рынка статистических и эконометрических услуг" был на порядок меньше, чем в США, что совпадает с оценками и по другим показателям, например, по числу специалистов.

Публикации по новым статистическим методам, по их применениям в технико-экономических исследованиях, в инженерном деле постоянно появляются, например, в журнале "Заводская лаборатория", в секции "Математические методы исследования". Надо назвать также журналы "Автоматика и телемеханика" (издается Институтом проблем управления Российской академии наук), "Экономика и математические методы" (издается Центральным экономико-математическим институтом РАН).

Однако необходимо констатировать, что для большинства менеджеров, экономистов и инженеров эконометрика является экзотикой. Это объясняется тем, что в вузах современным статистическим методам почти не учат. Во всяком случае, по состоянию на 2001 г. каждый квалифицированный специалист в этой области - самоучка.

Этому выводу не мешает то, что в вузовских программах обычно есть два курса, связанных со статистическими методами. Один из них - "Теория вероятностей и математическая статистика". Этот небольшой курс читают специалисты с математических кафедр и успевают дать лишь общее представление об основных понятиях математической статистики. Кроме того, внимание математиков обычно сосредоточено на внутриматематических проблемах, их больше интересует доказательства теорем, а не применение современных статистических методов в задачах экономики и менеджмента. Другой курс - "Статистика" или "Общая теория статистики", входящий в стандартный блок экономических дисциплин. Его читают экономисты, не всегда хорошо подкованные в математике. Фактически он является введением в прикладную статистику и содержит первые начала эконометрических методов (по состоянию на 1900 г.). Учебники по "Общей теории статистики" являются неисчерпаемой копилкой математико-статистических ошибок, они порождают поток публикаций, разоблачающих эти ошибки (см., например, [2]). Ничего удивительного в этом нет - такие учебники писали и пишут высококвалифицированные в своей области экономисты, однако они, как правило, плохо знают математику.

Эконометрика (как учебный предмет) призвана, опираясь на два названных вводных курса, вооружить экономиста, менеджера, инженера современным эконометрическим инструментарием, разработанным за последние 50-70 лет. Не владея эконометрикой, отечественный специалист - менеджер и инженер - оказывается неконкурентоспособным по сравнению с зарубежным. Во многих странах мира - Японии и США, Франции и Швейцарии, Перу и Ботсване и др. - статистическим методам обучают в средней школе, ЮНЕСКО постоянно проводят конференции по вопросам такого обучения [3] . В СССР и СЭВ, а теперь - по плохой традиции - и в России игнорируют этот предмет в средней школе и лишь слегка затрагивают его в высшей. Результат на рынке труда очевиден - снижение конкурентоспособности специалистов.

Обсудим сложившуюся ситуацию, уделив основное внимание статистическим методам в экономических и технико-экономических исследованиях, т.е. эконометрике.

1.2.Эконометрика = экономика + метрика

Сначала необходимо выяснить, что обычно понимают под эконометрикой. Затем обсудим современное состояние эконометрики как научно-практической дисциплины.

Во вводных монографиях по экономической теории, как правило, выделяют в качестве ее разделов макроэкономику, микроэкономику и эконометрику. При этом о макроэкономике и микроэкономике обычно подробно рассказывается в тех же монографиях или в дальнейших учебных пособиях, в то время как об эконометрике узнать что-либо самостоятельно российскому студенту почти невозможно. Лишь в последнее время появились отдельные курсы в нескольких московских экономических вузах и соответствующие учебники, увы, трактующие ее крайне узко.

В одном из наиболее распространенных в России вводном курсе западной экономической теории сказано: "Статистический анализ экономических данных называется эконометрикой, что буквально означает: наука об экономических измерениях" [4, с.25]. Действительно, термин "эконометрика" состоит из двух частей: "эконо-" - от "экономика" и "-метрика" - от "измерение". Эконометрика (в другом русско- и англоязычном варианте названия этой дисциплины - эконометрия) входит в обширное семейство дисциплин, посвященных измерениям и применению статистических методов в различных областях науки и практики. К этому семейству относятся, в частности, биометрика (или биометрия), технометрика, наукометрия, психометрика, хемометрика (наука об измерениях и применении статистических методов в химии). Особняком стоит социометрия - этот термин закрепился за статистическими методами анализа взаимоотношений в малых группах, т.е. за небольшой частью такой дисциплины, как статистический анализ в социологии. Эконометрика, как и другие "метрики", посвящена развитию и применению статистических методов в конкретной области науки и практики - в экономике, прежде всего в теории и практике менеджмента.

В мировой науке эконометрика занимает достойное место. Нобелевские премии по экономике получили эконометрики Ян Тильберген, Рагнар Фриш, Лоуренс Клейн, Трюгве Хаавельмо. В 2000 г. к ним добавились еще двое - Джеймс Хекман и Дэниель Мак-Фадден. Выпускается ряд научных журналов, полностью посвященных эконометрике, в том числе:

Journal of Econometrics (Швеция),
Econometric Reviews (США),
Econometrica (США),
Sankhya. Indian Journal of Statistics. Ser.D. Quantitative Economics (Индия),
Publications Econometriques (Франция).

Однако в нашей стране по ряду причин эконометрика не была сформирована как самостоятельное направление научной и практической деятельности, в отличие, например, от Польши, которая стараниями О.Ланге и его коллег покрыта сетью эконометрических "институтов" (в российской терминологии - кафедр вузов). В настоящее время в России начинают развертываться эконометрические исследования, в частности, начинается широкое преподавание этой дисциплины.

Кратко рассмотрим в настоящей главе современную структуру эконометрики. Знакомство с ней необходимо для обоснованных суждений о возможностях применения эконометрических методов и моделей в экономических и технико-экономических исследованиях.

1.3. Структура эконометрики

В эконометрике, как дисциплине на стыке экономики (включая менеджмент) и статистического анализа, естественно выделить три вида научной и прикладной деятельности (по степени специфичности методов, сопряженной с погруженностью в конкретные проблемы):

а) разработка и исследование эконометрических методов (методов прикладной статистики) с учетом специфики экономических данных;

б) разработка и исследование эконометрических моделей в соответствии с конкретными потребностями экономической науки и практики;

в) применение эконометрических методов и моделей для статистического анализа конкретных экономических данных.

Кратко рассмотрим три только что выделенных вида научной и прикладной деятельности. По мере движения от а) к в) сужается широта области применения конкретного эконометрического метода, но при этом повышается его значение для анализа конкретной экономической ситуации. Если работам вида а) соответствуют научные результаты, значимость которых оценивается по общеэконометрическим критериям, то для работ вида в) основное - успешное решение задач конкретной области экономики. Работы вида б) занимают промежуточное положение, поскольку, с одной стороны, теоретическое изучение эконометрических моделей может быть весьма сложным и математизированным (см., например, монографию [5]), с другой -

результаты представляют интерес не для всей экономической науки, а лишь для некоторого направления в ней.

Прикладная статистика - другая область знаний, чем математическая статистика. Это четко проявляется и при преподавании. Курс математической статистики состоит в основном из доказательств теорем, как и соответствующие учебные пособия. В курсах прикладной статистики и эконометрики основное - методология анализа данных и алгоритмы расчетов, а теоремы приводятся как обоснования этих алгоритмов, доказательства же, как правило, опускаются (их можно найти в научной литературе).

Внутренняя структура статистики как науки была выявлена и обоснована при создании в 1990 г. Всесоюзной статистической ассоциации (см., например, статью [6]). Прикладная статистика - методическая дисциплина, являющаяся центром статистики. При применении к конкретным областям знаний и отраслям народного хозяйства получаем научно-практические дисциплины типа "статистика в промышленности", "статистика в медицине" и др. С этой точки зрения эконометрика - это "статистические методы в экономике". Математическая статистика играет роль математического фундамента для прикладной статистики. К настоящему времени очевидно четко выраженное размежевание этих двух научных направлений. Математическая статистика исходит из сформулированных в 1930-50 гг. постановок математических задач, происхождение которых связано с анализом статистических данных. В настоящее время исследования по математической статистике посвящены обобщению и дальнейшему математическому изучению этих задач. Поток новых математических результатов (теорем) не ослабевает, но новые практические рекомендации по обработке статистических данных при этом не появляются. Можно сказать, что математическая статистика как научное направление замкнулась внутри себя. Сам термин "прикладная статистика", используемый с 1960-х годов, возник как реакция на описанную выше тенденцию. Прикладная статистика нацелена на решение реальных задач. Поэтому в ней возникают новые постановки математических задач анализа статистических данных, развиваются и обосновываются новые методы. Обоснование часто проводится математическими методами, т.е. путем доказательства теорем. Большую роль играет методологическая составляющая - как именно ставить задачи, какие предположения принять с целью дальнейшего математического изучения. Велика роль современных информационных технологий, в частности, компьютерного эксперимента.

Рассматриваемое соотношение математической и прикладной статистик отнюдь не являются исключением. Как правило, математические дисциплины проходят в своем развитии ряд этапов. Вначале в какой-либо прикладной области возникает необходимость в применении математических методов и накапливаются соответствующие эмпирические приемы (для геометрии это - "измерение земли" в т.н. Древнем Египте). Затем возникает математическая дисциплина со своей аксиоматикой (для геометрии это - время Евклида). Затем идет внутриматематическое развитие и преподавание (считается, что большинство результатов элементарной геометрии получено учителями гимназий в XIX в.). При этом на запросы исходной прикладной области перестают обращать внимание, и та порождает новые научные дисциплины (сейчас "измерением земли" занимается не геометрия, а геодезия и картография). Затем научный интерес к исходной дисциплине иссякает, но преподавание по традиции продолжается (элементарная геометрия до сих пор изучается в средней школе, хотя трудно понять, в каких практических задачах может понадобиться, например, теорема о том, что высоты треугольника пересекаются в одной точке). Следующий этап - окончательное вытеснение дисциплины из реальной жизни в историю науки (объем преподавания элементарной геометрии в настоящее время постепенно сокращается, в частности, ей все меньше уделяется внимания на вступительных экзаменах в вузах). К интеллектуальным дисциплинам, закончившим свой жизненный путь, относится средневековая схоластика. Как отмечает проф. МГУ им. М.В. Ломоносова В.Н.Тутубалин [7], теория вероятностей и математическая статистика успешно двигаются по ее пути - вслед за элементарной геометрией.

Подведем итог. Хотя статистические данные собираются и анализируются с незапамятных времен (см., например, Книгу Чисел в Ветхом Завете), современная математическая статистика как наука была создана, по общему мнению специалистов, сравнительно недавно - в первой половине XX в. Именно тогда были разработаны основные идеи и получены результаты, излагаемые ныне в учебных курсах математической статистики. После чего специалисты по математической статистике занялись внутриматематическими проблемами, а для теоретического обслуживания проблем практического анализа статистических данных стала формироваться новая дисциплина - прикладная статистика. (Ее центральным печатным органом в нашей стране является упомянутая выше секция "Математические методы исследования" журнала "Заводская лаборатория", где за последние 30 лет опубликовано более 1000 статей по прикладной статистике.)

В настоящее время статистическая обработка данных проводится, как правило, с помощью соответствующих программных продуктов. Разрыв между математической и прикладной статистикой проявляется, в частности, в том, что большинство методов, включенных в статистические пакеты программ (например, в заслуженные Statgraphics и SPSS или в более новую систему Statistica), даже не упоминается в учебниках по математической статистике. В результате специалист по математической статистике оказывается зачастую беспомощным при обработке реальных данных, а пакеты программ применяют (что еще хуже - и разрабатывают) лица, не имеющие необходимой теоретической подготовки. Естественно, что они допускают разнообразные ошибки (напомним, анализ типовых ошибок при применении критериев согласия Колмогорова и омега-квадрат дан в [2]), в том числе в таких ответственных документах, как государственные стандарты по статистическим методам (ниже подробнее рассказано об удручающих результатах анализа этих стандартов; итоги суммированы в статье [8]).

Ситуация с внедрением современных статистических (эконометрических) методов на предприятиях и в организациях различных отраслей народного хозяйства противоречива. К сожалению, при развале отечественной промышленности в 1990-е годы больше всего пострадали структуры, наиболее нуждающиеся в эконометрических методах - службы качества, надежности, центральные заводские лаборатории и др. Однако толчок к развитию получили службы маркетинга и сбыта, сертификации, прогнозирования, инноваций и инвестиций, которым также полезны различные эконометрические методы, в частности, методы экспертных оценок.

1.4. Специфика экономических данных

Для анализа экономических данных могут применяться все разделы прикладной статистики, а именно:

статистика случайных величин;

многомерный статистический анализ;

статистика временных рядов и случайных процессов;

статистика объектов нечисловой природы, в том числе статистика интервальных данных.

Перечисленные четыре области выделены на основе математической природы элементов выборки: в первой из них это - числа, во второй - вектора, в третьей - функции, в четвертой - объекты нечисловой природы, т.е. элементы пространств, в которых нет операций сложения и умножения на число. Примерами объектов нечисловой природы являются значения качественных признаков, бинарные отношения (ранжировки, разбиения, толерантности), последовательности из 0 и 1, множества, нечеткие множества, интервалы, тексты (см. главы 8 и 9 ниже)..

Как и для применений статистических методов в иных областях, в эконометрике решаются задачи описания данных (в том числе усреднения), оценивания, проверки гипотез, восстановления зависимостей, классификации объектов и признаков, прогнозирования, принятия статистических решений и др.

Однако в некоторых отношениях экономические данные отличаются от технических или астрономических, и эти отличия необходимо учитывать при выборе методов анализа конкретных экономических данных.

Многие экономические показатели неотрицательны. Значит, их надо описывать неотрицательными случайными величинами. А вот нормальные распределения принципиально не подходят, поскольку для них вероятность отрицательных значений всегда положительна.

Экономические процессы развиваются во времени, поэтому большое место в эконометрике занимают вопросы анализа и прогнозирования временных рядов, в том числе многомерных. При этом в одних задачах больше внимания уделяют изучению трендов (средних значений, математических ожиданий), например, при анализе динамики цен. В других же - важны отклонения от средней тенденции, например, при применении контрольных карт (карт Шухарта, кумулятивных сумм и др.). Однако в целом спектральный анализ и выделение различных периодов, циклов и типов волн менее распространены, чем, скажем, в биометрике и медицине.

В экономике доля нечисловых данных существенно выше, чем в технике и технологии, соответственно больше применений для статистики объектов нечисловой природы (ниже разберем это утверждение подробнее).

Количество изучаемых объектов в экономическом исследовании часто ограничено в принципе, поэтому обоснование вероятностных моделей в ряде случаев затруднено. Уникальные объекты, например, город Москва, трудно рассматривать как элемент выборки из генеральной совокупности с каким-то определенным распределением, поскольку подобное рассмотрение противоречит здравому смыслу. Вспоминается давняя обложка журнала "Крокодил", на которой изображены два хозяйственника с монетой в руках: "Если упадет орлом, будем строить завод, если решкой - не будем". Подобная рандомизация решений выглядит бессмысленной при принятии ровно одного решения, однако при контроле качества в массовом производстве такой подход оправдан.

Поэтому в эконометрике часто применяются детерминированные методы анализа данных, в отличие от, например, технических наук, в которых обычным является использование вероятностных моделей. Неопределенность приходится описывать не в терминах вероятностно-статистических моделей, а иными способами, например, в терминах теории нечеткости (fuzzy sets theory) или математики и статистики интервальных данных.

Есть два принципиально различных подхода к изучению поведения организаций и людей. Согласно первому из них вполне допустимо описывать действия человека в вероятностных терминах, например, считать его ответ на заданный вопрос случайной величиной. Сторонники второго подхода полагают, что поведение человека или организации является детерминированным, определяется теми или иными причинами, а случайность при анализе выборки возникает лишь из-за случайности при отборе лиц для опроса или предприятий для изучения. Если ответ на вопрос имеет вид "да" - "нет", то число ответов "да" при первом подходе, как известно, имеет биномиальное распределение, а при втором - гипергеометрическое. К счастью для эконометриков, при увеличении объема генеральной совокупности эти два распределения сближаются (если доля выборки в генеральной совокупности мала, например, меньше 10%, то вместо гипергеометрического распределения можно использовать биномиальное), так что при обоих подходах можно применять одни и те же эконометрические методы, не тратя сил на решение философского вопроса о детерминированности или случайности поведения экономического агента- человека или организации.

Итак, специфика эконометрики проявляется не в перечне применяемых для анализа конкретных экономических данных статистических методов, а в частоте использования тех или иных методов.

1.5. Нечисловые экономические величины

В теоретических и практических задачах экономики и менеджмента постоянно используются различные величины, обычно рассматриваемые как числовые. Например, рыночная цена товара, прибыль предприятия, индекс инфляции, валовой внутренний продукт, чистая приведенная величина для потока платежей и т.д. При более тщательном анализе оказывается, что подобные величины не имеют определенного численного значения, они размыты, имеют нечисловой характер, и описывать их следует с помощью нечисловых математических понятий, относящихся к тем или иным классам объектов нечисловой природы, таким, как нечеткие множества, интервалы, распределения вероятностей и др.

Действительно, можно ли считать, что существует рыночная цена на некоторый товар, выраженная числом? Рассмотрим всем привычный товар - хлеб. Для определенности рассмотрим стандартный батон белого хлеба, который стоил 25 копеек в 1990 г. В настоящее время (июнь 2001 г.) в различных торговых точках Москвы его можно купить по ценам от 6 руб. 50 коп. до 7 руб. 30 коп. Сотрудники Института высоких статистических технологий и эконометрики в течение нескольких лет собирала информацию о ценах на 35 продовольственных товаров в 11 "точках" Москвы и Подмосковья (итоги подведены в статье [9]), и максимальная из отмеченных цен превышала минимальную, как правило, на 30-50%. Можно говорить о цене товара при конкретном акте купли-продажи, при покупке в конкретном магазине, но нельзя говорить о конкретном числовом значении рыночной цены товара. Так, говорить о "рыночной цене" конкретной квартиры (не в новостройке) бессмысленно. Цена выявится только в результате соглашения продавца и покупателе при совершении акта купли-продажи. С другой стороны, полностью отказываться от этого укоренившегося в литературе понятия нецелесообразно. Мы предлагаем принять, что рыночная цена - объект нечисловой природы, и описывать ее для стандартного батона белого хлеба, например, в виде интервала [6,50; 7,30] руб.

Анализируя реальные данные, убеждаемся, что интервальный характер имеют рыночные цены на двигатели, черный и цветной металл, сплавы, электроэнергию, нефть, бензин, автоприборы и автомобили, трактора, различные виды приводной техники и другие промышленные товары, точно так же как и на разнообразные услуги. Цены зависят от конкретного договора между поставщиком и потребителем. Часто появляется дополнительный мешающий фактор - инфляция. Так, с сентября 1995 г. по январь 1996 г. доллар США подешевел в нашей стране почти в 2 раза (если сравнивать по покупательной способности в области продовольственных товаров).

Нечисловой характер имеют не только цены. При обсуждении понятия "прибыль предприятия" начнем с очевидной бессмысленности выражения "максимизация прибыли" без указания интервала времени, за который прибыль максимизируется. Только задав интервал времени, можно принять оптимальные решения и рассчитать ожидаемую прибыль. Ясно, что оптимальные решения зависят от интервала планирования. Известная в экономической теории проблема "горизонта планирования" состоит в том, что оптимальное поведение зависит от того, на какое время вперед планируют, а выбор этого горизонта не имеет рационального обоснования. В монографии [5] рассмотрен ряд примеров указанной зависимости и предложено использовать асимптотически оптимальные планы. Дополнительная сложность состоит в том, что будущая прибыль не может быть определена точно, а потому сама должна описываться как объект нечисловой природы. Итак, задача "максимизации прибыли" может приобрести точный смысл, например, лишь как максимизация нечеткой прибыли на нечетком интервале времени. Оптимизация в случае нечетких переменных рассматривалась в литературе (см., например, [10]), однако пока не получила широкого практического внедрения.

Для приведения экономических величин к одному моменту времени (к сопоставимым ценам) используются индексы инфляции, в другой терминологии, дефляторы. Рассчитывают их с помощью тех или иных потребительских корзин. При этом на нечеткость "рыночных цен" товаров накладывается произвол в выборе состава потребительской корзины и объемов потребления. Теоретический анализ этой ситуации привел нобелевского лауреата по экономике В.В.Леонтьева к выводу о принципиальной невозможности сравнения экономических величин, относящихся к различным моментам времени [11]. Возможный выход состоит в задании индекса инфляции в интервальном виде. Так, расчеты по собранным Институтом высоких статистических технологий и эконометрики данным о ценах показывают, что для Москвы индекс инфляции с марта 1991 г. по апрель 1999 г. описывается интервалом [21,5; 24,0] (при использовании деноминированных рублей).

Еще более размыты обобщенные макроэкономические показатели типа "валового внутреннего продукта" (ВВП), особенно при их сравнении по годам и странам. По мнению известного экономиста О.Моргенштерна [12] подобные макроэкономические показатели могут быть определены лишь с точностью 5-10%. Однако, если пользоваться одной и той же методикой расчета, то можно заметить и изменения в 0,1 %. Проблема в том, что сама методика может вызывать сомнения. Например, по применяемой Госкомстатом РФ "системе национальных счетов" банковские услуги составляют 13% ВВП. С точки зрения здравого смысла это -

абсурдно высокая величина. Она объясняется тем, что, например, выдача кредита в 1 миллион рублей рассматривается как услуга стоимостью в 1 миллион рублей, эквивалентная выпечке и продаже 150 000 батонов хлеба. При всей высокой оценке тяжкого труда банковских боссов, клерков и охранников трудозатраты крестьян, мукомолов, пекарей, транспортников и продавцов 150 000 батонов хлеба, очевидно, несоизмеримо выше.

Нечеткость в неявной форме присутствует и в натуральных показателях. Пусть, например, выпущена партия из 1000 автомашин определенной марки. Нечеткость, связанная с этой партией, состоит в неопределенности реального срока службы автомашин, полезных и вредных эффектов от их эксплуатации. Для снятия этих неопределенностей необходимо, в частности, экономически оценить потери от гибели людей в автокатастрофах. Сколько стоит жизнь человека? При всем уважении к оценкам страховых компаний сама постановка этого вопроса вызывает неловкость. Многие этические и религиозные учения исходят из бесценности человеческой жизни. Из-за принципиальной недопустимости выражения стоимости человеческой жизни в денежных единицах не получили распространения, в частности, методы статистического контроля качества, основанные на учете народнохозяйственного ущерба от пропуска дефектных изделий при контроле.

Более подробно рассмотрим проблемы управления инвестиционными процессами. Одна из них - проблема сравнения инвестиционных проектов. С чисто финансовой точки зрения такой проект - это финансовый поток (cash flow), другими словами, поток платежей и поступлений, т.е. последовательность моментов времени, каждому из которых соответствует некоторая величина платежей (для определенности учитываем их со знаком "минус") или поступлений (учитываем со знаком "плюс"). Как оценивать такие потоки в целом, как их сравнивать? Из многих характеристик потоков платежей рассмотрим здесь две - чистую приведенную величину, называемую в отечественных публикациях также чистой текущей стоимостью или чистым дисконтированным доходом (есть и иные названия) и обозначаемую NPV (Net Present Value), и внутреннюю норму доходности, или прибыли IRR (Internal Rate of Return).

При определении NPV, как известно, для приведения величин платежей и поступлений к одному моменту времени используется постоянный дисконт-фактор. В реальности дисконт-фактор не является заранее известной функцией от времени и зависит от динамики как макроэкономических показателей - ставки рефинансирования Центрального банка РФ и индекса инфляции, так и микроэкономических - финансового положения инвестора, кредитной и депозитной ставок конкретного банка и др.. Кроме того, размеры и моменты осуществления платежей и поступлений также могут быть известны лишь с некоторой точностью. Следовательно, как функция от неопределенных (размытых) величин такая характеристика инвестиционного проекта, как NPV, сама является неопределенной. Лишь частично эту неопределенность можно снять, рассматривая NPV как функцию одной независимой переменной - дисконт-фактора. Если все перечисленные неопределенности можно описать интервалами (т.е. задать границы - "от" и "до"), то NPV также описывается интервалом, границы которого можно рассчитать с помощью подходов, развитых в статистике интервальных данных (см. главу 9 ниже). В результате в ряде случаев становится невозможным сделать однозначный выбор при сравнении двух инвестиционных проектов по NPV. Дело в том, что сравнение чисел можно провести всегда, а сравнение интервалов - лишь тогда, когда они не пересекаются. Если же пересекаются - целесообразно заявить об эквивалентности двух рассматриваемых инвестиционных проектов по чистой текущей стоимости NPV.

Внутренняя норма доходности IRR - это значение постоянного дисконт-фактора q, при котором NPV как функция q обращается в 0. К сожалению, как хорошо известно, при "неудачном" распределении поступлений и платежей уравнение NPV(q) = 0 может иметь не одно, а много решений. В литературе указывают и некоторые иные причины, по которым IRR нецелесообразно использовать для сравнения потоков платежей. Кроме того, в случае IRR имеются те же источники неопределенности, что и для NPV - размытость дисконт-фактора, моментов и величин поступлений и платежей. Эта размытость приводит к необходимости рассматривать IRR как интервал, а при непустоте пересечения интервалов, соответствующих двум инвестиционным проектам, сравнение этих проектов сводится к утверждению об их равноценности.

Итак, рассмотренные характеристики инвестиционных проектов NPV и IRR, как и любые иные, имеют неустранимые неопределенности. Игнорировать это объективное обстоятельство, завышать точность экономических расчетов - это значит обманываться самому либо вводить в заблуждение заказчиков расчетов.

Как же поступать при анализе инвестиционных проектов? Рассмотрим два корректных подхода к такому анализу. Во-первых, можно постараться явным образом учесть имеющиеся неопределенности (в том числе перечисленные выше) и применить те или иные способы анализа неопределенных величин, в частности, разработанные в теории нечеткости и в статистике объектов нечисловой природы (см., например, монографии [5,10]). Другими словами, требуется более тщательный экономико-математический анализ ситуации, предполагающий построение соответствующих эконометрических моделей, разработку и/или применение необходимого программного обеспечения. А для этого нужны обученные кадры, время и деньги.

Во-вторых, вместо расчетов можно обратиться к интуиции специалистов, применив современные методы экспертных оценок (см. ниже главу 12), в частности, основанные на сборе оценок экспертами нечисловых экономических величин и их анализе методами статистики объектов нечисловой природы. Для практического использования представляется перспективным оценивание в виде интервалов (частный случай применения теории нечетких множеств) и соответственно их анализ методами статистики интервальных данных. Применение комбинированных подходов, предполагающих использование систем, интегрирующих как эконометрические и экономико-математические модели, так и методы экспертных оценок - пока дело будущего.

1.6. Статистика интервальных данных - научное направление на стыке метрологии и математической статистики

В статистике интервальных данных (СИД) элементами выборки являются не числа, а интервалы, в частности, порожденные наложением ошибок измерения на значения случайных величин. Подробнее этот сравнительно новый, но весьма перспективный раздел эконометрики рассмотрим в главе 9. Здесь дадим лишь общее представление о статистике интервальных данных в сравнении с классической математической статистикой. Прежде всего отметим, что СИД входит в теорию устойчивости (робастности) статистических процедур и примыкает к интервальной математике. В СИД изучены практически все задачи классической прикладной математической статистики, в частности, задачи регрессионного анализа, планирования эксперимента, сравнения альтернатив и принятия решений в условиях интервальной неопределенности и др. Основная идея СИД является общеинженерной - каждая величина должна приводиться вместе с погрешностью ее определения. К сожалению, эта идея еще не стала общеэкономической.

Рассмотрим развитие в течение последних 15 лет асимптотических методов статистического анализа интервальных данных при больших объемах выборок и малых погрешностях измерений. В отличие от классической математической статистики, сначала устремляется к бесконечности объем выборки и только потом - уменьшаются до нуля погрешности. Разработана общая схема исследования, включающая расчет двух основных характеристик - нотны (максимально возможного отклонения статистики, вызванного интервальностью исходных данных) и рационального объема выборки (превышение которого не дает существенного повышения точности оценивания и статистических выводов, связанных с проверкой гипотез). Она применена к оцениванию математического ожидания и дисперсии, медианы и коэффициента вариации, параметров гамма-распределения в ГОСТ 11.011-83 и характеристик аддитивных статистик, для проверки гипотез о параметрах нормального распределения, в т.ч. с помощью критерия Стьюдента, а также гипотезы однородности двух выборок по критерию Смирнова, и т.д.. Разработаны подходы к учету интервальной неопределенности в основных постановках регрессионного, дискриминантного и кластерного анализов.

Многие утверждения СИД отличаются от аналогов из классической математической статистики. В частности, не существует состоятельных оценок: средний квадрат ошибки оценки, как правило, асимптотически равен сумме дисперсии этой оценки, рассчитанной согласно классической теории, и квадрата нотны. Метод моментов иногда оказывается точнее метода максимального правдоподобия (см. ГОСТ 11.011-83). Нецелесообразно с целью повышения точности выводов увеличивать объем выборки сверх некоторого предела. В СИД классические доверительные интервалы должны быть расширены вправо и влево на величину нотны, и длина их не стремится к 0 при росте объема выборки.

СИД позволяет снять некоторые противоречия между метрологией и классической математической статистикой. Например, вторая из названных дисциплин утверждает, что путем увеличения числа измерений можно сколь угодно точно оценить параметр, а первая вполне справедливо оспаривает это утверждение. Результаты СИД уточняют интуитивные представления метрологов (которые сосредотачивались, впрочем, вокруг весьма частного с точки зрения эконометрики вопроса - оценивания математического ожидания) и развенчивают "гордыню" математической статистики.

1.7.Эконометрические модели

Статистические и математические модели экономических явлений и процессов определяются спецификой той или иной области экономических исследований. Так, в экономике качества модели, на которых основаны статистические методы сертификации и управления качеством - модели статистического приемочного контроля, статистического контроля (статистического регулирования) технологических процессов (обычно с помощью контрольных карт Шухарта или кумулятивных контрольных карт), планирования экспериментов, оценки и контроля надежности и другие - используют как технические, так и экономические характеристики, а потому относятся к эконометрике, равно как и многие модели теории массового обслуживания (теории очередей). Экономический эффект только от использования статистического контроля в промышленности США оценивается как 0,8% валового национального продукта (20 миллиардов долларов в год), что существенно больше, чем от любого иного экономико-математического или эконометрического метода.

К эконометрике качества относятся многие публикации научно-технического журнал "Заводская лаборатория (диагностика материалов)". Этот журнал посвящен аналитической химии, физическим, математическим и механическим методам исследования, а также сертификации материалов. Он создан в 1932 г. и адресован специалистам черной и цветной металлургии, химической промышленности и др. Кроме сотрудников центральных заводских лабораторий, служб качества, надежности и других заводских подразделений, он ориентирован в основном на работников прикладных научно-исследовательских организаций. Сейчас журнал базируется в Институте металлургии им.А.А.Байкова Российской академии наук. С 60-х годов в нем действует секция редколлегии "Математические методы исследования", отвечающая за публикацию статей по статистическим методам в промышленности, в частности, в метрологии, диагностике материалов, стандартизации, управлении качеством и сертификации. Технические и экономические вопросы обычно рассматриваются в неразрывном единстве. С рассматриваемой тематикой должен быть знаком каждый специалист по эконометрике, а также по экономике и организации производства.

Ввиду важности статистических методов в стандартизации и управления качеством в СССР с начала 70-х годов разрабатывались государственные стандарты по статистическим методам в рассматриваемой области. По мнению ряда специалистов, из-за неграмотности разработчиков государственные стандарты содержали многочисленные ошибки. Для анализа ситуации в 1985 г. была организована т.н. Рабочая группа по упорядочению системы стандартов по прикладной статистике и другим статистическим методам. В этот научный коллектив входили 66 научных работников и специалистов из различных отраслей народного хозяйства и вузов, в том числе более 20 докторов наук. Оказалось, что существенная часть стандартов по статистическим методам действительно содержала грубые ошибки. Основная часть ошибочных стандартов была отменена, некоторые действуют до сих пор. Затем с целью исправления положения был организован Всесоюзный центр по статистическим методам и информатике (ныне - Институт высоких статистических технологий и эконометрики МГТУ им. Н.Э. Баумана), который разработал около 30 компьютерных систем по современным статистическим методам управления качеством. Наибольшее распространение получила система НАДИС (НАДежность и ИСпытания), созданная под руководством проф. О.И.Тескина (МГТУ им. Н.Э.Баумана). Итоги описанного направления работ подведены в журнале "Заводская лаборатория" в статье [8].

Работы по эконометрическим моделям статистического контроля постоянно публикуются в "Заводской лаборатории". Эти модели мы рассмотрим в главе 13. Рассмотрим здесь только одну конкретную рекомендацию, основанную на сравнении по экономическим показателям различных схем организации контроля и технического обслуживания. Этот подход приводит к принципиальному изменению технико-экономической политики при контроле качества. Он позволяет "снять" парадокс классической теории статистического контроля - чем выше достигнутый уровень качества, тем больше необходимый объем контроля. Предлагаемый выход состоит в переходе к расширению возможностей менеджера при выборе технической политики на основе учета экономических рисков. "Перекладывание" контроля на потребителя может быть экономически выгодно, если производитель организовал защиту от риска методом пополнения партий (путем включение запасных изделий) или путем развития технического обслуживания, позволяющего быстро заменять дефектное изделие.

Другой важный раздел эконометрики - теория и практика экспертных оценок. Экспертные оценки используют для решения ряда экономических задач, например, выбора оптимального направления инвестиций, или наилучшего образца определенного вида продукции для организации массового выпуска, или при прогнозировании развития экономической ситуации, или при распределении финансирования... Следовательно, используемые в теории экспертных оценок модели [являются эконометрическими. Они рассматриваются в главе 12.

Менее полезными практически (с точки зрения достигаемого экономического эффекта), но более известными в теоретических и учебных публикациях являются различные эконометрические модели, предназначенные для прогнозирования макроэкономических показателей. Это обычно модели весьма частного вида, имеющие целью прогнозирование многомерного временного ряда. Они представляют собой систему линейных зависимостей между прошлыми и настоящими значениями переменных. В таких задачах оценивают как структуру модели, т.е. вид зависимости между значениями известных координат вектора в прежние моменты времени и их значениями в прогнозируемый момент (т.е. проводят т.н. идентификацию модели), так и коэффициенты, входящие в эту зависимость. Структура такой модели - объект нечисловой природы, что и объясняет сложность соответствующей теории.

Каждой области экономических исследований, связанной с анализом эмпирических данных, как правило, соответствуют свои эконометрические модели. Например, для моделирования процессов налогообложения с целью оценки результатов применения управляющих воздействий (например, изменения ставок налогов) на процессы налогообложения должен быть разработан комплекс соответствующих эконометрических моделей. Кроме системы уравнений, описывающей динамику системы налогообложения под влиянием общей экономической ситуации, управляющих воздействий и случайных отклонений, необходим блок экспертных оценок. Полезен блок статистического контроля, включающий как методы выборочного контроля правильности уплаты налогов (налогового аудита), так и блок выявления резких отклонений параметров, описывающих работу налоговых служб. Подходам к проблеме математического моделирования процессов налогообложения посвящена монография [13], содержащая также информацию о современных статистических (эконометрических) методах и экономико-математических моделях, в том числе имитационных.

С помощью эконометрических методов следует оценивать различные величины и зависимости, используемые при построении имитационных моделей процессов налогообложения, в частности, функции распределения предприятий по различным параметрам налоговой базы. При анализе потоков платежей необходимо использовать эконометрические модели инфляционных процессов, поскольку без оценки индекса инфляции невозможно вычислить дисконт-функцию, а потому нельзя установить реальное соотношение авансовых и "итоговых" платежей. Прогнозирование сбора налогов может осуществляться с помощью системы временных рядов - на первом этапе по каждому одномерному параметру отдельно, а затем - с помощью некоторой линейной эконометрической системы уравнений, дающей возможность прогнозировать векторный параметр с учетом связей между координатами и лагов, т.е. влияния значений переменных в определенные прошлые моменты времени. Возможно, более полезными окажутся имитационные модели более общего вида, основанные на интенсивном использовании современной вычислительной техники.

1.8. Применения эконометрических методов

Эконометрика не так сильно оторвалась от реальных задач, как математическая статистика, специалисты в области которой зачастую ограничиваются доказательством теорем, не утруждая себя вопросом о том, для решения каких практических задач эти теоремы могут быть нужны. Поэтому эконометрические модели обычно доводятся "до числа", т.е. применяются для обработки конкретных эмпирических данных. Так, эконометрические методы нужны для оценки параметров экономико-математических моделей, например, моделей логистики (в частности, управления запасами [5]).

Приведение к сопоставимым ценам - составная часть любого экономического расчета, связанного более чем с одним моментом времени. Как показали наши наблюдения над ценами, использование публикуемых Госкомстатом РФ значений индексов инфляции приводит к систематическим ошибкам. Так, по нашим данным цены за 5 лет (с декабря 1990 г. по декабрь 1995 г.) выросли в среднем в 9989 раз, а по данным Госкомстата РФ - в 4700 раз. Различие - в 2 раза! Оно сохраняется и в настоящее время. Сказанное определяет актуальность использования независимой информации о ценах и индексах инфляции при анализе экономического положения российских предприятий и граждан России.

В частности, инфляцию необходимо учитывать при анализе результатов финансовой деятельности предприятий и их подразделений за год или более длительные интервалы времени. Постепенно эта простая мысль становится все более близкой специалистам в указанной области, хотя до сих пор в большинстве случаев оперируют номинальными значениями, как будто инфляция полностью отсутствует.

Эконометрические методы следует использовать как составную часть научного инструментария практически любого технико-экономического исследования. Оценка точности и стабильности технологических процессов, разработка адекватных методов статистического приемочного контроля и статистического контроля технологических процессов, оптимизация выхода полезного продукта методами планирования экстремального эксперимента в химико-технологических системах, повышение качества и надежности изделий, сертификация продукции, диагностика материалов, изучение предпочтений потребителей в маркетинговых исследованиях, применение современных методов экспертных оценок в задачах принятия решений, в частности, в стратегическом, инновационном, инвестиционном менеджменте, при прогнозировании - везде полезна эконометрика.

Бесспорно совершенно, что практически любая область экономики и менеджмента имеет дело со статистическим анализом эмпирических данных, а потому имеет те или иные эконометрические методы в своем инструментарии. Например, перспективно применение этих методов для анализа научного потенциала России, при изучении рисков инновационных исследований, в задачах контроллинга [14], при проведении маркетинговых опросов, сравнении инвестиционных проектов, эколого-экономических исследований в области химической безопасности биосферы и уничтожения химического оружия, в задачах страхования, в том числе экологического, при разработке стратегии производства и продажи специальной техники и во многих других областях.

1.9. Эконометрика как область научно-практической деятельности

Подводя итоги сказанному выше, обратимся к вопросам подготовки кадров в области эконометрики. В настоящее время в классификаторах специальностей научных работников и специальностей, по которым идет подготовка студентов, эконометрика не представлена вообще, а статистика - двумя отдельными позициями: в специальности "теория вероятностей и математическая статистика" как часть математики и как одна из экономических специальностей. Такие практически важные области, как статистические методы в промышленности, в частности, статистические методы управления качеством и надежностью (т.е. обеспечения, повышения качества промышленной продукции) , технической диагностики, планирования эксперимента, а также статистические методы в менеджменте, в экологии, в химии, в геологии, в медицине и т.д., и т.п. вообще не представлены в рассматриваемых классификаторах. Можно сказать, что они существуют нелегально, потому что, например, научным работникам при защите диссертаций приходится "маскироваться" под другие специальности.

Поскольку кадры по статистическим методам и эконометрике не готовятся, то каждый специалист - самоучка, то общее их число на порядок меньше, чем в Великобритании. США и других странах, в которых науки "эконометрика" и "статистика" рассматривается в одном ряду с такими общепризнанными науками, как математикой, физикой, химией, биологией и др.

Разрыв между математической статистикой и статистикой как экономической дисциплиной обернулся тем, что математики "замкнулись в себе", доказывая теоремы на основе постановок 30-50 гг. и почти ничего не давая для анализа реальных данных, а экономисты, хотя и помещают математико-статистические методы в свои книги, но, не зная математики, дают непрекращающийся поток ошибок в учебниках.

Давно стало ясно, что положение в области статистических методов и эконометрики надо менять. В 1985-90 гг. была проведена большая работа по анализу положения дел в области теории и практики статистики и эконометрики в нашей стране. В итоге в октябре 1990 г. создана Всесоюзная статистическая ассоциация (ВСА). Как единое целое ВСА после развала СССР перестала действовать, хотя де-юре продолжает существовать, поскольку решение о роспуске ВСА в соответствии с ее Уставом может принять только съезд ВСА. Такого съезда не было.

В соответствии с реальной структурой статистики ВСА делилась на четыре секции, а именно: 1) практической статистики, 2) статистических методов и их применений, 3) статистики надежности (состояла из работников оборонной промышленности), 4) социально-экономической статистики. Названия секций, зафиксированные в документах ВСА, не вполне соответствуют действительности. Первая секция состояла из работников Госкомстата, большинство членов второй и третьей занимаются научной и практической деятельностью, в том числе в социально-экономической области (в частности, ведут научные и практические работы по эконометрике), а четвертая состояла из преподавателей статистических дисциплин в рамках экономического образования. Вторая секция (во взаимодействии с третьей) "породила" в 1992 г. Российскую ассоциацию статистических методов, а в 1996 г. - Российскую академию статистических методов. В настоящее время эти структуры занимаются в основном поддержкой проведения научных исследований и публикацией их результатов.

По ряду исторических причин отечественная статистика расколота на кланы, практически не взаимодействующие друг с другом. Создание ВСА преследовало, в частности, цель налаживания контактов между секциями 1 и 4, с одной стороны, и секциями 2 и 3, с другой. К сожалению, в обстановке, наступившей с 1992 г., было не до перестройки теории статистики, ее применений и преподавания.

Однако необходимость налаживания контактов не отпала. Так, вряд ли можно считать допустимой ситуацию, когда практически в каждом учебнике по "общей теории статистики" даются абсолютно неверные рекомендации по применению критерия Колмогорова, используемого для проверки согласия эмпирического распределения с теоретическим [2]. Очевидно, необходимы постоянные контакты между специалистами по социально-экономическим применениям статистических методов, с одной стороны, и математической статистике, с другой стороны. Эконометрика находится именно на этом стыке.

1.10. Эконометрические методы в практической и учебной деятельности

Компьютер на рабочем месте менеджера, экономиста, инженера - уже реальность. Практическое применение эконометрических методов обычно осуществляется с помощью диалоговых систем, соответствующих решаемым экономическим и технико-экономическим задачам. Для конкретных наборов задач таких систем разработано уже много, некоторые перечислены в статье [8]. Создание подобных систем должны быть продолжено. Так, для налоговых служб должны быть подготовлены соответствующие оригинальные системы на базе действующих автоматизированных информационных систем (АИС).

Однако для того, чтобы грамотно применять компьютерную систему, надо иметь некоторые предварительные знания по эконометрике. В отсутствии подобных знаний у подавляющего большинства российских экономистов и инженеров, в том числе у менеджеров - директоров предприятий, государственных служащих, а также, например, у работников налоговых органов, - основная проблема. Лицо, ничего не знающее об эконометрике, не в состоянии понять, что эта научно-практическая дисциплина может помочь решить проблемы его организации, а потому ему и в голову не приходит пригласить бригаду эконометриков к сотрудничеству.

Эта проблема наглядно выявилась в ходе работ Всесоюзного центра статистических методов и информатики (ныне - Институт высоких статистических технологий и эконометрики МГТУ им. Н.Э. Баумана). Центром был разработан широкий спектр программных систем по эконометрике. Однако число их продаж было явно неадекватно проведенным оценкам емкости рынка, т.е. числу предприятий, которым были бы полезны эти системы. Это объяснялось попросту отсутствием на подавляющем числе предприятий специалистов, знакомых с эконометрическими методами хотя бы на том элементарном уровне, который позволяет понять, что им такие системы нужны. Например, нужны для того, чтобы обоснованно анализировать и выбирать планы статистического приемочного контроля, что необходимо делать практически на любом предприятии, независимо от отрасли и форм собственности. В любом договоре на поставку есть раздел "Правила приемки и методы контроля", и подготовлен он обычно отнюдь не на современном уровне. Если же на предприятии были квалифицированные специалисты, то они стремились расширить свой инструментарий за счет программных систем по эконометрике Всесоюзного центра статистических методов и информатики.

Поэтому надо широко преподавать эконометрику. Без этого разработанные для нужд организаций и предприятий имитационные компьютерные модели на основе эконометрических методов останутся омертвленным капиталом, не будут грамотно использоваться.

Но не следует сосредотачиваться лишь на подготовке специалистов по разработке эконометрических методов, умеющих доказывать теоремы и писать программы. Прежде всего нужны пользователи, понимающие, для решения каких задач годится тот или иной эконометрический метод, какая нужна исходная информация, как интерпретировать выдаваемые компьютером результаты.

Современное обучение эконометрическим методам возможно лишь при использовании компьютерных систем статистического анализа, включающих, в частности, методы статистики объектов нечисловой природы и другие идеи последних десятилетий. Большой интерес у студентов вызывает использование конкретных эконометрических данных, например, таких: на июнь 2001 г. индекс инфляции составил, по нашим данным, более 42,5 (по сравнению с декабрем 1990 г.), следовательно, средняя начисленная зарплата по стране (2260 руб. в месяц) в ценах декабря 1990 г. равна 2260/42,5 = 53 руб.18 коп., т.е. за 10,5 лет уменьшилась в 5,6 раз (в декабре 1990 г. средняя зарплата составляла 297 руб.). Прежняя минимальная зарплата в 70 руб. (декабрь 1990 г.) при индексации соответствует примерно 3000 руб., т.е. заметно больше средней зарплаты июня 2001 г. Эконометрическому анализу инфляции посвящена глава 7 ниже.

Эконометрические методы - эффективный инструмент в работе менеджера и инженера, занимающегося конкретными проблемами, и задача высшей школы - дать его в руки выпускников экономических и технических специальностей. Кроме теоретических знаний, менеджеры и инженеры должны иметь практические инструменты - сделанные на основе современных достижений эконометрической науки компьютерные системы, предназначенные для анализа статистических данных и построения эконометрических моделей конкретных экономических и технико-экономических явлений и процессов.

Подведем некоторые итоги. В настоящей главе продемонстрирована необходимость обучения эконометрическим методам будущих менеджеров, экономистов, инженеров . Рассмотрено место курса эконометрики в системе высшего технического образования: опираясь на курсы "Теория вероятностей и математическая статистика" и "Статистика", он призван довести знания студентов до уровня современности. Указаны связи курса эконометрики со многими иными учебными предметами - менеджментом, маркетингом, экологией, стандартизацией, метрологией и управлением качеством, инвестиционной, инновационной, контрольной и контроллинговой деятельностью, оценкой финансового состояния предприятия, прогнозированием и технико-экономическим планированием, экономико-математическим моделированием производственных систем и др.

Эконометрика - эффективный инструмент научного анализа и моделирования в руках квалифицированного менеджера, экономиста, инженера.

Цитированная литература

1. Комаров Д.М., Орлов А.И. Роль методологических исследований в разработке методоориентированных экспертных систем (на примере оптимизационных и статистических методов). - В сб.: Вопросы применения экспертных систем. - Минск: Центросистем, 1988. С.151-160.
2. Орлов А.И. Распространенная ошибка при использовании критериев Колмогорова и омега-квадрат. - Журнал " Заводская лаборатория". 1985. Т.51. №1. С.60-62.
3. The teaching of statistics / Studies in mathematical education, vol.7. - Paris, UNESCO, 1991. - 258 pp.
4. Долан Э.Дж., Линдсей Д.Е. Рынок: микроэкономическая модель. - СПб: СП "Автокомп", 1992. - 496 с.
5. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979. - 296 с.
6. Орлов А.И. О перестройке статистической науки и её применений. - Журнал "Вестник статистики". 1990. No.1. С.65 - 71.
7. Тутубалин В.Н. Границы применимости (вероятностно-статистические методы и их возможности). - М.: Знание, 1977. - 64 с.
8. Орлов А.И. Сертификация и статистические методы. - Журнал "Заводская лаборатория". 1997. Т.63. № 3. С.55-62.
9. Орлов А.И., Жихарев В.Н., Цупин В.А., Балашов В.В. Как оценивать уровень жизни? (На примере московского региона). – Журнал «Обозреватель-Observer». 1999. No.5 (112). С. 80-83.
10. Орлов А.И. Задачи оптимизации и нечеткие переменные. - М.: Знание, 1980. -64 с.
11. Леонтьев В. Экономические эссе. Теория, исследования, факты и политика: Пер. с англ. - М.: Политиздат, 1990. - 415 с.
12. Моргенштерн О. О точности экономико-статистических наблюдений. - М.: Статистика, 1968. - 324 с.
13. Математическое моделирование процессов налогообложения (подходы к проблеме) / Коллективная монография под ред. Н.Ю.Ивановой, А.И.Орлова и др. - М.: ЦЭО Минобразования РФ, 1997. - 232 с.
14. Карминский А.М., Оленев Н.И., Примак А.Г., Фалько С.Г. Контроллинг в бизнесе. Методологические и практические основы построения контроллинга в организациях. - М.: Финансы и статистика, 1998. - 256 с.

Глава 2. Выборочные исследования

Термин "выборочные исследования" применяют, когда невозможно изучить все единицы представляющей интерес совокупности. Приходится знакомиться с частью совокупности - с выборкой, а затем с помощью эконометрических методов и моделей переносить выводы с выборки на всю совокупность. В качестве примера рассмотрим выборочные исследования предпочтений потребителей, которые часто проводят специалисты по маркетингу.

2.1. Построение выборочной функции спроса

Функция спроса часто встречается в экономических учебниках, но при этом обычно не рассказывается, как она получена. Между тем оценить ее по эмпирическим данным не так уж трудно. Мы часто выясняем ожидаемый спрос с помощью следующего простого приема - спрашиваем потенциальных потребителей: "Какую максимальную цену Вы заплатили бы за такой-то товар?" Пусть для определенности речь идет о конкретном учебном пособии по менеджменту. В одном из экспериментов выборка состояла из 20 опрошенных. Они назвали следующие максимально допустимые для них цены (в рублях по состоянию на сентябрь 1998 г.):

40, 25, 30, 50, 35, 20, 50, 32, 15, 40, 20, 40, 45, 30, 50, 25, 35, 20, 35, 40.

Первым делом названные величины надо упорядочить в порядке возрастания. Результаты представлены в табл.1. В первом столбце - номера различных численных значений (в порядке возрастания), названных потребителями. Во втором столбце приведены сами значения цены, названные ими. В третьем столбце указано, сколько раз названо то или иное значение.

Табл.1. Эмпирическая оценка функции спроса и ее использование

№ п/п (i)	Цена p_i	N_i	Спрос $D(p_i)$	Прибыль (p-10)D(p)	Прибыль (p-15)D(p)	Прибыль (p-25)D(p)
1	**15**	**1**	**20**	**100**	**0**	**-**
2	**20**	**3**	**19**	**190**	**95**	**-**
3	**25**	**2**	**16**	**240**	**160**	**0**
4	**30**	**2**	**14**	**280**	**210**	**70**
5	**32**	**1**	**12**	**264**	**204**	**84**
6	**35**	**3**	**11**	**275**	**220**	**110**
7	**40**	**4**	**8**	**240**	**200**	**120**
8	**45**	**1**	**4**	**140**	**120**	**80**
9	**50**	**3**	**3**	**120**	**105**	**75**

Таким образом, 20 потребителей назвали 9 конкретных значений цены (максимально допустимых, или приемлемых для них значений), каждое из значений, как видно из третьего столбца, названо от 1 до 4 раз. Теперь легко построить выборочную функцию спроса в зависимости от цены. Она будет представлена в четвертом столбце, который заполним снизу вверх. Если мы будем предлагать товар по цене свыше 50 руб., то его не купит никто из опрошенных. При цене 50 руб. появляются 3 покупателя. Записываем 3 в четвертый столбец в девятую строку. А если цену понизить до 45? Тогда товар купят четверо – тот единственный, для кого максимально возможная цена - 45, и те трое, кто был согласен на большую цену – 50 руб. Таким образом, легко заполнить столбец 4, действуя по правилу: значение в клетке четвертого столбца равно сумме значений в находящейся слева клетке третьего столбца и в лежащей снизу клетке четвертого столбца. Например, за 30 руб. купят товар 14 человек, а за 20 руб. - 19.

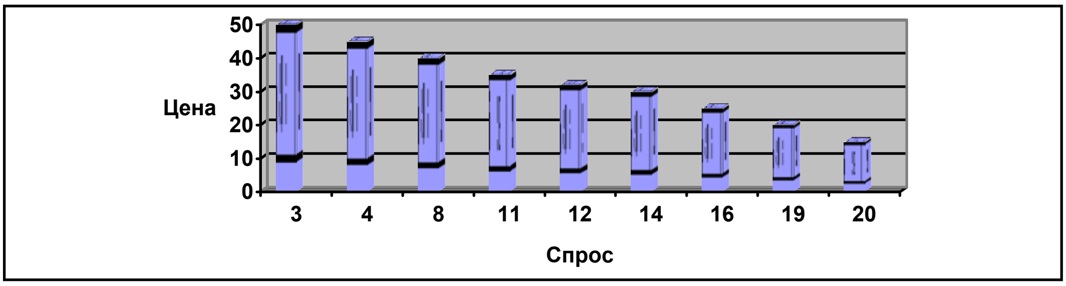
Зависимость спроса от цены - это зависимость четвертого столбца от второго. Табл.1 дает нам девять точек такой зависимости. Зависимость можно представить на рисунке, в координатах «спрос – цена». Если абсцисса - это спрос, а ордината - цена, то девять точек на кривой спроса, перечисленные в порядке возрастания абсциссы, имеют вид:

(3; 50), (4; 45), (8; 40), (11; 35), (12; 32), (14; 30), (16; 25), (19; 20), (20; 15).

Эти девять точек можно использовать для построения кривой спроса каким-либо графическим или расчетным способом, например, методом наименьших квадратов (см. ниже главу 5). Кривая спроса, как и должно быть согласно учебникам экономической теории, убывает, имея направления от левого верхнего угла чертежа к правому. Однако заметны отклонения от гладкого вида функции, связанные, в частности, с естественным пристрастием потребителей к круглым числам. Заметьте, все опрошенные, кроме одного, назвали числа, кратные 5 руб.

Данные табл.1 могут быть использованы для выбора цены продавцом-

Рис.1. Кривая спроса



монополистом (или действующем на рынке монополистической конкуренции). Пусть расходы на изготовление единицы товара равны 10 руб. (например, оптовая цена книги - 10 руб.). По какой цене ее продавать на том рынке, функцию спроса для которого мы только что нашли? Для ответа на этот вопрос вычислим суммарную прибыль, т.е. произведение прибыли на одном экземпляре (р-10) на число проданных (точнее, запрошенных) экземпляров D(p). Результаты приведены в пятом столбце табл.1. Максимальная прибыль, равная 280 руб., достигается при цене 30 руб. за экземпляр. При этом из 20 потенциальных покупателей окажутся в состоянии заплатить за книгу 14, т.е. 70% .

Если же удельные издержки производства, приходящиеся на одну книгу (или оптовая цена), повысятся до 15 руб., то данные столбца 6 табл.1 показывают, что максимальная прибыль, равная 220 руб. (она, разумеется, меньше, чем в предыдущем случае), достигается при более высокой цене - 35 руб. Эта цена доступна 11 потенциальным покупателям, т.е. 55% от всех возможных покупателей. При дальнейшем повышении издержек, скажем, до 25 руб., как вытекает из данных столбца 7 табл.1, максимальная прибыль, равная 120 руб., достигается при цене 40 руб. за единицу товара, что доступно 8 лицам, т.е. 40% покупателей. Отметьте, что при повышении оптовой цены на 10 руб. оказалось выгодным увеличить розничную лишь на 5, поскольку более резкое повышение привело бы к такому сокращению спроса, которое перекрыло бы эффект от повышения удельной прибыли (т.е. прибыли, приходящейся на одну проданную книгу).

Представляет интерес анализ оптимального объема выпуска при различных значениях удельных издержек (табл.2).

Табл.2. Прибыль при различных значениях издержек

№ п/п (i)	Цена p_i	Спрос $D(p_i)$	Прибыль (p-5)D(p)	Прибыль (p-20)D(p)	Прибыль (p-30)D(p)	Прибыль (p-35)D(p)	Прибыль (p-40)D(p)
1	**15**	**20**	**200**	**-**	**-**	**-**	**-**
2	**20**	**19**	**285**	**0**	**-**	**-**	**-**
3	**25**	**16**	**320**	**80**	**-**	**-**	**-**
4	**30**	**14**	**350 ***	**140**	**0**	**-**	**-**
5	**32**	**12**	**324**	**144**	**24**	**-**	**-**
6	**35**	**11**	**330**	**165 ***	**55**	**0**	**-**
7	**40**	**8**	**280**	**160**	**80 ***	**40**	**0**
8	**45**	**4**	**160**	**100**	**60**	**40**	**20**
9	**50**	**3**	**135**	**90**	**60**	**45 ***	**30 ***

В табл.2 звездочками указаны максимальные значения прибыли при том или ином значении издержек, не включенном в табл.1. Для легкости обозрения результаты об оптимальных объемах выпуска и соответствующих ценах из табл. 1 и 2 приведены в табл.3.

Табл.3. Зависимость оптимального выпуска и цены от издержек

Издержки	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**
Оптимальный выпуск	**14**	**14**	**11**	**11**	**8**	**8**	**3**	**3**
Цена	**30**	**30**	**35**	**35**	**40**	**40**	**50**	**50**

Как видно из табл.3, с ростом издержек оптимальный выпуск падает, а цена растет. При этом изменение издержек на 5 единиц может вызывать, а может и не вызывать повышения цены. В этом проявляется микроструктура функции спроса – небольшое повышение цены может привести к тому, что значительные группы покупателей откажутся от покупок, и прибыль упадет.

Этот эффект напоминает известное в экономической теории разделение налогового бремени между производителем и потребителем. Неверно говорить, что производитель перекладывает издержки или, конкретно, налоги, на потребителя, повышая цену на их величину, поскольку при этом сокращается спрос (и выпуск), а потому и прибыль производителя.

Дальнейшее ясно - если оптовая цена будет повышаться, то и дающая максимальную прибыль розничная цена также будет повышаться, и все меньшая доля покупателей сможет приобрести товар. Крайняя точка - оптовая цена, равная 45 руб. Тогда только трое (15 %) купят товар за 50 руб., а прибыль продавца составит только 15 руб. Наглядно видно, что повышение издержек производства приводит к ориентации производителя на наиболее богатые слои населения, но и повышение цен (до оптимального для монополиста-производителя уровня) не приводит к повышению прибыли, напротив, она снижается, и при этом большинство потенциальных потребителей не в состоянии купить товар. Таково влияние инфляции издержек на экономическую жизнь. (Об инфляции мы подробнее поговорим позже.)

Отметим, что рыночные структуры не в состоянии обеспечить всех желающих – это просто не выгодно. Так, из 20 опрошенных лишь 14, т.е. 70%, могут рассчитывать на покупку, даже при минимальных издержках и ценах. Если общество желает чем-либо обеспечить всех граждан, оно должно раздавать это благо бесплатно, как это делается, например, с учебниками в школах.

2.2. Маркетинговые опросы потребителей

Потенциального покупателя интересует не только цена, но и качество товара, красота упаковки (например, для подарочных наборов конфет) и многое другое. Хочешь узнать, чего желает потребитель - спроси его. Эта простая мысль объясняет популярность маркетинговых опросов.

Бесспорно, что основная цель производственной и торговой деятельности - удовлетворение потребностей людей. Как получить представление об этих потребностях? Очевидно, необходимо опросить потребителей. В американском учебнике по рекламному делу [1] подробно рассматриваются различные методы опроса потребителей и обработки результатов с помощью методов эконометрики. Расскажем о результатах опроса потребителей растворимого кофе. Исследование проведено Институтом высоких статистических технологий и эконометрики по заказу АОЗТ "Д-2" в апреле 1994 г. в Москве.

Сбор данных. Обсудим постановку задачи. Заказчика интересуют предпочтения как продавцов кофе (розничных и мелкооптовых), так и непосредственно потребителей. В результате совместного обсуждения было признано целесообразным использовать для опроса и тех, и других одну и ту же анкету из 14 основных и 4 социально-демографических вопросов с добавлением двух вопросов специально для продавцов. Анкета была разработана совместно представителями заказчика и исполнителя и утверждена заказчиком. В табл.4 приведен несколько сокращенный вариант этой анкеты.

Табл.4. Анкета для потребителей растворимого кофе

Дорогой потребитель растворимого кофе,

Институт высоких статистических технологий и эконометрики просит Вас ответить на несколько простых вопросов о том, какой кофе Вы любите. Ваши ответы позволят составить объективное представление о вкусах российских любителей кофе и будут способствовать повышению качества этого товара на российском рынке.

1.Часто ли Вы пьете растворимый кофе: иногда, каждый день 1 чашку, 2-3 чашки, больше, чем 3 чашки.

(Здесь и далее подчеркните нужное.)

2. Что Вы цените в кофе: вкус, аромат, крепость, цвет, отсутствие вредных для здоровья веществ, что-либо еще (сообщите нам, что именно).

3. Как часто покупаете кофе: по мере надобности или по возможности?

4. Любите ли Вы бразильский растворимый кофе? Да, нет, не знаю.

5. Какой объем упаковки Вы предпочитаете: в пакетиках, маленькая банка, средняя банка, большая банка, обязательно стеклянная банка, все равно.

6. Где покупаете растворимый кофе: в ларьках, в продуктовых магазинах, в специализированных отделах и магазинах, все равно, где купить, где-либо еще (опишите, пожалуйста).

7. Были ли случаи, когда купленный Вами кофе оказывался низкого качества? Да, нет.

8. Согласны ли Вы, что за высокое и гарантированное качество продукта можно и заплатить несколько дороже? Да, нет.

9. Какой кофе Вы предпочтете купить: банка неизвестного качества за 2000 руб. или продукт того же веса, безопасность которого гарантирована Минздравом России, за 2500 руб.? Первый, второй.

10. Считаете ли Вы нужным, чтобы производитель принял меры для того, чтобы вредные для здоровья вещества, в частности, ионы тяжелых металлов, не проникали из материала упаковки непосредственно в растворимый кофе? Да, нет.

Институт высоких статистических технологий и эконометрики предполагает сравнить потребительские предпочтения различных категорий россиян. Поэтому просим ответить еще на несколько вопросов.

11. Пол: женский, мужской.

12. Возраст: до 20, 20-30, 30-50, более 50.

13. Род занятий: учащийся, работающий, пенсионер, инженер, врач, преподаватель, служащий, менеджер, предприниматель, научный работник, рабочий, др. (пожалуйста, расшифруйте).

14. Вся Ваша семья любит растворимый кофе или же Вы - единственный любитель этого восхитительного напитка современного человека? Вся семья, я один (одна).

15. Согласились бы Вы и в дальнейшем участвовать в опросах потребителей относительно качества различных пищевых продуктов (чай, джем и др.). Если "да", то сообщите свой адрес, телефон, имя и отчество.

Спасибо за Ваше содействие работе по повышению качества продуктов на российском рынке!

Выбор метода опроса. Широко применяются процедуры опроса, когда респонденты (так социологи и маркетологи называют тех, от кого получают информацию, т.е. опрашиваемых) самостоятельно заполняют анкеты (розданные им или полученные по почте), а также личные и телефонные интервью. Из этих процедур нами было выбрано личное интервью по следующим причинам.

Возврат почтовых анкет сравнительно невелик (в данном случае можно было ожидать не более 5-10%), оттянут по времени и искажает структуру совокупности потребителей (наиболее динамичные люди вряд ли найдут время для ответа на подобную анкету). Кроме того, есть проблемы с почтовой связью (постоянное изменение тарифов затрудняет возмещение респондентам почтовых расходов и др.).

Самостоятельное заполнение анкеты, как показали специально проведенные эксперименты, не позволяет получить полные ответы на поставленные вопросы (респондент утомляется или отвлекается, отказывается отвечать на часть вопросов, иногда не понимает их или отвечает не по существу). Некоторые категории респондентов, например, продавцы в киосках, отказываются заполнять анкеты, но готовы устно ответить на вопросы.

Телефонный опрос искажает совокупность потребителей, поскольку наиболее активных индивидуумов трудно застать дома и уговорить ответить на вопросы анкеты. Репрезентативность нарушается также и потому, что на один номер телефона может приходиться различное количество продавцов и потребителей растворимого кофе, а некоторые из них не имеют телефонов вообще. Анкета достаточно длинна, и разговор по домашнему и тем более служебному телефону респондента может быть прекращен досрочно по его инициативе. Иногородних продавцов и потребителей растворимого кофе, приехавших в Москву, по телефону опросить практически невозможно.

Метод личного интервью лишен перечисленных недостатков. Соответствующим образом подготовленный интервьюер, получив согласие на интервью, удерживает внимание собеседника на анкете, добивается получения ответов на все её вопросы, контролируя при этом соответствие ответов реальной позиции респондента. Ясно, что успех интервьюирования зависит от личных качеств и подготовки интервьюера. Однако расходы на получение одной анкеты при использовании этого метода больше, чем для других рассмотренных методов.

Формулировки вопросов. В маркетинговых и социологических опросах используют три типа вопросов - закрытые, открытые и полузакрытые, они же полуоткрытые. При ответе на закрытые вопросы респондент может выбирать лишь из сформулированных составителями анкеты вариантов ответа. В качестве ответа на открытые вопросы респондента просят изложить свое мнение в свободной форме.

Полузакрытые, они же полуоткрытые вопросы занимают промежуточное положение - кроме перечисленных в анкете вариантов, респондент может добавить свои соображения.

В социологических публикациях продолжается дискуссия по поводу "мягких" и "жестких" форм сбора данных, т.е. фактически о том, какого типа вопросы более целесообразно использовать - открытые или закрытые (см., например, статью директора Института социологии РАН В.А. Ядова [2]). Преимущество открытых вопросов состоит в том, что респондент может свободно высказать свое мнение так, как сочтет нужным. Их недостаток - в сложности сопоставления мнений различных респондентов. Для такого сопоставления и получения сводных характеристик организаторы опроса вынуждены сами шифровать ответы на открытые вопросы, применяя разработанную ими схему шифровки. Преимущество закрытых вопросов в том и состоит, что такую шифровку проводит сам респондент. Однако при этом организаторы опроса уподобляются древнегреческому мифическому персонажу Прокрусту. Как известно, Прокруст приглашал путников заночевать у него. Укладывал их на кровать. Если путник был маленького роста, он вытягивал его ноги так, чтобы они доставали до конца кровати. Если же путник оказывался высоким и ноги его торчали - он обрубал их так, чтобы достигнуть стандарта: "рост" путника должен равняться длине кровати. Так и организаторы опроса, применяя закрытые вопросы, заставляют респондента "вытягивать" или "обрубать" свое мнение, чтобы выразить его с помощью приведенных в формулировке вопроса возможных ответов.

Ясно, что для обработки данных по группам и сравнения групп между собой нужны формализованные данные, и фактически речь может идти лишь о том, кто - респондент или маркетолог (социолог, психолог и др.) - будет шифровать ответы. В проекте "Потребители растворимого кофе" практически для всех вопросов варианты ответов можно перечислить заранее, т.е. можно широко использовать закрытые вопросы. В отличие от опросов с вопросами типа: "Одобряете ли Вы идущие в России реформы?", в которых естественно просить респондента расшифровать, что он понимает под "реформами" (открытый вопрос). Поэтому в используемой в описываемом проекте анкете использовались в основном закрытые и полузакрытые вопросы. Как показали результаты обработки, этот подход оказался правильным - лишь в небольшом числе анкет оказались вписаны свои варианты ответов. Вместе с тем демонстрировалось уважение к мнению респондента, не выдвигалось требование обязательного выбора из заданного множества ответов - респондент мог добавить свое, но редко пользовался этой возможностью (не более чем в 5% случаев).

В последнем вопросе анкеты респонденту предлагалось стать постоянным участником опросов о качестве товаров народного потребления. Ряд респондентов откликнулся на это предложение, в результате стало возможным развертывание постоянной сети "экспертов по качеству", подобной аналогичным в США.

Обоснование объема выборки и проведение опроса. Математико-статистические вероятностные модели выборочных маркетинговых и социологических исследований часто опираются на предположение о том, что выборку можно рассматривать как "случайную выборки из конечной совокупности" (см. терминологическое приложение). Типа той, когда из списков избирателей с помощью датчика случайных чисел отбирается необходимое число номеров для формирования жюри присяжных заседателей. В рассматриваемом проекте нельзя обеспечить формирование подобной выборки - не существует реестра потребителей растворимого кофе. Однако в этом и нет необходимости. Поскольку гипергеометрическое распределение хорошо приближается биномиальным, если объем выборки по крайней мере в 10 раз меньше объема всей совокупности (в рассматриваемом случае это так), то правомерно использование биномиальной модели, согласно которой мнение респондента (ответы на вопросы анкеты) рассматривается как случайный вектор, а все такие вектора независимы между собой. Другими словами, можно использовать модель простой случайной выборки. Таким образом, позиция в давней дискуссии в среде специалистов, изучающих поведение человека (маркетологов, социологов, психологов, политологов и др.) о том, есть ли случайность в поведении отдельно взятого человека или же случайность проявляется лишь в отборе выборки из генеральной совокупности, практически не влияет на алгоритмы обработки данных.

В биномиальной модели выборки оценивание характеристик происходит тем точнее, чем объем выборки больше. Часто спрашивают: "Какой объем выборки нужен?" В математической статистике есть методы определения необходимого объема выборки. Они основаны на разных подходах. Либо на задании необходимой точности оценивания параметров. Либо на явной формулировке альтернативных гипотез, между которыми необходимо сделать выбор. Либо на учете погрешностей измерений (методы статистики интервальных данных, см. ниже). Ни один из этих подходов нельзя применить в рассматриваемом случае.

Биномиальная модель выборки. Она применяется для описания ответов на закрытые вопросы, имеющие две подсказки, например, "да" и "нет". Конечно, пары подсказок могут быть иными. Например, "согласен" и "не согласен". Или при опросе потребителей кондитерских товаров первая подсказка может иметь такой вид: "Больше люблю "Марс", чем "Сникерс". А вторая тогда такова: "Больше люблю "Сникерс", чем "Марс".

Пусть объем выборки равен n. Тогда ответы опрашиваемых можно представить как $X_1, X_2, ..., X_n$, где $X_i = 1$, если i-й респондент выбрал первую подсказку, и $X_i = 0$, если i-й респондент выбрал вторую подсказку, $i=1,2,\ldots,n$. В вероятностной модели предполагается, что случайные величины $X_1, X_2, ..., X_n$ независимы и одинаково распределены. Поскольку эти случайные величины принимают два значения, то ситуация описывается одним параметром p - долей выбирающих первую подсказку во всей генеральной совокупности. Тогда

$$P(X_i = 1) = p,\ P(X_i = 0)= 1-p,\ i=1,2,\ldots,n.$$

Пусть $m = X_1 + X_2 +\ldots+X_n$. Оценкой вероятности р является частота $p^*=m/n$. При этом математическое ожидание $M(p^*)$ и дисперсия $D(p^*)$ имеют вид

$$M(p^*) = p,\ D(p^*)= p(1-p)/n.$$

По Закону Больших Чисел (ЗБЧ) теории вероятностей (в данном случае - про теореме Бернулли) частота p^* сходится (т.е. безгранично приближается) к вероятности p при росте объема выборки. Это и означает, что оценивание проводится тем точнее, чем больше объем выборки. Точность оценивания можно указать. Займемся этим.

По теореме Муавра-Лапласа теории вероятностей

$$\lim_{n\to\infty} P\{\frac{m-np}{\sqrt{np(1-p)}} \le x\} = \Phi(x),$$

где $\Phi(x)$ - функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy,$$

где π = 3,1415925…-отношение длины окружности к ее диаметру, e= 2,718281828… - основание натуральных логарифмов. График плотности стандартного нормального распределения

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

очень точно изображен на германской денежной банкноте в 10 немецких марок. Эта банкнота посвящена великому немецкому математику Карлу Гауссу (1777-1855), среди основных работ которого есть относящиеся к нормальному распределению. В настоящее время нет необходимости вычислять функцию стандартного нормального распределения и ее плотность по приведенным выше формулам, поскольку давно составлены подробные таблицы (см., например, [3]), а распространенные программные продукты содержат алгоритмы нахождения этих функций.

С помощью теоремы Муавра-Лапласа могут быть построены доверительные интервалы для неизвестной эконометрику вероятности. Сначала заметим, что из этой теоремы непосредственно следует, что

$$\lim_{n\to\infty} P\{-x \le \frac{m-np}{\sqrt{np(1-p)}} \le x\} = \Phi(x) - \Phi(-x).$$

Поскольку функция стандартного нормального распределения симметрична относительно 0, т.е. $\Phi(x) + \Phi(-x) = 1$, то $\Phi(x) - \Phi(-x) = 2\Phi(x) - 1$.

Зададим доверительную вероятность γ. Пусть $U(\gamma)$ удовлетворяет условию

$$\Phi(U(\gamma)) - \Phi(-U(\gamma)) = \gamma,$$

т.е.

$$U(\gamma) = \Phi^{-1}(\frac{1+\gamma}{2}).$$

Из последнего предельного соотношения следует, что

$$\lim_{n\to\infty} P\{p^* - U(\gamma)\frac{\sqrt{p(1-p)}}{\sqrt{n}} \le p \le p^* + U(\gamma)\frac{\sqrt{p(1-p)}}{\sqrt{n}}\} = \gamma.$$

К сожалению, это соотношение нельзя непосредственно использовать для доверительного оценивания, поскольку верхняя и нижняя границы зависят от неизвестной вероятности. Однако с помощью метода наследования сходимости [4, п.2.4] можно доказать, что
Следовательно, нижняя доверительная граница имеет вид

$$\lim_{n\to\infty} P\{p^* - U(\gamma)\frac{\sqrt{p^*(1-p^*)}}{\sqrt{n}} \le p \le p^* + U(\gamma)\frac{\sqrt{p^*(1-p^*)}}{\sqrt{n}}\} = \gamma.$$

$$p_{нижн} = p^* - U(\gamma)\frac{\sqrt{p^*(1-p^*)}}{\sqrt{n}},$$

в то время как верхняя доверительная граница такова:

$$p_{верх} = p^* + U(\gamma)\frac{\sqrt{p^*(1-p^*)}}{\sqrt{n}}.$$

Наиболее распространенным (в прикладных исследованиях) значением доверительной вероятности является $\gamma = 0{,}95$. Иногда употребляют термин "95% доверительный интервал". Тогда $U(\gamma) = 1{,}96$.

Пример. Пусть *n*=500, *m*=200. Тогда p^* =0,40. Найдем доверительный интервал для $\gamma = 0{,}95$:

$$p_{нижн} = 0{,}40 - 1{,}96\frac{\sqrt{0{,}4\times 0{,}6}}{\sqrt{500}} = 0{,}40 - 0{,}043 = 0{,}357, \quad p_{верх} = 0{,}40 + 0{,}043 = 0{,}443.$$

Таким образом, хотя в достаточно большой выборке 40% респондентов говорят "да", можно утверждать лишь, что во всей генеральной совокупности таких от 35,7% до 44,3% - крайние значения отличаются на 8,6%.

Замечание. С достаточной для практики точностью можно заменить 1,96 на 2.

Удобные для использования в практической работе маркетолога и социолога таблицы точности оценивания разработаны во ВЦИОМ (Всероссийском центре по изучению общественного мнения). Приведем здесь несколько модифицированный вариант одной из них.

Табл.5. Допустимая величина ошибки выборки (в процентах)

Объем группы Доля p*	1000	750	600	400	200	100
Около 10% или 90%	2	3	3	4	5	7
Около 20% или 80%	3	4	4	5	7	9
Около 30% или 70%	4	4	4	6	9	10
Около 40% или 60%	4	4	5	6	8	11
Около 50%	4	4	5	6	8	11

В условиях рассмотренного выше примера надо взять вторую снизу строку. Объема выборки 500 нет в таблице, но есть объемы 400 и 600, которым соответствуют ошибки в 6% и 5% соответственно. Следовательно, в условиях примера целесообразно оценить ошибку как ((5+6)/2)% = 5,5%. Эта величина несколько больше, чем рассчитанная выше (4,3%). С чем связано это различие? Дело в том, что таблица ВЦИОМ связана не с доверительной вероятностью $\gamma = 0{,}95$, а с доверительной вероятностью $\gamma = 0{,}99$, которой соответствует множитель $U(\gamma) = 2{,}58.$ Расчет ошибки по приведенным выше формулам дает 5,65%, что практически совпадает со значением, найденным по табл.5.

Минимальный из обычно используемых объемов выборки n в маркетинговых или социологических исследованиях - 100, максимальный - до 5000 (обычно в исследованиях, охватывающих ряд регионов страны, т.е. фактически разбивающихся на ряд отдельных исследований - как в ряде исследований ВЦИОМ). По данным Института социологии Российской академии наук [5], среднее число анкет в социологическом исследовании не превышает 700. Поскольку стоимость исследования растет по крайней мере как линейная функция объема выборки, а точность повышается как квадратный корень из этого объема, то верхняя граница объема выборки определяется обычно из экономических соображений. Объемы пилотных исследований (т.е. проводящихся впервые, предварительно или как первые в сериях подобных) обычно ниже, чем объемы исследований по обкатанной программе.

Нижняя граница определяется тем, что в минимальной по численности анализируемой подгруппе должно быть несколько десятков человек (не менее 30), поскольку по ответам попавших в эту подгруппу необходимо сделать обоснованные заключения о предпочтениях соответствующей подгруппы в совокупности всех потребителей растворимого кофе. Учитывая деление опрашиваемых на продавцов и покупателей, на мужчин и женщин, на четыре градации по возрасту и восемь - по роду занятий, наличие 5 - 6 подсказок во многих вопросах, приходим к выводу о том, что в рассматриваемом проекте объем выборки должен быть не менее 400 - 500. Вместе с тем существенное превышение этого объема нецелесообразно, поскольку исследование является пилотным.

Поэтому объем выборки был выбран равным 500. Анализ полученных результатов (см. ниже) позволяет утверждать, что в соответствии с целями исследования выборку следует считать репрезентативной.

Организация опроса. Интервьерами работали молодые люди – студенты первого курса экономико-математического факультета Московского государственного института электроники и математики (технического университета) и лицея No.1140, проходившие обучение по экономике, всего 40 человек, имеющих специальную подготовку по изучению рынка и проведению маркетинговых опросов потребителей и продавцов (в объеме 8 часов). Опрос продавцов проводился на рынках г. Москвы, действующих в Лужниках, у Киевского вокзала и в других местах. Опрос покупателей проводился на рынках, в магазинах, на улицах около киосков и ларьков, а также в домашней и служебной обстановке.

Большое внимание уделялось качеству заполнения анкет. Интервьюеры были разбиты на шесть бригад, бригадиры персонально отвечали за качество заполнения анкет. Второй уровень контроля осуществляла специально созданная "группа организации опроса", третий происходил при вводе информации в базу данных. Каждая анкета заверена подписями интервьюера и бригадира, на ней указано место и время интервьюирования. Поэтому необходимо признать высокую достоверность собранных анкет.

Обработка данных. В соответствии с целью исследования основной метод первичной обработки данных - построение частотных таблиц для ответов на отдельные вопросы. Кроме того, проводилось сравнение различных групп потребителей и продавцов, выделенных по социально-демографическим данным, с помощью критериев проверки однородности выборок (см. ниже). При более углубленном анализе применялись различные методы статистики объектов нечисловой природы (более 90 % маркетинговых и социологических данных имеют нечисловую природу [6]). Использовались средства графического представления данных.

Итоги опроса. Итак, по заданию одной из торговых фирм были изучены предпочтения покупателей и мелкооптовых продавцов растворимого кофе. Совместно с представителями заказчика был составлен опросный лист (анкета типа социологической) из 16 основных вопросов и 4 дополнительных, посвященных социально-демографической информации. Опрос проводился в форме интервью с 500 покупателями и продавцами кофе. Места опроса - рынки, лотки, киоски, продуктовые и специализированные магазины. Другими словами, были охвачены все виды мест продаж кофе. Интервью проводили более 40 специально подготовленных (примерно по 8-часовой программе) студентов, разбитых на 7 бригад. После тщательной проверки бригадирами и группой обработки информация была введена в специально созданную базу данных. Затем проводилась разнообразная статистическая обработка, строились таблицы и диаграммы, проверялись статистические гипотезы и т.д. Заключительный этап - осмысление и интерпретация данных, подготовка итогового отчета и предложений для заказчиков.

Технология организации и проведения маркетинговых опросов лишь незначительно отличается от технологии социологических опросов, многократно описанной в литературе. Так, мы предпочли использовать полуоткрытые вопросы, в которых для опрашиваемого дан перечень подсказок, а при желании он может высказать свое мнение в свободной форме. Не уложившихся в подсказки оказалось около 5 % , их мнения были внесены в базу данных и анализировались дополнительно. Для повышения надежности опроса о наиболее важных с точки зрения маркетинга моментах спрашивалось в нескольких вопросах. Были вопросы - ловушки, с помощью которых контролировалась "осмысленность" заполнения анкеты. Например, в вопросе: "Что Вы цените в кофе: вкус, аромат, крепость, наличие пенки..." ловушкой является включение "крепости" - ясно, что крепость зависит не от кофе самого по себе, а от его количества в чашке. В ловушку никто из 500 не попался - никто не отметил "крепость". Этот факт свидетельствует о надежности выводов проведенного опроса. Мы считали нецелесообразным задавать вопрос об уровне доходов (поскольку в большинстве случаев отвечают "средний", что невозможно связать с определенной величиной). Вместо такого вопроса мы спрашивали: "Как часто Вы покупаете кофе: по мере надобности или по возможности?". Поскольку кофе не является дефицитным товаром, первый ответ свидетельствовал о наличии достаточных денежных средств, второй - об их ограниченности (потребитель не всегда имел возможность позволить себе купить банку растворимого кофе).

Стоимость подобных исследований - 5-10 долларов США на одного обследованного. При этом трудоемкость (и стоимость) начальной стадии - подготовки анкеты и интервьюеров, пробный опрос и др. - 30 % от стоимости исследования, стоимость непосредственно опроса - тоже 30 %, ввод информации в компьютер и проведение расчетов, построение таблиц и графиков - 20 %, интерпретация результатов, подготовка итогового отчета и предложений для заказчиков - 20 % . Таким образом, стоимость собственно опроса в два с лишним раза меньше стоимости остальных стадий исследования. И в выполнении работы участвуют различные специалисты. На первой стадии – в основном нужны высококвалифицированные аналитики. На второй – многочисленные интервьюеры, в роли которых могут выступать студенты и школьники, прошедшие конкретный курс обучения в 8-10 часов. На третьей – работа с компьютером (надо уметь строить и обсчитывать электронные таблицы или базы данных, использовать статистические пакеты, составлять и печатать таблицы и диаграммы и т.п.). На четвертой – опять в основном нужны высококвалифицированные аналитики.

Приведем некоторые из полученных результатов.

а) В отличие от западных потребителей, отечественные не отдавали предпочтения стеклянным банкам по сравнению с жестяными. Поскольку жестяные банки дешевле стеклянных, то можно было порекомендовать (в 1994 г., когда проходил опрос) с целью снижения расходов закупку кофе в жестяных банках.

б) Отечественные потребители готовы платить на 10-20% больше за экологически безопасный кофе более высокого качества, имеющий сертификат Минздрава и символ экологической безопасности на упаковке.

в) Средний объем потребления растворимого кофе - 850 г в месяц (на семью потребителя).

г) Потребители растворимого кофе делятся на классы. Есть "продвинутые" потребители, обращающие большое внимание на качество и экологическую безопасность, марку и страну производства, терпимо относящиеся к изменению цены. Эти "тонкие ценители" - в основном женщины от 30 до 50 лет, служащие, менеджеры, научные работники, преподаватели, врачи (т.е. лица с высшим образованием), пьющие кофе как дома, так и на работе, причем "кофейный ритуал" зачастую входит в процедуру деловых переговоров или совещаний. Противоположный по потребительскому поведению класс состоит из мужчин двух крайних возрастных групп - школьников и пенсионеров. Для них важна только цена, что очевидным образом объясняется недостатком денег.

Результаты были использованы заказчиком в рекламной кампании. В частности, обращалось внимание на сертификат Минздрава и на экологическую безопасность упаковки.

Приведем пример еще одной анкеты из нашего опыта, предназначенной для изучения спроса на образовательные услуги (табл.6).

Табл.6. Исследование рынка образовательных услуг

ИССЛЕДОВАНИЕ РЫНКА ОБРАЗОВАТЕЛЬНЫХ УСЛУГ

Анкета студентов первого курса экономико-математического факультета МГИЭМ(ту).

А. Объективные данные

1. Группа
2. Пол
3. Год рождения
4. Женат(замужем) - да/нет

Б. Общее изучение рынка

5. Почему Вы выбрали специальность экономиста?
6. Почему Вы выбрали именно МГИЭМ(ту) среди всех вузов Москвы, готовящих экономистов?
7. Как Вы представляете себе будущую деятельность по окончании МГИЭМ(ту)?
8. Есть ли у Вас надежда на то, что приобретаемые сейчас знания окажутся полезными в практической работе? Если нет, то зачем Вы учитесь?

В. Отношение к платному образованию

9. Если бы обучение в МГИЭМ(ту) было платным (порядка 1 миллиона руб. в год в ценах февраля 1994 г.), стали бы Вы поступать в МГИЭМ(ту)?

10. Если обучение в МГИЭМ(ту) станет платным, то останетесь ли Вы учиться в МГИЭМ(ту)? (Например, организация оплаты за учебу такова: некоторая фирма заключает контракт со студентом и оплачивает его учебу; студент самостоятельно ищет такую фирму.)

11. Представляет ли для Вас интерес возможность параллельно с дипломом МГИЭМ(ту) получить диплом бакалавра Межкультурного открытого университета (штаб-квартира в Нидерландах) по специальности "бизнес администрейшн" (обучение заочное, стоимость 1780 долларов США за курс)?

Г. О курсе "Основы экономики"

12. Нужно ли рассказывать содержание реферата-дайджеста учебника К. Макконнелла и С. Брю "Экономикс: Принципы, проблемы и политика" или считать его общеизвестным и говорить о том, чего в нём нет?

13. Полезен ли электронный учебник? Если нет, то почему?

14. Нужны ли Вам индивидуальные занятия в аудитории (а не в компьютерном классе с электронным учебником) и в каком виде?

15. Какие темы Вы считаете полезным рассмотреть дополнительно?

16. Сформулируйте иные Ваши замечания и предложения по курсу "Основы экономики": по лекциям, практическим и индивидуальным занятиям.

Д. Дополнительная информация

17. Какие предметы обучения - самые трудные, какие - самые легкие на первом семестре?

18. Подрабатываете ли Вы? Если согласны, укажите примерную (среднюю) сумму в месяц.

19. Существенна ли для Вас стипендия?

20. Есть ли у Вас дома компьютер?

21. Участвуете ли Вы в каких-либо политических движениях, партиях? Если согласны, назовите.

2.3. Проверка однородности двух биномиальных выборок

Как сравнить две группы - мужчин и женщин, молодых и пожилых, и т.п.? В маркетинге это важно для сегментации рынка. Если две группы не отличаются по ответам, значит, их можно объединить в один сегмент и проводить по отношению к ним одну и туже маркетинговую политику, в частности, осуществлять одни и те же рекламные воздействия. Если же две группы различаются, то и относиться к ним надо по-разному. Это - представители двух разных сегментов рынка, требующих разного подхода при борьбе за их завоевание.

Эконометрическая постановка такова. Рассматривается вопрос с двумя возможными ответами, например, "да" и "нет". В первой группе из n_1 опрошенных m_1 человек сказали "да", а во второй группе из n_2 опрошенных m_2 сказали "да". В вероятностной модели предполагается, что m_1 и m_2 - биномиальные случайные величины $B(n_1, p_1)$ и $B(n_2, p_2)$ соответственно. (Запись $B(n, p)$ означает, что случайная величина m, имеющая биномиальное распределение $B(n, p)$ с параметрами n - объем выборки и p - вероятность определенного ответа (скажем, ответа "да"), может быть представлена в виде $m = X_1 + X_2 + ... + X_n$, где случайные величины $X_1, X_2, ..., X_n$ независимы, одинаково распределены, принимают два значения1 и 0, причем $P(X_i = 1) = p$, $P(X_i = 0) = 1-p$, $i=1,2,...,n$.)

Однородность двух групп означает, что соответствующие им вероятности равны, неоднородность - что эти вероятности отличаются. В терминах математической статистики: необходимо проверить гипотезу однородности

$$H_0 : p_1 = p_2$$

при альтернативной гипотезе

$$H_1 : p_1 \neq p_2 .$$

(Иногда представляют интерес односторонние альтернативные гипотезы $H_1^{'} : p_1 > p_2$ и $H_1^{''} : p_1 < p_2$.)

Оценкой вероятности p_1 является частота $p_1^*=m_1/n_1$, а оценкой вероятности p_2 является частота $p_2^*=m_2/n_2$. Даже при совпадении вероятностей p_1 и p_2 частоты, как правило, различаются, как говорят, "по чисто случайным причинам". Рассмотрим случайную величину $p_1^* - p_2^*$. Тогда

$$M(p_1^* - p_2^*) = p_1 - p_2,\ D(p_1^* - p_2^*) = p_1 (1 - p_1)/ n_1 + p_2 (1-p_2)/ n_2 .$$

Из теоремы Муавра-Лапласа и теоремы о наследовании сходимости [4, п.2.4] следует, что

$$\lim_{n_1\to\infty, n_2\to\infty} P\{\frac{p_1^* - p_2^* - M(p_1^* - p_2^*)}{\sqrt{D(p_1^* - p_2^*)}} \le x\} = \Phi(x),$$

где $\Phi(x)$ - функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. Для практического применения этого соотношения следует заменить неизвестную эконометрику дисперсию разности частот на оценку этой дисперсии:

$$D^*(p_1^* - p_2^*) = p^*_1 (1 - p^*_1)/ n_1 + p^*_2 (1-p^*_2)/ n_2 .$$

С помощью указанной выше математической техники можно показать, что

$$\lim_{n_1\to\infty, n_2\to\infty} P\{\frac{p_1^* - p_2^* - M(p_1^* - p_2^*)}{\sqrt{D^*(p_1^* - p_2^*)}} \le x\} = \Phi(x).$$

При справедливости гипотезы однородности $M(p_1^* - p_2^*) = 0$. Поэтому правило принятия решения при проверке однородности двух выборок выглядит так:

1. Вычислить статистику

$$Q = \frac{p_1^* - p_2^*}{\sqrt{\frac{p_1^*(1-p_1^*)}{n_1} + \frac{p_2^*(1-p_2^*)}{n_2}}}.$$

2. Сравнить значение модуля статистика $|Q|$ с граничным значением K. Если $|Q|\le K$, то принять гипотезу однородности H_0 . Если же $|Q|>K$, то заявить об отсутствии однородности и принять альтернативную гипотезу H_1 .

Граничное значение K определяется выбором уровня значимости статистического критерия проверки однородности. Из приведенных выше предельных соотношений следует, что при справедливости гипотезы однородности H_0 для уровня значимости $\alpha = P(|Q| > K)$ имеем (при $n_1 \to \infty, n_2 \to \infty$)

$$\alpha \to \Phi(K) - \Phi(-K) = 2\Phi(K) - 1.$$

Следовательно, граничное значение в зависимости от уровня значимости целесообразно выбирать из условия

$$K = K(\alpha) = \Phi^{-1}(\frac{1+\alpha}{2}).$$

Здесь $\Phi^{-1}(\cdot)$ - функция, обратная к функции стандартного нормального распределения. В социально-экономических исследованиях наиболее распространен 5% уровень значимости, т.е. $\alpha = 0{,}05$. Для него $K = 1{,}96$.

Пример. Пусть в первой группе из 500 опрошенных ответили "да" 200, а во второй группе из 700 опрошенных сказали "да" 350. Есть ли разница между генеральными совокупностями, представленными этими двумя группами, по доле отвечающих "да"?

Уберем из формулировки примера термин "генеральная совокупность".

Пусть из 500 опрошенных мужчин ответили "да, я люблю пепси-колу" 200, а из 700 опрошенных женщин 350 сказали "да, я люблю пепси-колу". Есть ли разница между мужчинами и женщинами по доле отвечающих "да" на вопрос о любви к пепси-коле?

В рассматриваемом примере нужные для расчетов величины таковы: $n_1 = 500, p_1^* = 200/500 = 0{,}4; n_2 = 700, p_2^* = 350/700 = 0{,}5.$ Вычислим статистику

$$Q = \frac{0{,}4 - 0{,}5}{\sqrt{\dfrac{0{,}4 \cdot 0{,}6}{500} + \dfrac{0{,}5 \cdot 05}{700}}} = \frac{-0{,}1}{\sqrt{\dfrac{0{,}24}{500} + \dfrac{0{,}25}{700}}} = \frac{-0{,}1}{\sqrt{0{,}00048 + 0{,}0003571}}$$

$$= \frac{-0{,}1}{\sqrt{0{,}0008371}} = \frac{-0{,}1}{0{,}029} = -3{,}45.$$

Поскольку $|Q|$ = 3,45 > 1,96, то необходимо отклонить нулевую гипотезу т принять альтернативную. Таким образом, мужчины и женщины отличаются по рассматриваемому признаку - любви к пепси-коле.

Необходимо отметить, что результат проверки гипотезы однородности зависит не только от частот, но и от объемов выборок. Предположим, что частоты (доли) зафиксированы, а объемы выборок растут. Тогда числитель статистики Q не меняется, а знаменатель уменьшается, значит, вся дробь возрастает. Поскольку знаменатель стремится к 0, то дробь возрастает до бесконечности и рано или поздно превзойдет любую границу. Есть только одно исключение - когда в числителе стоит 0. Следовательно, вывод эконометрика должен выглядеть так: "различие обнаружено" или "различие не обнаружено". Во втором случае различие, возможно, было бы обнаружено при увеличении объемов выборок.

Как и для доверительного оценивания вероятности, во ВЦИОМ разработаны две полезные таблицы, позволяющие оценить вызванные чисто случайными причинами допустимые расхождения между частотами в группах. Эти таблицы рассчитаны при выполнении нулевой гипотезы однородности и соответствуют ситуациям, когда частоты близки к 50% (табл.7) или к 20% (табл.8). Если наблюдаемые частоты - от 30% до 70%, то рекомендуется пользоваться первой из этих таблиц, если от 10% до 30% или от 70% до 90% - то второй. Если наблюдаемые частоты меньше 10% или больше 90%, то теорема Муавра-Лапласа и основанные на ней асимптотические формулы дают не очень хорошие приближения, целесообразно применять иные, более продвинутые математические средства, в частности, приближения с помощью распределения Пуассона.

Табл.7. Допустимые расхождения (в %) между частотами в двух группах в случае, когда наблюдаются частоты от 30% до 70%

Объемы Групп	750	600	400	200	100
750	6	7	7	10	12
600	7	8	8	11	13
400	7	8	10	11	14
200	10	11	11	13	16
100	12	13	14	16	18

Табл.8. Допустимые расхождения (в %) между частотами в двух группах в случае, когда наблюдаются частоты от 10% до30% или от 70% до 90%

Объемы Групп	750	600	400	200	100
750	5	5	6	8	10
600	5	6	7	8	10
400	6	7	8	9	11
200	8	8	9	10	12
100	10	10	11	12	14

В условиях разобранного выше примера табл.7 дает допустимое расхождение 7%. Действительно, объем первой группы 500 отсутствует в таблице, но строки, соответствующие объемам 400и 600, совпадают для первых двух столбцов слева. Эти столбцы соответствуют объемам второй группы 750 и 600, между которыми расположен объем 700, данный в примере. Он ближе к 750, поэтому берем величину расхождения, стоящую на пересечении первого столбца и второй (и третьей) строк, т.е. 7%. Поскольку реальное расхождение (10%) больше, чем 7%, то делаем вывод о наличии значимого различия между группами. Естественно, этот вывод совпадает с полученным ранее расчетным путем.

Допустимое расхождение $\Delta = \Delta(\alpha)$ между частотами нетрудно получить расчетным путем. Для этого достаточно воспользоваться формулой для статистики Q и определить, при каком максимальном расхождении частот все еще делается вывод о том, что верна гипотеза однородности. Следовательно, допустимое расхождение $\Delta = \Delta(\alpha)$ находится из уравнения

$$K(\alpha) = \frac{\Delta(\alpha)}{\sqrt{\frac{p_1^*(1-p_1^*)}{n_1} + \frac{p_2^*(1-p_2^*)}{n_2}}}.$$

Таким образом,

$$\Delta(\alpha) = K(\alpha)\sqrt{\frac{p_1^*(1-p_1^*)}{n_1} + \frac{p_2^*(1-p_2^*)}{n_2}}.$$

Для данных примера $\Delta = \Delta(\alpha) =$ 1,96 $\times$0,029 = 0,057, или 5,7%, для уровня значимости 0,05. .

Для других уровней значимости надо использовать другие коэффициенты $K(\alpha)$. Так, K(0,01) = 2,58 для уровня значимости 1% и K(0,10) = 1,64 для уровня значимости 10%. Для данных примера $\Delta = \Delta(\alpha) =$ 2,58 $\times$0,029 = = 0,7482 $\approx$ 0,075, или 7,5%, для уровня значимости 0,01. Если округлить до ближайшего целого числа процентов, то получим 7%, как при использовании таблицы 7 выше.

Анализ таблиц 7 и 8 показывает, что для констатации различия частоты должны отличаться не менее чем на 6%, а при некоторых объемах выборок - более чем на 10%, при объемах выборок 100 и 100 - на 19%. Если частоты отличаются на 5% или менее, можно сразу сказать, что эконометрический анализ приведет к выводу о том, что различие не обнаружено (для выборок объемов не более 750).

В связи с этим возникает вопрос: каково типовое отличие частот в двух выборках из одной и той же совокупности? Разность частот в этом случае имеет нулевое математическое ожидание и дисперсию

$$p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right) = \frac{p(1-p)(n_1+n_2)}{n_1 n_2}.$$

Величина p(1-p) достигает максимума при p=1/2, и этот максимум равен 1/4. Если p=1/2, а объемы двух выборок совпадают и равны 500, то дисперсия разности частот равна

$$\frac{0,25 \times 1000}{500 \times 500} = \frac{250}{250 \times 1000} = \frac{1}{1000}.$$

Следовательно, среднее квадратического отклонение σ равно 0,032, или 3,2%. Поскольку для стандартной нормальной случайной величины в 50% случаев ее значение не превосходит по модулю 0,67 (а в 50% случаев - больше 0,67), то типовой разброс равен $0,67\sigma$, а в рассматриваемом случае- 2,1%. Приведенные соображения дают метод контроля за правильностью проведения повторных опросов. Если частоты излишне устойчивы, это подозрительно!

Цитированная литература

1. Сэндидж Ч., Фрайбургер В., Ротцолл К. Реклама: теория и практика: Пер. с англ. - М.: Прогресс, 1989. - 630 с.
2. Ядов В.А. Стратегии и методы качественного анализа данных. - Журнал "Социология: методология, методы, математические модели", 1991, No.1, с.14-31.
3. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. - М.: Наука, 1983. - - 416 с.
4. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979. - 296 с.
5. Опыт применения ЭВМ в социологических исследованиях. - М.: Институт социологических исследований АН СССР, Советская социологическая ассоциация, 1977. - 158 с.
6. Орлов А.И. Общий взгляд на статистику объектов нечисловой природы. - В сб.: Анализ нечисловой информации в социологических исследованиях (научные редакторы: В.Г. Андреенков, А.И.Орлов, Ю.Н.Толстова). - М.: Наука, 1985. = С.58-92.

Глава 3. Основы теории измерений

Теория измерений (в дальнейшем сокращенно ТИ) является одной из составных частей эконометрики. Она входит в состав *статистики объектов нечисловой природы*. Необходимость использования ТИ при анализе экономических данных рассмотрим на примере экспертного оценивания, в частности, в связи с агрегированием мнений экспертов, построением обобщенных показателей и рейтингов.

Использование чисел в жизни и хозяйственной деятельности людей отнюдь не всегда предполагает, что эти числа можно складывать и умножать, производить иные арифметические действия. Что бы вы сказали о человеке, который занимается умножением телефонных номеров? И отнюдь не всегда 2+2=4. Если вы вечером поместите в клетку двух животных, а потом еще двух, то отнюдь не всегда можно утром найти в этой клетке четырех животных. Их может быть и много больше - если вечером вы загнали в клетку овцематок или беременных кошек. Их может быть и меньше - если к двум волкам вы поместили двух ягнят. Числа используются гораздо шире, чем арифметика.

Так, например, мнения экспертов часто выражены в *порядковой шкале* (подробнее о шкалах говорится ниже), т.е. эксперт может сказать (и обосновать), что один показатель качества продукции более важен, чем другой, первый технологический объект более опасен, чем второй, и т.д. Но он не в состоянии сказать, *во сколько раз* или *на сколько* более важен, соответственно, более опасен. Экспертов часто просят дать ранжировку (упорядочение) объектов экспертизы, т.е. расположить их в порядке возрастания (или убывания) интенсивности интересующей организаторов экспертизы характеристики. Ранг - это номер (объекта экспертизы) в упорядоченном ряду значений характеристики у различных объектов. Такой ряд в статистике называется вариационным. Формально ранги выражаются числами 1, 2, 3, ..., но с этими числами нельзя делать привычные арифметические операции. Например, хотя в арифметике 1 + 2 = 3, но нельзя утверждать, что для объекта, стоящем на третьем месте в упорядочении, интенсивность изучаемой характеристики равна сумме интенсивностей объектов с рангами 1 и 2. Так, один из видов экспертного оценивания - оценки учащихся. Вряд ли кто-либо будет утверждать, что знания отличника равны сумме знаний двоечника и троечника (хотя 5 = 2 + 3), хорошист соответствует двум двоечникам (2 + 2 = 4), а между отличником и троечником такая же разница, как между хорошистом и двоечником (5 - 3 = 4 - 2). Поэтому очевидно, что для анализа подобного рода качественных данных необходима не всем известная арифметика, а другая теория, дающая базу для разработки, изучения и применения конкретных методов расчета. Это и есть ТИ. (При чтении литературы надо иметь в виду, что в настоящее время термин "теория измерений" применяется для обозначения целого ряда научных дисциплин: классической метрологии (науки об измерениях физических величин), рассматриваемой здесь ТИ, некоторых других направлений, например, алгоритмической теории измерений. Обычно из контекста понятно, о какой конкретно теории идет речь.))

Краткая история теории измерений. Сначала ТИ развивалась как теория психофизических измерений. В послевоенных публикациях американский психолог С.С. Стивенс основное внимание уделял шкалам измерения. Во второй половине XX в. сфера применения ТИ стремительно расширяется. Посмотрим, как это происходило. Один из томов выпущенной в США в 1950-х годах "Энциклопедии психологических наук" назывался "Психологические измерения". Значит, составители этого тома расширили сферу применения РТИ с психофизики на психологию в целом. А в основной статье в этом сборнике под названием, обратите внимание, "Основы теории измерений", изложение шло на абстрактно-математическом уровне, без привязки к какой-либо конкретной области применения. В этой статье [1] упор был сделан на "гомоморфизмах эмпирических систем с отношениями в числовые" (в эти математические термины здесь вдаваться нет необходимости), и математическая сложность изложения возросла по сравнению с работами С.С. Стивенса.

Уже в одной из первых отечественных статей по РТИ (конец 1960-х годов) было установлено, что баллы, присваиваемые экспертами при оценке объектов экспертизы, как правило, измерены в порядковой шкале. Отечественные работы, появившиеся в начале 1970-х годов, привели к существенному расширению области использования РТИ. Ее применяли к педагогической квалиметрии (измерению качества знаний учащихся), в системных исследованиях, в различных задачах теории экспертных оценок, для агрегирования показателей качества продукции, в социологических исследованиях, и др.

Итоги этого этапа были подведены в монографии [2]. В качестве двух основных проблем РТИ наряду с *установлением типа шкалы* измерения конкретных данных был выдвинут поиск алгоритмов анализа данных, результат работы которых не меняется при любом допустимом преобразовании шкалы (т.е. является *инвариантным* относительно этого преобразования).

Метрологи вначале резко возражали против использования термина "измерение" для качественных признаков. Однако постепенно возражения сошли на нет, и к концу XX в. ТИ стала рассматриваться как общенаучная теория.

3.1. Основные шкалы измерения

В соответствии с ТИ при математическом моделировании реального явления или процесса следует прежде всего установить *типы шкал*, в которых измерены те или иные переменные. Тип шкалы задает *группу допустимых преобразований шкалы*. Допустимые преобразования не меняют соотношений между объектами измерения. Например, при измерении длины переход от аршин к метрам не меняет соотношений между длинами рассматриваемых объектов - если первый объект длиннее второго, то это будет установлено и при измерении в аршинах, и при измерении в метрах. Обратите внимание, что при этом численное значение длины в аршинах отличается от численного значения длины в метрах - не меняется лишь результат сравнения длин двух объектов.

Укажем основные виды шкал измерения и соответствующие группы допустимых преобразований.

В *шкале наименований* (другое название этой шкалы - *номинальная;* это - переписанное русскими буквами английское название шкалы) **допустимыми** являются все взаимно-однозначные преобразования. В этой шкале числа используются лишь как метки. Примерно так же, как при сдаче белья в прачечную, т.е. лишь для различения объектов. В шкале наименований измерены, например, номера телефонов, автомашин, паспортов, студенческих билетов. Номера страховых свидетельств государственного пенсионного страхования, медицинского страхования, ИНН (индивидуальный номер налогоплательщика) измерены в шкале наименований. Пол людей тоже измерен в шкале наименований, результат измерения принимает два значения - мужской, женский. Раса, национальность, цвет глаз, волос - номинальные признаки. Номера букв в алфавите - тоже измерения в шкале наименований. Никому в здравом уме не придет в голову складывать или умножать номера телефонов, такие операции не имеют смысла. Сравнивать буквы и говорить, например, что буква П лучше буквы С, также никто не будет. Единственное, для чего годятся измерения в шкале наименований - это различать объекты. Во многих случаях только это от них и требуется. Например, шкафчики в раздевалках для взрослых различают по номерам, т.е. числам, а в детских садах используют рисунки, поскольку дети еще не знают чисел.

В *порядковой шкале* числа используются не только для различения объектов, но и для установления порядка между объектами. Простейшим примером являются оценки знаний учащихся. Символично, что в средней школе применяются оценки 2, 3, 4, 5, а в высшей школе ровно тот же смысл выражается словесно - неудовлетворительно, удовлетворительно, хорошо, отлично. Этим подчеркивается "нечисловой" характер оценок знаний учащихся. В порядковой шкале **допустимыми** являются все строго возрастающие преобразования.

Установление типа шкалы, т.е. задания группы допустимых преобразований шкалы измерения - дело специалистов соответствующей прикладной области. Так, оценки привлекательности профессий мы в монографии [2], выступая в качестве социологов, считали измеренными в порядковой шкале. Однако отдельные социологи не соглашались с нами, полагая, что выпускники школ пользуются шкалой с более узкой группой допустимых преобразований, например, интервальной шкалой. Очевидно, эта проблема относится не к математике, а к наукам о человеке. Для ее решения может быть поставлен достаточно трудоемкий эксперимент. Пока же он не поставлен, целесообразно принимать порядковую шкалу, так как это гарантирует от возможных ошибок.

Оценки экспертов, как уже отмечалось, часто следует считать измеренными в порядковой шкале. Типичным примером являются задачи ранжирования и классификации промышленных объектов, подлежащих экологическому страхованию.

Почему мнения экспертов естественно выражать именно в порядковой шкале? **Как показали многочисленные опыты, человек более правильно (и с меньшими затруднениями) отвечает на вопросы качественного, например, сравнительного, характера, чем количественного.** Так, ему легче сказать, какая из двух гирь тяжелее, чем указать их примерный вес в граммах.

В различных областях человеческой деятельности применяется много других видов порядковых шкал. Так, например, в минералогии используется шкала Мооса, по которому минералы классифицируются согласно критерию твердости. А именно: тальк имеет балл 1, гипс - 2, кальций - 3, флюорит - 4, апатит - 5, ортоклаз - 6, кварц - 7, топаз - 8, корунд - 9, алмаз - 10. Минерал с большим номером является более твердым, чем минерал с меньшим номером, при нажатии царапает его.

Порядковыми шкалами в географии являются - бофортова шкала ветров ("штиль", "слабый ветер", "умеренный ветер" и т.д.), шкала силы землетрясений. Очевидно, нельзя утверждать, что землетрясение в 2 балла (лампа качнулась под потолком - такое бывает и в Москве) ровно в 5 раз слабее, чем землетрясение в 10 баллов (полное разрушение всего на поверхности земли).

В медицине порядковыми шкалами являются - шкала стадий гипертонической болезни (по Мясникову), шкала степеней сердечной недостаточности (по Стражеско-Василенко-Лангу), шкала степени выраженности коронарной недостаточности (по Фогельсону), и т.д. Все эти шкалы построены по схеме: заболевание не обнаружено; первая стадия заболевания; вторая стадия; третья стадия… Иногда выделяют стадии 1а, 1б и др. Каждая стадия имеет свойственную только ей медицинскую характеристику. При описании групп инвалидности числа используются в противоположном порядке: самая тяжелая - первая группа инвалидности, затем - вторая, самая легкая - третья.

Номера домов также измерены в порядковой шкале - они показывают, в каком порядке стоят дома вдоль улицы. Номера томов в собрании сочинений писателя или номера дел в архиве предприятия обычно связаны с хронологическим порядком их создания.

При оценке качества продукции и услуг, в т.н. квалиметрии (буквальный перевод: измерение качества) популярны порядковые шкалы. А именно, единица продукции оценивается как годная или не годная. При более тщательном анализе используется шкала с тремя градациями: есть значительные дефекты - присутствуют только незначительные дефекты - нет дефектов. Иногда применяют четыре градации: имеются критические дефекты (делающие невозможным использование) - есть значительные дефекты - присутствуют только незначительные дефекты - нет дефектов. Аналогичный смысл имеет сортность продукции - высший сорт, первый сорт, второй сорт,…

При оценке экологических воздействий первая, наиболее обобщенная оценка - обычно порядковая, например: природная среда стабильна - природная среда угнетена (деградирует). Аналогично в эколого-медицинской шкале: нет выраженного воздействия на здоровье людей - отмечается отрицательное воздействие на здоровье.

Порядковая шкала используется и во многих иных областях. В эконометрике это прежде всего различные методы экспертных оценок (см. посвященную им главу 12).

Все шкалы измерения делят на две группы - шкалы качественных признаков и шкалы количественных признаков.

Порядковая шкала и шкала наименований - основные шкалы качественных признаков. Поэтому во многих конкретных областях результаты качественного анализа можно рассматривать как измерения по этим шкалам.

Шкалы количественных признаков - это шкалы интервалов, отношений, разностей, абсолютная. По шкале *интервалов* измеряют величину потенциальной энергии или координату точки на прямой. В этих случаях на шкале нельзя отметить ни естественное начало отсчета, ни естественную единицу измерения. Исследователь должен сам задать точку отсчета и сам выбрать единицу измерения. Допустимыми преобразованиями в шкале интервалов являются линейные возрастающие преобразования, т.е. линейные функции. Температурные шкалы Цельсия и Фаренгейта связаны именно такой зависимостью: $^0C = 5/9\ (^0F - 32)$, где ^0C - температура (в градусах) по шкале Цельсия, а ^0F - температура по шкале Фаренгейта.

Из количественных шкал наиболее распространенными в науке и практике являются шкалы *отношений.* В них есть естественное начало отсчета - нуль, т.е. отсутствие величины, но нет естественной единицы измерения. По шкале отношений измерены большинство физических единиц: масса тела, длина, заряд, а также цены в экономике. Допустимыми преобразованиями шкале отношений являются подобные (изменяющие только масштаб). Другими словами, линейные возрастающие преобразования без свободного члена. Примером является пересчет цен из одной валюты в другую по фиксированному курсу. Предположим, мы сравниваем экономическую эффективность двух инвестиционных проектов, используя цены в рублях. Пусть первый проект оказался лучше второго. Теперь перейдем на валюту самой экономически мощной державы мира - юани, используя фиксированный курс пересчета. Очевидно, первый проект должен опять оказаться более выгодным, чем второй. Это очевидно из общих соображений. Однако алгоритмы расчета не обеспечивают автоматически выполнения этого очевидного условия. Надо проверять, что оно выполнено. Результаты подобной проверки для средних величин описаны ниже.

В шкале разностей есть естественная единица измерения, но нет естественного начала отсчета. Время измеряется по шкале *разностей*, если год (или сутки - от полудня до полудня) принимаем естественной единицей измерения, и по шкале интервалов в общем случае. На современном уровне знаний естественного начала отсчета указать нельзя. Дату сотворения мира различные авторы рассчитывают по-разному, равно как и момент рождества Христова. Так, согласно новой статистической хронологии, разработанной группой акад. РАН А.Т.Фоменко, Господь Иисус Христос родился примерно в 1054 г. по принятому ныне летоисчислению в Стамбуле (он же - Царьград, Византия, Троя, Иерусалим, Рим).

Только для *абсолютной* шкалы результаты измерений - числа в обычном смысле слова. Примером является число людей в комнате. Для абсолютной шкалы допустимым является только тождественное преобразование.

В процессе развития соответствующей области знания тип шкалы может меняться. Так, сначала температура измерялась по *порядковой* шкале (холоднее - теплее). Затем - по *интервальной* (шкалы Цельсия, Фаренгейта, Реомюра). Наконец, после открытия абсолютного нуля температуру можно считать измеренной по шкале *отношений* (шкала Кельвина). Надо отметить, что среди специалистов иногда имеются разногласия по поводу того, по каким шкалам следует считать измеренными те или иные реальные величины. Другими словами, процесс измерения включает в себя и определение типа шкалы (вместе с обоснованием выбора определенного типа шкалы). Кроме перечисленных шести основных типов шкал, иногда используют и иные шкалы.

3.2. Инвариантные алгоритмы и средние величины

Основное требование к алгоритмам анализа данных формулируется в ТИ так: **выводы, сделанные на основе данных, измеренных в шкале определенного типа, не должны меняться при допустимом преобразовании шкалы измерения этих данных.** Другими словами, выводы должны быть *инвариантны* по отношению к допустимым преобразованиям шкалы.

Таким образом, одна из основных целей теории измерений - борьба с субъективизмом исследователя при приписывании численных значений реальным объектам. Так, расстояния можно измерять в аршинах, метрах, микронах, милях, парсеках и других единицах измерения. Массу (вес) - в пудах, килограммах, фунтах и др. Цены на товары и услуги можно указывать в юанях, рублях, тенге, гривнах, латах, кронах, марках, долларах США и других валютах (при условии заданных курсов пересчета). Подчеркнем очень важное, хотя и вполне очевидное обстоятельство: выбор единиц измерения зависит от исследователя, т.е. субъективен. *Статистические выводы могут быть адекватны реальности только тогда, когда они не зависят от того, какую единицу измерения предпочтет исследователь, т.е. когда они инвариантны относительно допустимого преобразования шкалы.*

Оказывается, сформулированное условие является достаточно сильным. Из многих алгоритмов эконометрического анализа данных ему удовлетворяют лишь некоторые. Покажем это на примере сравнения средних величин.

Пусть $X_1, X_2, \dots, X_n$ - выборка объема n. Часто используют среднее арифметическое

$$X_{cp} = \frac{X_1 + X_2 + \dots + X_n}{n}.$$

Использование среднего арифметического настолько привычно, что второе слово в термине часто опускают. И говорят о средней зарплате, среднем доходе и других средних для конкретных экономических данных, подразумевая под "средним" среднее арифметическое. Такая традиция может приводить к ошибочным выводам. Покажем это на примере расчета средней заработной платы (среднего дохода) работников условного предприятия (табл.1).

Табл.1. Численность работников различных категорий, их заработная плата и доходы (в условных единицах).

№ п/п	Категория работников	Число работников	Заработная плата	Суммарные доходы
1	Низкоквалифицированные рабочие	40	100	4000
2	Высококвалифицированные рабочие	30	200	6000
3	Инженеры и служащие	25	300	7500
4	Менеджеры	4	1000	4000
5	Генеральный директор (владелец)	1	18500	18500
6	Всего	100		40000

Первые три строки в табл.1 вряд ли требуют пояснений. Менеджеры - это директора по направлениям, а именно, по производству (главный инженер), по финансам, по маркетингу и сбыту, по персоналу (по кадрам). Владелец сам руководит предприятием в качестве генерального директора. В столбце "заработная плата" указаны доходы одного работника соответствующей категории, а в столбце "суммарные доходы" - доходы всех работников соответствующей категории.

Фонд оплаты труда составляет 40000 единиц, работников всего 100, следовательно, средняя заработная плата составляет 40000/100 = 400 единиц. Однако эта средняя арифметическая величина явно не соответствует интуитивному представлению о "средней зарплате". Из 100 работников лишь 5 имеют заработную плату, ее превышающую, а зарплата остальных 95 существенно меньше средней арифметической. Причина очевидна - заработная плата одного человека - генерального директора - превышает заработную плату 95 работников - низкоквалифицированных и высококвалифицированных рабочих, инженеров и служащих.

Ситуация напоминает описанную в известном рассказе о больнице, в которой 10 больных, из них у 9 температура 40 ^{0}С, а один уже отмучился, лежи в морге с температурой 0 ^{0}С. Между тем средняя температура по больнице равна 36 ^{0}С - лучше не бывает!

Сказанное показывает, что среднее арифметическое можно использовать лишь для достаточно однородных совокупностей (без больших выбросов в ту или иную сторону). А какие средние использовать для описания заработной платы? Вполне естественно использовать медиану. Для данных табл.1 медиана - среднее арифметическое 50-го и 51-го работника, если их заработные платы расположены в порядке неубывания. Сначала идут зарплаты 40 низкоквалифицированных рабочих, а затем - с 41-го до 70-го работника - заработные платы высококвалифицированных рабочих. Следовательно, медиана попадает именно на них и равна 200. У 50-ти работников заработная плата не превосходит 200, и у 50-ти - не менее 200, поэтому медиана показывает "центр", около которого группируется основная масса исследуемых величин. Еще одна средняя величина - мода, наиболее часто встречающееся значение. В рассматриваемом случае это заработная плата низкоквалифицируемых рабочих, т.е. 100. Таким образом, для описания зарплаты имеем три средние величины - моду (100 единиц), медиану (200 единиц) и среднее арифметическое (400 единиц). Для наблюдающихся в реальной жизни распределений доходов и заработной платы справедлива та же закономерность: мода меньше медианы, а медиана меньше среднего арифметического.

Для чего в экономике используются средние величины? Обычно для того, чтобы заменить совокупность чисел одним числом, чтобы сравнивать совокупности с помощью средних.

Пусть, например, $Y_1, Y_2,...,Y_n$ - совокупность оценок экспертов, "выставленных" одному объекту экспертизы (например, одному из вариантов стратегического развития фирмы), $Z_1, Z_2,...,Z_n$ - второму (другому варианту такого развития). Как сравнивать эти совокупности? Очевидно, самый простой способ - по средним значениям.

А как вычислять средние? Известны различные виды средних величин: среднее арифметическое, медиана, мода, среднее геометрическое, среднее гармоническое, среднее квадратическое. Напомним, что общее понятие средней величины введено французским математиком первой половины XIX в. академиком О. Коши. Оно таково: средней величиной является любая функция $f(X_1, X_2,...,X_n)$ такая, что при всех возможных значениях аргументов значение этой функции не меньше, чем минимальное из чисел $X_1, X_2,...,X_n$, и не больше, чем максимальное из этих чисел. Все перечисленные выше виды средних являются средними по Коши.

При допустимом преобразовании шкалы значение средней величины, очевидно, меняется. Но выводы о том, для какой совокупности среднее больше, а для какой - меньше, не должны меняться (в соответствии с требованием инвариантности выводов, принятом как основное требование в ТИ). Сформулируем соответствующую математическую задачу поиска вида средних величин, результат сравнения которых устойчив относительно допустимых преобразований шкалы.

Пусть $f(X_1, X_2,...,X_n)$ - среднее по Коши. Пусть среднее по первой совокупности меньше среднего по второй совокупности:

$$f(Y_1, Y_2,..., Y_n) < f(Z_1, Z_2,...,Z_n).$$

Тогда согласно ТИ для устойчивости результата сравнения средних необходимо, чтобы для любого допустимого преобразования g из группы допустимых преобразований в соответствующей шкале было справедливо также неравенство

$$f(g(Y_1), g(Y_2),...,g(Y_n)) < f(g(Z_1), g(Z_2),...,g(Z_n)).$$

т.е. среднее преобразованных значений из первой совокупности также было меньше среднего преобразованных значений для второй совокупности. Причем сформулированное условие должно быть верно для любых двух совокупностей $Y_1, Y_2,...,Y_n$и $Z_1, Z_2,...,Z_n$ и, напомним, любого допустимого преобразования. Средние величины, удовлетворяющие сформулированному условию, назовем допустимыми (в соответствующей шкале). Согласно ТИ только такими средними можно пользоваться при анализе мнений экспертов и иных данных, измеренных в рассматриваемой шкале.

С помощью математической теории, развитой в монографии [2], удается описать вид допустимых средних в основных шкалах. Сразу ясно, что для данных, измеренных в шкале наименований, в качестве среднего годится только мода.

3.3. Средние величины в порядковой шкале

Рассмотрим обработку мнений экспертов, измеренных в порядковой шкале. Справедливо следующее утверждение.

Теорема 1. *Из всех средних по Коши допустимыми средними в порядковой шкале являются только члены вариационного ряда (порядковые статистики).*

Теорема 1 справедлива при условии, что среднее $f(X_1, X_2,...,X_n)$ является непрерывной (по совокупности переменных) и симметрической функцией. Последнее означает, что при перестановке аргументов значение функции $f(X_1, X_2,...,X_n)$ не меняется. Это условие является вполне естественным, ибо среднюю величину мы находим для *совокупности (множества)*, а не для *последовательности*. Множество не меняется в зависимости от того, в какой последовательности мы перечисляем его элементы.

Согласно теореме 1 в качестве среднего для данных, измеренных в порядковой шкале, можно использовать, в частности, медиану (при нечетном объеме выборки). При четном же объеме следует применять один из двух центральных членов вариационного ряда - как их иногда называют, левую медиану или правую медиану. Моду тоже можно использовать - она всегда является членом вариационного ряда. Но никогда нельзя рассчитывать среднее арифметическое, среднее геометрическое и т.д.

Приведем численный пример, показывающий некорректность использования среднего арифметического $f(X_1, X_2) = (X_1 + X_2)/2$ в порядковой шкале. Пусть $Y_1= 1$, $Y_2 = 11$, $Z_1= 6$, $Z_2= 8$. Тогда $f(Y_1, Y_2) = 6$, что меньше, чем $f(Z_1, Z_2) = 7$. Пусть строго возрастающее преобразование g таково, что $g(1) = 1$, $g(6) = 6$, $g(8) = 8$, $g(11) = 99$. Таких преобразований много. Например, можно положить $g(x) = x$ при x, не превосходящих 8, и $g(x) = 99(x-8)/3 + 8$ для x, больших 8. Тогда $f(g(Y_1), g(Y_2)) = 50$, что больше, чем $f(g(Z_1), g(Z_2)) = 7$. Как видим, в результате допустимого, т.е. строго возрастающего преобразования шкалы упорядоченность средних изменилась.

Таким образом, ТИ выносит жесткий приговор среднему арифметическому - использовать его с порядковой шкале нельзя. Однако же те, кто не знает теории измерений, используют его. Всегда ли они ошибаются? Оказывается, можно в какой-то мере реабилитировать среднее арифметическое, если перейти к вероятностной постановке и к тому удовлетвориться результатами для больших объемов выборок. В монографии [2] получено также следующее утверждение.

Теорема 2. *Пусть $Y_1, Y_2,...,Y_m$ - независимые одинаково распределенные случайные величины с функцией распределения $F(x)$, а $Z_1, Z_2,...,Z_n$ - независимые одинаково* *распределенные случайные величины с функцией распределения H(x), причем выборки Y_1, Y_2,..., Y_m и Z_1, Z_2,...,Z_n независимы между собой и $MY_1 > MZ_1$. Для того, чтобы вероятность события*

$$\{\omega : \frac{g(Y_1) + g(Y_2) + ... + g(Y_m)}{m} > \frac{g(Z_1) + g(Z_2) + ... + g(Z_n)}{n}\}$$

стремилась к 1 при $\min(m, n) \to \infty$ *для любой строго возрастающей непрерывной функции g, удовлетворяющей условию*

$$\overline{\lim_{|x| \to \infty}} \left| \frac{g(x)}{x} \right| < \infty,$$

необходимо и достаточно, чтобы при всех x выполнялось неравенство $F(x) \leq H(x)$, причем существовало число x_0 , для которого $F(x_0) < H(x_0)$.

Примечание. *Условие с верхним пределом носит чисто внутриматематический характер. Фактически функция g - произвольное допустимое преобразование в порядковой шкале.*

Согласно теореме 2 средним арифметическим можно пользоваться и в порядковой шкале, если сравниваются выборки из двух распределений, удовлетворяющих приведенному в теореме неравенству. Проще говоря, одна из функций распределения должна всегда лежать над другой. Функции распределения не могут пересекаться, им разрешается только касаться друг друга. Это условие выполнено, например, если функции распределения отличаются только сдвигом:

$$F(x) = H(x+b)$$

при некотором b . Последнее условие выполняется, если два значения некоторой величины измеряются с помощью одного и того же средства измерения, у которого распределение погрешностей не меняется при переходе от измерения одного значения рассматриваемой величины к измерению другого.

3.4. Средние по Колмогорову

Обобщением нескольких из перечисленных выше средних является среднее по Колмогорову. Для чисел X_1, X_2,...,X_n среднее по Колмогорову вычисляется по формуле

$$G\{(F(X_1)+F(X_2)+...F(X_n))/n\},$$

где F - строго монотонная функция (т.е. строго возрастающая или строго убывающая), G - функция, обратная к F. Среди средних по Колмогорову - много хорошо известных персонажей. Так, если $F(x) = x$, то среднее по Колмогорову - это среднее арифметическое, если $F(x) = \ln x$, то среднее геометрическое, если $F(x) = 1/x$, то среднее гармоническое, если $F(x) = x^2$, то среднее квадратическое, и т.д. Среднее по Колмогорову - частный случай среднего по Коши. С другой стороны, такие популярные средние, как медиана и мода, нельзя представить в виде средних по Колмогорову. В монографии [2] доказаны следующие утверждения.

Теорема 3. *При справедливости некоторых внутриматематических условий регулярности в шкале интервалов из всех средних по Колмогорову допустимым является только среднее арифметическое.*

Таким образом, среднее геометрическое или среднее квадратическое температур (в шкале Цельсия) или расстояний не имеют смысла. В качестве среднего надо применять среднее арифметическое. А также можно использовать медиану или моду.

Теорема 4. *При справедливости некоторых внутриматематических условий регулярности в шкале отношений из всех средних по Колмогорову допустимыми являются только степенные средние с $F(x) = x^c$, $c \neq 0$, и среднее геометрическое.*

Замечание. *Среднее геометрическое является пределом степенных средних при* $c \to 0$.

Есть ли средние по Колмогорову, которыми нельзя пользоваться в шкале отношений? Конечно, есть. Например, с $F(x) = e^x$.

Аналогично средним величинам могут быть изучены и другие статистические характеристики - показатели разброса, связи, расстояния и др. (см., например, [2]). Нетрудно показать, например, что коэффициент корреляции не меняется при любом допустимом преобразовании в шкале интервалов, как и отношение дисперсий, дисперсия не меняется в шкале разностей, коэффициент вариации - в шкале отношений, и т.д.

Приведенные выше результаты о средних величинах широко применяются, причем не только в экономике, менеджменте, теории экспертных оценок или социологии, но и в инженерном деле, например, для анализа методов агрегирования датчиков в АСУ ТП доменных печей. Велико прикладное значение ТИ в задачах стандартизации и управления качеством, в частности, в квалиметрии. Здесь есть и интересные теоретические результаты. Так, например, любое изменение коэффициентов весомости единичных показателей качества продукции приводит к изменению упорядочения изделий по средневзвешенному показателю (эта теорема доказана проф. В.В. Подиновским).

Цитированная литература

1. Суппес П., Зинес Дж. Основы теории измерений. - В сб.: Психологические измерения. - М.: Мир, 1967. С. 9-110.
2. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979. - 296 с.
3. Пфанцагль И. Теория измерений. - М.: Мир, 1976. - 165 с.

Глава 4. Статистический анализ числовых величин (непараметрическая статистика)

В учебных курсах по теории вероятностей и математической статистике рассматривают различные параметрические семейства распределений числовых случайных величин. А именно, изучают семейства нормальных распределений, логарифмически нормальных, экспоненциальных, гамма-распределений, распределений Вейбулла-Гнеденко и др. Все они зависят от одного, двух или трех параметров. Поэтому для полного описания распределения достаточно знать или оценить одно, два или три числа. Очень удобно. Поэтому широко развита параметрическая теория математической статистики, в которой предполагается, что распределения результатов наблюдений принадлежат тем или иным параметрическим семействам.

К сожалению, параметрические семейства существуют лишь в головах авторов учебников по теории вероятностей и математической статистике. В реальной жизни их нет. Поэтому эконометрика использует в основном непараметрические методы, в которых распределения результатов наблюдений могут иметь произвольный вид.

Сначала на примере нормального распределения подробнее обсудим невозможность практического использования параметрических семейств для описания распределений конкретных экономических данных. Затем разберем параметрические методы отбраковки резко выделяющихся наблюдений и продемонстрируем невозможность практического использования ряда методов параметрической статистики, ошибочность выводов, к которым они приводят. Затем разберем непараметрические методы доверительного оценивания основных характеристик числовых случайных величин - математического ожидания, медианы, дисперсии, среднего квадратического отклонения, коэффициента вариации. Завершат главу методы проверки однородности двух выборок, независимых или связанных.

4.1. Часто ли распределение результатов наблюдений является нормальным?

В эконометрических и экономико-математических моделях, применяемых, в частности, при изучении и оптимизации процессов маркетинга и менеджмента, управления предприятием и регионом, точности и стабильности технологических процессов, в задачах надежности, обеспечения безопасности, в том числе экологической, функционирования технических устройств и объектов, разработки организационных схем часто применяют понятия и результаты теории вероятностей и математической статистики. При этом зачастую используют те или иные параметрические семейства распределений вероятностей. Наиболее популярно нормальное распределение. Используют также логарифмически нормальное распределение, экспоненциальное распределение, гамма-распределение, распределение Вейбулла-Гнеденко и т.д.

Очевидно, всегда необходимо проверять соответствие моделей реальности. Возникают два вопроса. Отличаются ли реальные распределения от используемых в модели? Насколько это отличие влияет на выводы?

Ниже на примере нормального распределения и основанных на нем методов отбраковки резко отличающихся наблюдений (выбросов) показано, что реальные распределения практически всегда отличаются от включенных в классические параметрические семейства, а имеющиеся отклонения от заданных семейств делают неверными выводы, в рассматриваемом случае, об отбраковке, основанные на использовании этих семейств.

Есть ли основания априори предполагать нормальность результатов измерений?

Иногда утверждают, что в случае, когда погрешность измерения (или иная случайная величина) определяется в результате совокупного действия многих малых факторов, то в силу Центральной Предельной Теоремы (ЦПТ) теории вероятностей эта величина хорошо приближается (по распределению) нормальной случайной величиной. Такое утверждение справедливо, если малые факторы действуют аддитивно и независимо друг от друга. Если же они действуют мультипликативно, то в силу той же ЦПТ аппроксимировать надо логарифмически нормальным распределением. В прикладных задачах обосновать аддитивность, а не мультипликативность действия малых факторов обычно не удается. Если же зависимость имеет общий характер, не приводится к аддитивному или мультипликативному виду, а также нет оснований принимать модели, дающие экспоненциальное, Вейбулла-Гнеденко, гамма или иные распределения, то о распределении итоговой случайной величины практически ничего не известно, кроме внутриматематических свойств типа регулярности.

При обработке конкретных данных иногда считают, что погрешности измерений имеют нормальное распределение. На предположении нормальности построены классические модели регрессионного, дисперсионного, факторного анализов, метрологические модели, которые еще продолжают встречаться как в отечественной нормативно-технической документации, так и в международных стандартах. На то же предположение опираются модели расчетов максимально достигаемых уровней тех или иных характеристик, применяемые при проектировании систем обеспечения безопасности функционирования экономических структур, технических устройств и объектов. Однако теоретических оснований для такого предположения нет. Необходимо экспериментально изучать распределения погрешностей.

Что же показывают результаты экспериментов? Сводка, данная в монографии [1], позволяет утверждать, что в большинстве случаев распределение погрешностей измерений отличается от нормального. Так, в Машинно-электротехническом институте (г. Варна в Болгарии) было исследовано распределение погрешностей градуировки шкал аналоговых электроизмерительных приборов. Изучались приборы, изготовленные в Чехословакии, СССР и Болгарии. Закон распределения погрешностей оказался одним и тем же. Он имеет плотность

$$f(x) = 0{,}534\exp(1-|x|^7).$$

Были проанализированы данные о параметрах 219 фактических распределениях погрешностей, исследованных разными авторами, при измерении как электрических, так и не электрических величин самыми разнообразными (электрическими) приборами. В результате этого исследования оказалось, что 111 распределений, т.е. примерно 50% , принадлежат классу распределений с плотностью

$$f(x;\alpha,b.\sigma) = \frac{\alpha}{2\lambda\sigma\Gamma(1/\alpha)}\exp\left(-\left|\frac{x-b}{\lambda\sigma}\right|^{\alpha}\right),$$

где α - параметр степени; b - параметр сдвига; σ - параметр масштаба; $\Gamma(\beta)$ - гамма-функция от аргумента β ;

$$\lambda = \sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}}$$

(см. [1, с. 56]); 63 распределения, т.е. 30%, имеют плотности с плоской вершиной и пологими длинными спадами и не могут быть описаны как нормальные или, например, экспоненциальные. Оставшиеся 45 распределений оказались двухмодальными.

В книге известного метролога проф. П. В. Новицкого [2] приведены результаты исследования законов распределения различного рода погрешностей измерения. Он изучил распределения погрешностей электромеханических приборов на кернах, электронных приборов для измерения температур и усилий, цифровых приборов с ручным уравновешиванием. Объем выборок экспериментальных данных для каждого экземпляра составлял 100-400 отсчетов. Оказалось, что 46 из 47 распределений значимо отличались от нормального. Исследована форма распределения погрешностей у 25 экземпляров цифровых вольтметров Щ-1411 в 10 точках диапазона. Результаты аналогичны. Дальнейшие сведения содержатся в монографии [1].

В лаборатории прикладной математики Тартуского государственного университета проанализировано 2500 выборок из архива реальных статистических данных. В 92% гипотезу нормальности пришлось отвергнуть.

Приведенные описания экспериментальных данных показывают, что погрешности измерений в большинстве случаев имеют распределения, отличные от нормальных. Это означает, в частности, что большинство применений критерия Стьюдента, классического регрессионного анализа и других статистических методов, основанных на нормальной теории, строго говоря, не является обоснованным, поскольку неверна лежащая в их основе аксиома нормальности распределений соответствующих случайных величин.

Очевидно, для оправдания или обоснованного изменения существующей практики анализа статистических данных требуется изучить свойства процедур анализа данных при "незаконном" применении. Изучение процедур отбраковки показало, что они крайне неустойчивы к отклонениям от нормальности, а потому применять их для обработки реальных данных нецелесообразно (см. ниже); поэтому нельзя утверждать, что произвольно взятая процедура устойчива к отклонениям от нормальности.

Иногда предлагают перед применением, например, критерия Стьюдента однородности двух выборок проверять нормальность. Хотя для этого имеется много критериев, но проверка нормальности - более сложная и трудоемкая статистическая процедура, чем проверка однородности (как с помощью статистик типа Стьюдента, так и с помощью непараметрических критериев). Для достаточно надежного установления нормальности требуется весьма большое число наблюдений. Так, чтобы гарантировать, что функция распределения результатов наблюдений отличается от некоторой нормальной не более, чем на 0,01 (при любом значении аргумента), требуется порядка 2500 наблюдений. В большинстве экономических, технических, медико-биологических и других прикладных исследований число наблюдений существенно меньше. Особенно это справедливо для данных, используемых при изучении проблем, связанных с обеспечением безопасности функционирования экономических структур и технических объектов.

Иногда пытаются использовать ЦПТ для приближения распределения погрешности к нормальному, включая в технологическую схему измерительного прибора специальные сумматоры. Оценим полезность этой меры. Пусть Z_1 , Z_2 ,..., Z_k - независимые одинаково распределенные случайные величины с функцией распределения $H = H(x)$ такие, что $M(Z_1) = 0,\ D(Z_1) = 1,\ M\,|\,Z_1\,|^3 = \rho < +\infty.$ Рассмотрим

$$w = \frac{Z_1 + Z_2 + \ldots + Z_k}{\sqrt{k}}.$$

Показателем обеспечиваемой сумматором близости к нормальности является

$$C = \sup_H \sup_x |\,P(w < x) - \Phi(x)\,|.$$

Тогда

$$0{,}3989\frac{\rho}{\sqrt{k}} \le C \le 0{,}7975\frac{\rho}{\sqrt{k}}.$$

Правое неравенство в последнем соотношении вытекает из оценок константы в неравенстве

Берри-Эссеена, полученном в книге [3, с.172], а левое - из примера в монографии [4, с.140-141]. Для нормального закона ρ =1,6, для равномерного ρ = 1,3, для двухточечного ρ =1 (это - нижняя граница для ρ). Следовательно, для обеспечения расстояния (в метрике Колмогорова) до нормального распределения не более 0,01 для "неудачных" распределений необходимо не менее k_0 слагаемых, где

$$0{,}4\sqrt{k_0} < 0{,}01, \ \ k_0 > 1600.$$

В обычно используемых сумматорах слагаемых значительно меньше. Сужая класс возможных распределений *H*, можно получить, как показано в монографии [5], более быструю сходимость, но теория здесь еще не смыкается с практикой. Кроме того, не ясно, обеспечивает ли близость распределения к нормальному (в определенной метрике) также и близость распределения статистики, построенной по случайным величинам с этим распределением, к распределению статистики, соответствующей нормальным результатам наблюдений. Видимо, для каждой конкретной статистики необходимы специальные теоретические исследования, Именно к такому выводу приходит автор монографии [5]. В задачах отбраковки выбросов ответ: "Не обеспечивает" (см. ниже).

Отметим, что результат любого реального измерения записывается с помощью конечного числа десятичных знаков, обычно небольшого (2-5), так что любые реальные данные целесообразно моделировать лишь с помощью дискретных случайных величин, принимающих конечное число значений. Нормальное распределение - лишь аппроксимация реального распределения. Так, например, данные конкретного исследования, приведенные в работе [6], принимают значения от 1,0 до 2,2, т.е. всего 13 возможных значений. Из принципа Дирихле следует, что в какой-то точке построенная по данным работы [6] функция распределения отличается от ближайшей функции нормального распределения не менее чем на 1/26, т.е. на 0,04. Кроме того, очевидно, что для нормального распределения случайной величины вероятность попасть в дискретное множество десятичных чисел с заданным числом знаков после запятой равна 0.

Из сказанного выше следует, что результаты измерений и вообще статистические данные имеют свойства, приводящие к тому, что моделировать их следует случайными величинами с распределениями, более или менее отличными от нормальных. В большинстве случаев распределения существенно отличаются от нормальных, в других нормальные распределения могут, видимо, рассматриваться как некоторая аппроксимация, но никогда нет полного совпадения. Отсюда вытекает как необходимость изучения свойств классических статистических процедур в неклассических вероятностных моделях (подобно тому, как это сделано ниже для критерия Стьюдента), так и необходимость разработки устойчивых (учитывающих наличие отклонений от нормальности) и непараметрических, в том числе свободных от распределения процедур, их широкого внедрения в практику статистической обработки данных.

Опущенные здесь рассмотрения для других параметрических семейств приводят к аналогичным выводам. Итог можно сформулировать так. Распределения реальных данных практически никогда не входят в какое-либо конкретное параметрическое семейство. Реальные распределения всегда отличаются от тех, что включены в параметрические семейства. Отличия могут быть большие или маленькие, но они всегда есть. Попробуем понять, насколько важны эти различия для проведения эконометрического анализа.

4.2. Неустойчивость параметрических методов отбраковки резко выделяющихся результатов наблюдений

При обработки реальных экономических данных, полученных в процессе наблюдений, измерений, расчетов, иногда один или несколько результатов наблюдений резко выделяются, т.е. далеко отстоят от основной массы данных. Такие резко выделяющиеся результаты наблюдений часто считают содержащими грубые погрешности, соответственно называют промахами или выбросами. В рассматриваемых случаях возникает естественная мысль о том, что подобные наблюдения не относятся к изучаемой совокупности, поскольку содержат грубую погрешность, а получены в результате ошибки, промаха. В метрологии об этом явлении говорят так: "Грубые погрешности и промахи возникают из-за ошибок или неправильных действий оператора (его психо-физиологического состояния, неверного отсчета, ошибок в записях или вычислениях, неправильного включения приборов и т.п.), а также при кратковременных резких изменений проведения измерений (вибрации, поступления холодного воздуха, толчка прибора оператором и т.п.). Если грубые погрешности и промахи обнаруживают в процессе измерений, то результаты, содержащие их, отбрасывают. Однако чаще всего их выявляют только при окончательной обработке результатов измерений с помощью специальных критериев оценки грубых погрешностей" [7, с.46-47].

Есть два подхода к обработке данных, которые могут быть искажены грубыми погрешностями и промахами:

1) отбраковка резко выделяющихся результатов наблюдений, т.е. обнаружение наблюдений, искаженных грубыми погрешностями и промахами, и исключение их из дальнейшей статистической обработки;

2) применение устойчивых (робастных) методов обработки данных, На результаты работы которых мало влияет наличие небольшого числа грубо искаженных наблюдений (см. ниже соответствующую главу).

В настоящем пункте обсуждаются методы отбраковки.

Наиболее изучена ситуация, когда результаты наблюдений - числа $x_1., x_2., \ldots, x_n.$, резко выделяется один результат наблюдения, для определенности, максимальный $x_{\max}$.

Простейшая вероятностно-статистическая модель такова [8]. При нулевой гипотезе H_0 результаты наблюдения $x_1., x_2., \ldots, x_n$ рассматриваются как реализация независимых одинаково распределенных случайных величин числа $X_1., X_2., \ldots, X_n.$ с функцией распределения $F(x)$. При альтернативной гипотезе H_1 случайные величины $X_1., X_2., \ldots, X_n.$ также независимы, $X_1., X_2., \ldots, X_{n-1}$ имеют распределение $F(x)$, а X_n - распределение $G(x)$, оно "существенно сдвинуто вправо" относительно $F(x)$, например, $G(x)=F(x - A)$, где A достаточно велико. Если альтернативная гипотеза справедлива, то при $A \to \infty$ вероятность равенства

$$X_n = \max(X_1, X_2, \ldots, X_n)$$

стремится к 1, поэтому естественно применять решающее правило следующего вида:

если $x_{max}. > d$, то принять $H_1.$,

если $x_{max}. \leq d$, то принять H_0 , (1)

где d - параметр решающего правила, который следует определять из вероятностно-статистических соображений.

При справедливости нулевой гипотезы

$$P\{\max_{1 \leq i \leq n} X_i \leq d\} = \{F(d)\}^n.$$

Статистический критерий проверки гипотезы H_0 , основанный на решающем правиле вида (1), имеет уровень значимости α , если

$$P\{\max_{1 \leq i \leq n} X_i > d\} = 1 - \{F(d)\}^n = \alpha,$$

т.е.

$$F(d) = \sqrt[n]{1-\alpha}. \qquad (2)$$

Из соотношения (2) определяют граничное значение *d=d(*α*, n)* в решающем правиле (1).

При больших n и малых α

$$F(d)=\sqrt[n]{1-\alpha}=1-\frac{\alpha}{n}+O\left(\frac{\alpha^2}{n^2}\right), \qquad (3)$$

поэтому в качестве хорошего приближения к *d(*α*, n)* рассматривают (1-α/*n*) - квантиль распределения *F(x*).

Пусть правило отбраковки задано в соответствии с выражениями (1) и (2) с некоторой функцией распределения F, однако выборка берется из функции распределения *G*, мало отличающейся от *F* в смысле расстояния Колмогорова

$$\rho(F,G)=\sup_x|F(x)-G(x)|\le\delta. \qquad (4)$$

С помощью соотношения (3) получаем, что величина γ= *G(d)* для *d* из уравнения (2) находится между $\gamma_1=\max(0,\ 1-\frac{\alpha}{n}-\delta)$ и $\gamma_2=\min(1-\frac{\alpha}{n}+\delta,\ 1)$. Уровень значимости критерия, построенного для *F*, при применении к наблюдениям из *G* есть 1-γ^n и может принимать любые значения в отрезке [1-γ_2^n; 1-γ_1^n]. В частности, при δ = 0,01, α=0,05, *n* = 5 возможные значения уровня значимости заполняют отрезок [0; 0,1], т.е. уровень значимости может быть в 2 раза выше номинального, а если *n* возрастает до 30, то максимальный уровень значимости есть 0,297, т.е. почти в 6 раз выше номинального. При дальнейшем росте *n* верхняя граница для уровня значимости, как нетрудно видеть, приближается к 1.

Рассмотрим и другой вопрос - насколько правило отбраковки с уровнем значимости α для *G* может отличаться от такового для *F* при справедливости неравенства (4). С использованием соотношения (3) заключаем, что из

$$G(d)=1-\frac{\alpha}{n} \qquad (5)$$

следует, что $\gamma_1\le F(d)\le\gamma_2$, где γ_1 и γ_2 выписаны выше. Решение уравнения (5) может принимать любое значение в отрезке [$F^{-1}(\gamma_1)$; $F^{-1}(\gamma_2)$]. В частности, при α=0,05 и *n* = 5 для стандартного нормального распределения *F* имеем *d(*α*, n)* = 2,319, при δ=0,01 решение уравнения (5) может принимать любое значение в отрезке [2,054; + ∞], при δ =0,005 - любое значение в [2,170; 2,576].

При использовании любого другого расстояния между функциями распределения выводы о неустойчивости правил отбраковки также справедливы. Отметим, что проведенные рассмотрения выполнены в рамках "общей схемы устойчивости" (см. ниже главу об устойчивости статистических процедур).

Рассмотренные примеры показывают, что при конкретном значении δ= 0,01 в неравенстве (4) весьма неустойчивы как уровни значимости при фиксированном правиле отбраковки, так и параметр *d* правила отбраковки при фиксированном уровне значимости. Обсудим, насколько реалистично определение функции распределения с точностью $\delta\le 0{,}01$.

Есть два подхода к определению функции распределения результатов наблюдений: эвристический подбор с последующей проверкой с помощью критериев согласия и вывод из некоторой вероятностной модели.

Пусть с помощью критерия согласия Колмогорова проверяется гипотеза о том, что выборка взята из распределения *F*. Пусть функции распределения *F* и *G* удовлетворяют соотношению (4). Пусть на самом деле выборка взята из распределения *G*, а не *F*. При каких δ не удастся различить *F* и *G*? Для определенности, при каких δ гипотеза согласия с *F* будет приниматься не менее чем в 50% случаев?

Критерий согласия Колмогорова основан на статистике

$$\lambda_n = \sqrt{n}\rho(F_n, H), \qquad (6)$$

где расстояние ρ между функциями распределения определено выше в формуле (4); H - та функция распределения, согласие с которой проверяется, а F_n - эмпирическая функция распределения (т.е. $F_n(x)$ равно доле наблюдений, меньших x, в выборке объема n). Как показал А.Н. Колмогоров в 1933 г., функция распределения случайной величины λ_n при росте объема выборки n сходится к некоторой функции распределения $K(x)$, которую ныне называют функцией Колмогорова. При этом $K(1,36)= 0,95$ и $K(0,83)=0,50$.

Поскольку выборка взята из распределения G, то с вероятностью 0,50

$$\rho(F_n, G) < 0{,}83/\sqrt{n} \qquad (7)$$

(при больших n). Тогда для рассматриваемой выборки с учетом неравенства (4) и неравенства треугольника для расстояния Колмогорова и симметричности этого расстояния имеем

$$\rho(F_n, F) \le \rho(F_n, G) + \rho(G\ , F) = \rho(F_n, G) + \rho(F\ , G) < 0{,}83/\sqrt{n} + \delta.$$

Если

$$0{,}83/\sqrt{n} + \delta \le 1{,}36/\sqrt{n},$$

т.е.

$$\delta\sqrt{n} \le 0{,}53, \qquad (8)$$

то, согласно формуле (6), гипотеза согласия принимается по крайней мере с той же вероятностью, с которой выполнено неравенств (7), т.е. с вероятностью не менее 0,50. Для $\delta = 0{,}01$ это условие выполняется при $n \le 2809$. Таким образом, для определения функции распределения с точностью $\delta \le 0{,}01$ с помощью критерия согласия Колмогорова необходимо несколько тысяч наблюдений, что для большинства эконометрических задач нереально.

При втором из названных выше подходов к определению функции распределения ее конкретный вид выводится из некоторой системы аксиом, в частности, из некоторой модели порождения соответствующей случайной величины. Например, из модели суммирования вытекает нормальное распределение, а из мультипликативной модели перемножения - логарифмически нормальное распределение. Как правило, при выводе используется предельный переход. Так, из Центральной Предельной Теоремы теории вероятностей вытекает, что сумма независимых случайных величин может быть приближена нормальным распределением. Однако более детальный анализ, в частности, с помощью неравенства Берри-Эссеена (см. предыдущий пункт) показывает, что для гарантированного достижения точности $\delta \le 0{,}01$ необходимо более полутора тысяч слагаемых. Такого количества слагаемых реально, конечно, указать почти никогда нельзя. Это означает, что при решении практических эконометрических задач теория дает возможность лишь сформулировать гипотезу о виде функции распределения, а проверять ее надо с помощью анализа реальной выборки объема, как показано выше, не менее нескольких тысяч.

Таким образом, в большинстве реальных ситуаций определить функцию распределения с точностью $\delta \le 0{,}01$ невозможно.

Итак, показано, что правила отбраковки, основанные на использовании конкретной функции распределения, являются крайне неустойчивыми к отклонениям от нее распределения элементов выборки, а гарантировать отсутствие подобных отклонений невозможно. Поэтому отбраковка по классическим правилам математической статистики не является научно обоснованной, особенно при больших объемах выборок. Указанные правила целесообразно применять лишь для выявления "подозрительных" наблюдений, вопрос об отброаковке которых должен решаться из соображений соответствующей предметной области, а не из формально-математических соображений.

Выше для простоты изложения рассмотрен лишь случай полностью известного распределения F, для которого изучено правило отбраковки, заданное формулами (1) и (2). Аналогичные выводы о крайней неустойчивости правил отбраковки справедливы, если "истинное распределение" принадлежит какому-либо параметрическому семейству, например, нормальному, Вейбулла-Гнеденко, гамма.

Параметрическим методам отбраковки, основанным на моделях тех или иных параметрических семейств распределений, посвящены тысячи книг и статей. Приходится признать, что они имеют в основном внутриматематический интерес. При обработке реальных данных следует применять устойчивые методы (см. соответствующую главу), в частности, непараметрические.

4.3. Непараметрическое доверительное оценивание характеристик распределения

Пусть исходные данные –это выборка $x_1, x_2, \dots, x_n$, где n – объем выборки. Выборочные значения $x_1, x_2, \dots, x_n$ рассматриваются как реализации независимых одинаково распределенных случайных величин $X_1, X_2, \dots, X_n$ с общей функцией распределения $F(x) = P(X_i < x)$, $i = 1,2, \dots, n$. Поскольку функция распределения произвольна (с точностью до условий регулярности типа существования моментов), то рассматриваемые задачи доверительного оценивания характеристик распределения являются *непараметрическими*. Существование моментов является скорее математическим ограничением, чем реальным, поскольку практически все реальные статистические данные финитны (ограничены сверху и снизу, например, шкалой прибора).

В расчетах будут использоваться выборочное среднее арифметическое

$$M = (X_1 + X_2 + \dots + X_n) / n,$$

выборочная дисперсия

$$S^2 = \{ (X_1 - M)^2 + (X_2 - M)^2 + \dots + (X_n - M)^2 \} / (n-1)$$

и некоторые другие выборочные характеристики, которые мы введем позже.

Точечное и интервальное оценивание математического ожидания. Точечной оценкой для математического ожидания в силу закона больших чисел является выборочное среднее арифметическое M.

Нижняя доверительная граница для математического ожидания имеет вид

$$M - U(p)\, S / n^{1/2},$$

где:

M – выборочное среднее арифметическое,

p – доверительная вероятность (истинное значение математического ожидания находится между нижней доверительной границей и верхней доверительной границей с вероятностью, равной доверительной);

$U(p)$ – число, заданное равенством $\Phi(U(p)) = (1+ p)/2$, где $\Phi(x)$ – функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. Например, при p = 95% (т.е. при p = 0,95) имеем $U(p)$ = 1,96. Функция U(p) имеется в большинстве литературных источников по теории вероятностей и математической статистике (см., например, [8]);

S – выборочное среднее квадратическое отклонение (квадратный корень из описанной выше выборочной дисперсии).

Верхняя доверительная граница для математического ожидания имеет вид

$$M + U(p)\, S / n^{1/2}.$$

Выражения для верхней и нижней доверительных границ получены с помощью Центральной Предельной Теоремы теории вероятностей. Они являются асимптотическими, т.е. становятся тем точнее, чем больше объем выборки. В частности, вероятность попадания истинного значения математического ожидания между нижней и верхней доверительными границами асимптотически приближается к доверительной вероятности, но, вообще говоря, может отличаться от нее. Это – недостатки непараметрического подхода. Достоинством же является то, что его можно применять всегда, когда случайная величина имеет математическое ожидание и дисперсию, что в силу финитности (ограниченности шкал) имеет быть практически всегда в реальных ситуациях.

Интересно сопоставить с параметрическим подходом. Обычно в таких случаях предполагают нормальность результатов наблюдений (которой, как уже было обосновано в первом пункте настоящей главы, практически никогда нет). Тогда формулы для нижней и верхней доверительных границ для математического ожидания имеют похожий вид, только вместо *U(p)* стоят квантили распределению Стьюдента (а не нормального распределения, как в приведенных выше формулах), соответствующие объему выборки. Как известно, при росте объема выборки квантили распределения Стьюдента сходятся к соответствующим квантилям стандартного нормального распределения, так что при больших объемах выборок оба подхода дают близкие результаты. Отметим, что классические доверительные интервалы несколько длиннее, поскольку квантили распределения Стьюдента больше квантилей стандартного нормального распределения, хотя это различие, на наш взгляд, и невелико.

Точечное и интервальное оценивание медианы. В случае медианы по доверительной вероятности *p* находят *U(p)*, как разъяснено выше. Затем вычисляют натуральное число

$$C(p) = [n/2 - U(p)n^{1/2}/2]\ ,$$

где [.] – знак целой части числа. Нижняя доверительная граница для медианы имеет вид

$$X\,(C(p)),$$

где *X(i)* – член вариационного ряда с номером *i*, построенного по исходной выборке (т.е. *i*-я порядковая статистика). Верхняя доверительная граница для медианы имеет вид

$$X\,(n + 1 - C(p)).$$

Теоретическое основание для приведенных доверительных границ содержится в литературе по порядковым статистикам (см., например, монографию [9, с.68]).

Поскольку в случае нормального распределения медиана совпадает с математическим ожиданием, то каких-либо специальных способов ее оценивания в классическом случае нет.

Точечное и интервальное оценивание дисперсии. Точечной оценкой дисперсии является выборочная дисперсия S^2. Доверительные границы находятся с помощью величины

$$d^2 = (m_4 - ((n-1)/n)^4\ S^4)/n\ ,$$

где m_4 - выборочный четвертый центральный момент, т.е.

$$m_4 = \{(X_1 - M)^4 + (X_2 - M)^4 + \ldots + (X_n - M)^4\}/n\ .$$

Нижняя доверительная граница для дисперсии случайной величины имеет вид

$$S^2 - U(p)d\ ,$$

где S^2 – выборочная дисперсия,

U(*p*) – квантиль нормального распределения порядка *(1+p)/2* (как и раньше),

d – положительный квадратный корень из величины d^2, введенной выше.

Верхняя доверительная граница для дисперсии случайной величины имеет вид

$$S^2 + U(p)d\ ,$$

где все составляющие имеют тот же смысл, что и выше.

При выводе приведенных соотношений используется асимптотическая нормальность выборочной дисперсии, установленная, например, в [10, с.419]. Соответственно доверительный интервал является непараметрическим и асимптотическим. В классическом случае точечная оценка имеет тот же вид, а вот доверительные границы находят с помощью квантилей распределения хи-квадрат с числом степеней свободы, на 1 меньшим объема выборки. Отметим, что в случае нормального распределения четвертый момент в 3 раза больше квадрата дисперсии, а потому можно оценить d^2 как $(2 S^4) / n$. Это дает быстрый способ для интервальной оценки дисперсии в нормальном случае.

Точечное и интервальное оценивание среднего квадратического отклонения. Дисперсия рассматриваемой случайной величины - выборочного среднего квадратического отклонения S – оценивается как дробь

$$d^2 / (4 S^2) .$$

Нижняя доверительная граница для среднего квадратического отклонения исходной случайной величины имеет вид

$$S - U(p)d / (2S) ,$$

где S^2 – выборочная дисперсия,

$U(p)$ – квантиль нормального распределения порядка $(1+p)/2$ (как и раньше),

d – положительный квадратный корень из величины d^2, введенной выше.

Верхняя доверительная граница для среднего квадратического отклонения исходной случайной величины имеет вид

$$S + U(p)d / (2S) ,$$

где все составляющие имеют тот же смысл, что и выше.

Правила расчетов настоящего подпункта получены из правил предыдущего подпункта с помощью метода линеаризации (см., например, [11, п.2.4]). В рассматриваемом случае доверительный интервал также является непараметрическим и асимптотическим, а классический подход связан с использованием распределения хи-квадрат.

Точечное и интервальное оценивание коэффициента вариации. Коэффициент вариации широко используется при анализе конкретных экономических данных (поскольку они, как правило, положительны), но не очень популярен среди теоретиков. Дисперсия выборочного коэффициента вариации

$$V_n = S / M$$

оценивается с помощью вспомогательной величины

$$D^2 = (V_n^4 - V_n^2 / 4 + m_4 / (4 S^2 M^2) - m_3 / M^3) / n ,$$

где M – выборочное среднее арифметическое,

S^2 – выборочная дисперсия,

m_3 - выборочный третий центральный момент, т.е.

$$m_3 = \{ (X_1 - M)^3 + (X_2 - M)^3 + \ldots + (X_n - M)^3 \} / n ,$$

m_4 - выборочный четвертый центральный момент (см. выше),

V_n – выборочный коэффициент вариации,

n - объем выборки.

Нижняя доверительная граница для (теоретического) коэффициента вариации исходной случайной величины имеет вид

$$V_n - U(p) D,$$

где V_n – выборочный коэффициент вариации,

$U(p)$ – квантиль нормального распределения порядка $(1+p)/2$ (как и ранее),

D – положительный квадратный корень из величины D^2, введенной выше.

Верхняя доверительная граница для (теоретического) коэффициента вариации исходной случайной величины имеет вид

$$V_n + U(p) D,$$

где все составляющие имеют тот же смысл, что и выше.

Как и в предыдущих случаях, доверительный интервал является непараметрическим и асимптотическим. Он получен в результате применения специальной технологии вывода асимптотических соотношений прикладной статистики. Эта технология в качестве первого шага использует многомерную центральную предельную теорему, примененную к сумме векторов, координаты которых – степени исходных случайных величин. Второй шаг – преобразование предельного многомерного нормального вектора с целью получения интересующего исследователя вектора. При этом используются соображения линеаризации и отбрасываются бесконечно малые величины. Третий шаг – строгое обоснование полученных результатов на стандартном для асимптотических математико-статистических рассуждений уровне. При этом обычно оказывается необходимым использовать необходимые и достаточные условия наследования сходимости, полученные в монографии [11, п.2.4]. Именно таким образом были получены приведенные выше результаты для выборочного коэффициента вариации. Формулы оказались существенно более сложными, чем в предыдущих случаях. Это объясняется тем, что выборочный коэффициент вариации - функция двух выборочных моментов, а ранее рассматривались либо выборочные моменты поодиночке, либо функция от одного выборочного момента - выборочной дисперсии.

4.4. О проверке однородности двух независимых выборок

В математико-статистических терминах постановка задачи такова: имеются две выборки $x_1, x_2,...,x_m$ и $y_1, y_2,...,y_n$ (т. е. наборы из m и n действительных чисел), требуется проверить их однородность. Термин «однородность» уточняется ниже.

Противоположным понятием является «различие». Можно переформулировать задачу: требуется проверить, есть ли различие между выборками. Если различия нет, то для дальнейшего изучения часто выборки объединяют.

Например, в маркетинге важно выделить сегменты потребительского рынка. Если установлена однородность двух выборок, то возможно объединение сегментов, из которых они взяты, в один. В дальнейшем это позволит осуществлять по отношению к ним одинаковую маркетинговую политику (проводить одни и те же рекламные мероприятия и т.п.). Если же установлено различие, то поведение потребителей в двух сегментах различно, объединять эти сегменты нельзя, и могут понадобиться различные маркетинговые стратегии, своя для каждого из этих сегментов.

Традиционный метод проверки однородности (критерий Стьюдента). Для дальнейшего критического разбора опишем традиционный статистический метод проверки однородности. Вычисляют средние арифметические в каждой выборке

$$\bar{x} = \frac{1}{m}\sum_{1\le i\le m} x_i, \qquad y = \frac{1}{n}\sum_{1\le i\le n} y_i \ ,$$

затем выборочные дисперсии

$$s_x^2 = \frac{1}{m-1}\sum_{1\le i\le m}(x_i-\bar{x})^2 \ , \ s_y^2 = \frac{1}{n-1}\sum_{1\le i\le n}(y_i-\bar{y})^2$$

и статистику Стьюдента t, на основе которой принимают решение,

$$t = \frac{\bar{x}-\bar{y}}{\sqrt{(m-1)s_x^2 + (n-1)s_y^2}}\sqrt{\frac{mn(m+n-2)}{m+n}} \ . \quad (1)$$

По заданному уровню значимости α и числу степеней свободы $(m+n-2)$ из таблиц распределения Стьюдента находят критическое значение $t_{кр}$. Если $|t|>t_{кр}$, то гипотезу однородности (отсутствия различия) отклоняют, если же $|t|\le t_{кр}$, то принимают. (При односторонних альтернативных гипотезах вместо условия $|t|>t_{кр}$ проверяют, что $t>t_{кр}$; эту постановку рассматривать не будем, так как в ней нет принципиальных отличий от обсуждаемой здесь.)

Рассмотрим условия применимости традиционного метода проверки однородности, основанного на использовании статистики t Стьюдента, а также укажем более современные методы.

Вероятностная модель порождения данных. Для обоснованного применения эконометрических методов необходимо прежде всего построить и обосновать вероятностную модель порождения данных. При проверке однородности двух выборок общепринята модель, в которой $x_1, x_2,...,x_m$ рассматриваются как результаты m независимых наблюдений некоторой случайной величины X с функцией распределения $F(x)$, неизвестной статистику, а $y_1, y_2,...,y_n$ - как результаты n независимых наблюдений, вообще говоря, другой случайной величины Y с функцией распределения $G(x)$, также неизвестной статистику. Предполагается также, что наблюдения в одной выборке не зависят от наблюдений в другой, поэтому выборки и называют независимыми.

Возможность применения модели в конкретной реальной ситуации требует обоснования. Независимость и одинаковая распределенность результатов наблюдений, входящих в выборку, могут быть установлены или исходя из методики проведения конкретных наблюдений, или путем проверки статистических гипотез независимости и одинаковой распределенности с помощью соответствующих критериев [8].

Если проведено $(m+n)$ измерений объемов продаж в $(m+n)$ торговых точках, то описанную выше модель, как правило, можно применять. Если же, например, x_i и y_i - объемы продаж одного и того же товара до и после определенного рекламного воздействия, то рассматриваемую модель применять нельзя. (В этом случае используют модель т.н. связанных выборок, в которой обычно строят новую выборку $z_i = x_i - y_i$ и используют статистические методы анализа одной выборки, а не двух. Проверка однородности для связанных выборок рассматривается ниже.)

При дальнейшем изложении принимаем описанную выше вероятностную модель двух выборок.

Уточнения понятия однородности. Понятие «однородность», т. е. «отсутствие различия», может быть формализовано в терминах вероятностной модели различными способами.

Наивысшая степень однородности достигается, если обе выборки взяты из одной и той же генеральной совокупности, т. е. справедлива нулевая гипотеза

$$H_0 : F(x)=G(x) \text{ при всех } x.$$

Отсутствие однородности означает, что верна альтернативная гипотеза, согласно которой

$$H_1 : F(x_0)\neq G(x_0)$$

хотя бы при одном значении аргумента x_0. Если гипотеза H_0 принята, то выборки можно объединить в одну, если нет - то нельзя.

В некоторых случаях целесообразно проверять не совпадение функций распределения, а совпадение некоторых характеристик случайных величин X и Y - математических ожиданий, медиан, дисперсий, коэффициентов вариации и др. Например, однородность математических ожиданий означает, что справедлива гипотеза

$$H'_0 : M(X)=M(Y),$$

где $M(X)$ и $M(Y)$ - математические ожидания случайных величин X и Y, результаты наблюдений над которыми составляют первую и вторую выборки соответственно. Доказательство различия между выборками в рассматриваемом случае - это доказательство справедливости альтернативной гипотезы

$$H'_1 : M(X) \neq M(Y) .$$

Если гипотеза H_0 верна, то и гипотеза H'_0 верна, но из справедливости H'_0 не следует справедливость H_0. В частности, если в результате обработки выборочных данных принята гипотеза H'_0, то отсюда *не следует,* что две выборки можно объединить в одну. Однако в ряде ситуаций целесообразна проверка именно гипотезы H'_0. Например, пусть функция спроса на определенный товар или услугу оценивается путем опроса потребителей (первая выборка) или с помощью данных о продажах (вторая выборка). Тогда маркетологу важно проверить гипотезу об отсутствии систематических расхождений результатов этих двух методов, т.е. гипотезу о равенстве математических ожиданий. Другой пример – из производственного менеджмента. Пусть изучается эффективность управления бригадами рабочих на предприятии с помощью двух организационных схем, результаты наблюдения - объем производства на одного члена бригады, а показатель эффективности организационной схемы - средний (по предприятию) объем производства на одного рабочего. Тогда для сравнения эффективности препаратов достаточно проверить гипотезу H'_0.

Классические условия применимости критерия Стьюдента. Пусть выполнены два классических условия применимости критерия Стьюдента, основанного на использовании статистики t, заданной формулой (1):

а) результаты наблюдений имеют нормальные распределения:

$$F(x)=N(x;\ m_1,\ \sigma_1^2),\ G(x)=N(x;\ m_2,\ \sigma_2^2)$$

с математическими ожиданиями m_1 и m_2 и дисперсиями σ_1^2 и σ_2^2 в первой и во второй выборках соответственно;

б) дисперсии результатов наблюдений в первой и второй выборках совпадают:

$$D(X)=\sigma_1^2=D(Y)=\sigma_2^2.$$

Если условия а) и б) выполнены, то нормальные распределения $F(x)$ и $G(x)$ отличаются только математическими ожиданиями, а поэтому *обе* гипотезы H_0 и H'_0 сводятся к гипотезе

$$H''_0 : m_1=m_2, ,$$

а *обе* альтернативные гипотезы H_1 и H'_1 сводятся к гипотезе

$$H''_1 : m_1 \neq m_2, .$$

Если условия а) и б) выполнены, то статистика t при справедливости H''_0 имеет распределение Стьюдента с $(m + n - 2)$ степенями свободы. Только в этом случае описанный выше традиционный метод обоснован безупречно. Если хотя бы одно из условий а) и б) не выполнено, то нет оснований считать, что статистика t имеет распределение Стьюдента, поэтому применение традиционного метода, строго говоря, не обосновано. Обсудим возможность проверки этих условий и последствия их нарушений.

О проверке условия нормальности. Априори нет оснований предполагать нормальность распределения результатов экономических, технико-экономических и иных наблюдений. Следовательно, нормальность надо проверять. Разработано много статистических критериев для проверки нормальности распределения результатов наблюдений [8]. Однако проверка нормальности - более сложная и трудоемкая статистическая процедура, чем проверка однородности (как с помощью статистики t Стьюдента, так и с использованием непараметрических критериев, рассматриваемых ниже).

Для достаточно надежного установления нормальности требуется весьма большое число наблюдений. Выше показано, что для того, чтобы гарантировать, что функция распределения результатов наблюдений отличается от некоторой нормальной не более чем на 0,01 (при любом значении аргумента), требуется порядка 2500 наблюдений. В большинстве экономических и технико-экономических исследований число наблюдений существенно меньше.

Как уже отмечалось, есть и одна общая причина отклонений от нормальности: любой результат наблюдения записывается конечным (обычно 2-5) количеством цифр, а с математической точки зрения вероятность такого события равна 0. Из сказанного выше следует, что в эконометрике распределение результатов экономических и технико-экономических наблюдений практически всегда более или менее отличается от нормального. Более подробно это утверждение выше.

Последствия нарушения условия нормальности. Если условие а) не выполнено, то распределение статистики t не является распределением Стьюдента. Однако при справедливости H'_0 и условии б) распределение статистики t при росте объемов выборок приближается к стандартному нормальному распределению $\Phi(x)=N(x;\ 0,\ 1)$. К этому же распределению приближается распределение Стьюдента при возрастании числа степеней свободы. Другими словами, несмотря на нарушение условия нормальности традиционный метод (критерий Стьюдента) можно использовать для проверки гипотезы H'_0 при больших объемах выборок. При этом вместо таблиц распределения Стьюдента достаточно пользоваться таблицами стандартного нормального распределения $\Phi(x)$.

Сформулированное в предыдущем абзаце утверждение справедливо для любых функций распределения $F(x)$ и $G(x)$ таких, что $M(X)=M(Y)$, $D(X)=D(Y)$ и выполнены некоторые внутриматематические условия, обычно считающиеся справедливыми в реальных задачах. Если же $M(X)\neq M(Y)$, то нетрудно вычислить, что при больших объемах выборок

$$P(t\leq x)\approx\Phi(x-a_{mn}), \qquad (2)$$

где

$$a_{mn}=\frac{\sqrt{mn}[M(X)-M(Y)]}{\sqrt{mD(X)+nD(Y)}}. \qquad (3)$$

Формулы (2) - (3) позволяют приближенно вычислять мощность t-критерия (точность возрастает при увеличении m и n).

О проверке условия равенства дисперсий. Иногда условие б) вытекает из методики получения результатов наблюдений, например, когда с помощью одного и того же прибора или методики m раз измеряют характеристику первого объекта и n раз-второго, а параметры распределения погрешностей измерения при этом не меняются. Однако ясно, что в постановках большинства исследовательских и практических задач нет основании априори предполагать равенство дисперсий.

Целесообразно ли проверять равенство дисперсий статистическими методами, например, как это иногда предлагают, с помощью F-критерия Фишера? Этот критерий основан на нормальности распределений результатов наблюдений, от которой неизбежны отклонения (см. выше), причем хорошо известно, что в отличие от t-критерия его распределение сильно меняется при малых отклонениях от нормальности [10]. Кроме того, F-критерий отвергает гипотезу $D(X)=D(Y)$ лишь при большом различии выборочных дисперсий. Так, для данных [8] о двух группах результатов химических анализов отношение выборочных дисперсий равно 1,95, т.е. существенно отличается от 1. Тем не менее гипотеза о равенстве теоретических дисперсий принимается на 1% уровне значимости. Следовательно, при проверке однородности применение F-критерия для предварительной проверки равенства дисперсий нецелесообразно.

Итак, в большинстве экономических и технико-экономических задач условие б) нельзя считать выполненным, а проверять его нецелесообразно.

Последствия нарушения условия равенства дисперсий. Если объемы выборок m и n велики, то можно показать, что распределение статистики t описывается с помощью только математических ожиданий $M(X)$ и $M(Y)$, дисперсий $D(X)$, $D(Y)$ и отношения объемов выборок, а именно:

$$P(t \leq x) \approx \Phi(b_{mn}x - a_{mn}), \qquad (4)$$

где a_{mn} определено формулой (3),

$$b_{mn}^{2} = \frac{\lambda D(X) + D(Y)}{D(X) + \lambda D(Y)}, \quad \lambda = \frac{m}{n}. \qquad (5)$$

Если $b_{mn} \neq 1$, то распределение статистики t отличается от распределения, заданного формулой (2), полученной в предположении равенства дисперсий. Когда $b_{mn}=1$? В двух случаях - при $m = n$ и при $D(X) = D(Y)$. Таким образом, при больших и равных объемах выборок требовать выполнения условия б) нет необходимости. Кроме того, ясно, что если объемы выборок мало различаются, то b_{mn} близко к 1. Так, для данных [8] имеем b^*_{mn}= 0,987, где b^*_{mn} - оценка b_{mn}, полученная заменой в формуле (5) теоретических дисперсий на выборочные.

Область применимости традиционного метода проверки однородности с помощью критерия Стьюдента. Подведем итоги рассмотрения t-критерия. Он позволяет проверять гипотезу H'_0 о равенстве математических ожиданий, но не гипотезу H_0 о том, что обе выборки взяты из одной и той же генеральной совокупности. Классические условия применимости критерия Стьюдента в подавляющем большинстве экономических и технико-экономических задач не выполнены. Тем не менее при больших и примерно равных объемах выборок его можно применять. При конечных объемах выборок традиционный метод носит неустранимо приближенный характер.

Критерий Крамера-Уэлча равенства математических ожиданий

Вместо критерия Стьюдента предлагаем для проверки H'_0 использовать критерий Крамера-Уэлча [12], основанный на статистике

$$T = \frac{\sqrt{mn}(\bar{x} - \bar{y})}{\sqrt{ns_x^2 + ms_y^2}}. \qquad (6)$$

Критерий Крамера-Уэлча имеет прозрачный смысл – разность выборочных средних арифметических для двух выборок делится на естественную оценку среднего квадратического отклонения этой разности. Естественность указанной оценки состоит в том, что неизвестные статистику дисперсии заменены их выборочными оценками. Из многомерной центральной предельной теоремы и из теорем о наследовании сходимости [11] вытекает, что при росте объемов выборок распределение статистики T Крамера-Уэлча сходится к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1. Итак, при справедливости H'_0 и больших объемах выборок распределение статистики T приближается с помощью стандартного нормального распределения $\Phi(x)$, из таблиц которого следует брать критические значения.

При $m=n$, как следует из формул (1) и (6), $t=T$. При $m \neq n$ этого равенства нет. В частности, при s_x^2 в (1) стоит множитель $(m-1)$, а в (6)- множитель n.

Если $M(X) \neq M(Y)$, то при больших объемах выборок

$$P(T \leq X) \approx \Phi(x - c_{mn}), \qquad (7)$$

где

$$c_{mn} = \frac{\sqrt{mn}[M(X) - M(Y)]}{\sqrt{nD(X) + mD(Y)}}. \qquad (8)$$

При $m=n$ или $D(X)=D(Y)$, согласно формулам (3) и (8), $a_{mn}=c_{mn}$, в остальных случаях равенства нет.

Из асимптотической нормальности статистики T, формул (7) и (8) следует, что правило принятия решения для критерия Крамера-Уэлча выглядит так:

- если $|T| \leq \Phi(1-\frac{\alpha}{2})$, то гипотеза однородности (равенства) математических ожиданий принимается на уровне значимости α,
- если же $|T| > \Phi(1-\frac{\alpha}{2})$, то гипотеза однородности (равенства) математических ожиданий отклоняется на уровне значимости α.

В эконометрике наиболее часто применяется уровень значимости $\alpha = 0{,}05$. Тогда значение модуля статистики T Крамера-Уэлча надо сравнивать с граничным значением $\Phi(1-\frac{\alpha}{2}) = 1{,}96$.

Из сказанного выше следует, что применение критерия Крамера-Уэлча не менее обосновано, чем применение критерия Стьюдента. Дополнительное преимущество - не требуется равенства дисперсий $D(X)=D(Y)$. Распределение статистики T не является распределением Стьюдента, однако и распределение статистики t, как показано выше, не является таковым в реальных ситуациях.

Распределение статистики T при объемах выборок $m=n=6, 8, 10,$ 12 и различных функциях распределений выборок F(x) и G(x) изучено нами совместно с Ю.Э. Камнем и Я.Э. Камнем методом статистических испытаний (Монте-Карло). Рассмотрены различные варианты функций распределения $F(x)$ и $G(x)$. Результаты показывают, что даже при таких небольших объемах выборок точность аппроксимации предельным стандартным нормальным распределением вполне удовлетворительна. Поэтому представляется целесообразным во всех тех случаях, когда в настоящее время используется критерий Стьюдента, заменить его на критерий Крамера-Уэлча. Конечно, такая замена потребует переделки ряда нормативно-технических и методических документов, исправления учебников и учебных пособий для вузов.

Пример. Пусть объем первой выборки $m = 120, \overline{x} = 13{,}7, s_x = 5{,}3.$ Для второй выборки $n = 541, \overline{y} = 14{,}1, s_y = 8{,}4.$ Вычислим величину статистики Крамера-Уэлча

$$T = \frac{\sqrt{mn}(\overline{x}-\overline{y})}{\sqrt{ns_x^2 + ms_y^2}} = \frac{\sqrt{120 \times 541}(13{,}7-14{,}1)}{\sqrt{541 \times 5{,}3^2 + 120 \times 8{,}4^2}} = \frac{\sqrt{64920}(-0{,}4)}{\sqrt{541 \times 28{,}09 + 120 \times 141{,}12}} =$$

$$= \frac{254{,}79 \times (-0{,}4)}{\sqrt{15196{,}69 + 16934.4}} = \frac{-101{,}916}{\sqrt{32131{,}09}} = \frac{-101{,}916}{179{,}25} = -0{,}57.$$

Поскольку полученное значение по абсолютной величине меньше 1,96, то гипотеза однородности математических ожиданий принимается на уровне значимости 0,05.

Непараметрические методы проверки однородности. В большинстве экономических и технико-экономических задач представляет интерес не проверка равенства математических ожиданий или иных характеристик распределения, а обнаружение различия генеральных совокупностей, из которых извлечены выборки, т.е. проверка гипотезы H_0. Методы проверки гипотезы H_0 позволяют обнаружить не только изменение математического ожидания, но и любые иные изменения функции распределения результатов наблюдений при переходе от одной выборки к другой (увеличение разброса, появление асимметрии и т. д.). Как установлено выше, методы, основанные на использовании статистик t Стьюдента и T Крамера-Уэлча, не позволяют проверять гипотезу H_0 . Априорное предположение о принадлежности функций распределения $F(x)$ и $G(x)$ к какому-либо определенному параметрическому семейству (например, семействам нормальных, логарифмически нормальных, распределений Вейбулла-Гнеденко, гамма-распределений и др.), как показано выше, обычно нельзя достаточно надежно обосновать. Поэтому для проверки H_0 следует использовать методы, пригодные при любом виде $F(x)$ и $G(x)$, т.е. непараметрические методы. (Термин «непараметрический метод» означает, что при использовании этого метода нет необходимости предполагать, что функции распределения результатов наблюдений принадлежат какому-либо определенному параметрическому семейству.)

Для проверки гипотезы H_0 разработано много непараметрических методов - критерии Смирнова, типа омега-квадрат (Лемана-Розенблатта), Вилкоксона (Манна-Уитни), Ван-дер-Вардена, Сэвиджа, хи-квадрат и др. [8, 9, 13]. Распределения статистик всех этих критериев при справедливости H_0 не зависят от конкретного вида совпадающих функций распределения $F(x)\equiv G(x)$. Следовательно, таблицами точных и предельных (при больших объемах выборок) распределений статистик этих критериев и их процентных точек [8, 9] можно пользоваться при любых непрерывных функциях распределения результатов наблюдений.

Каким из непараметрических критериев пользоваться? Как известно [10], для выбора одного из нескольких критериев необходимо сравнить их мощности, определяемые видом альтернативных гипотез. Сравнению мощностей критериев посвящена обширная литература.

Хорошо изучены свойства критериев при альтернативной гипотезе сдвига

$$H_{1c} : G(x)=F(x-d),\ d\neq 0.$$

Критерии Вилкоксона, Ван-дер-Вардена и ряд других ориентированы для применения именно в этой ситуации. Если m раз измеряют характеристику одного объекта и n раз - другого, а функция распределения погрешностей измерения произвольна, но не меняется при переходе от объекта к объекту (это более жесткое требование, чем условие равенства дисперсий), то рассмотрение гипотезы H_{1c} оправдано. Однако в большинстве экономических и технико-экономических исследований нет оснований считать, что функции распределения, соответствующие выборкам, различаются только сдвигом.

4.5. Какие гипотезы можно проверять с помощью двухвыборочного критерия Вилкоксона?

Покажем (и это - основной результат настоящего пункта), что двухвыборочный критерий Вилкоксона (в литературе его называют также критерием Манна-Уитни) предназначен для проверки гипотезы

$$H_0 :\ P(X < Y) = 1/2,$$

где X - случайная величина, распределенная как элементы первой выборки, а Y - второй.

В описанной выше вероятностной модели двух независимых выборок без ограничения общности можно считать, что объем первой из них не превосходит объема второй, $m \leq n$, в противном случае выборки можно поменять местами. Обычно предполагается, что функции $F(x)$ и $G(x)$ непрерывны и строго возрастают. Из непрерывности этих функций следует, что с вероятностью 1 все $m + n$ результатов наблюдений различны. В реальных эконометрических данных иногда встречаются совпадения, но сам факт их наличия - свидетельство нарушений предпосылок только что описанной базовой математической модели.

Статистика S двухвыборочного критерия Вилкоксона определяется следующим образом. Все элементы объединенной выборки $X_1, X_2, ..., X_m, Y_1, Y_2, ..., Y_n$ упорядочиваются в порядке возрастания. Элементы первой выборки $X_1, X_2, ..., X_m$ занимают в общем вариационном ряду места с номерами $R_1, R_2, ..., R_m$, другими словами, имеют ранги $R_1, R_2, ..., R_m$. Тогда статистика Вилкоксона - это сумма рангов элементов первой выборки

$$S = R_1 + R_2 + ... + R_m .$$

Статистика U Манна-Уитни определяется как число пар (X_i, Y_j) таких, что $X_i < Y_j$, среди всех mn пар, в которых первый элемент - из первой выборки, а второй - из второй. Как известно [13, с.160],

$$U = mn + m(m+1)/2 - S .$$

Поскольку S и U линейно связаны, то часто говорят не о двух критериях - Вилкоксона и Манна-Уитни, а об одном - критерии Вилкоксона (Манна-Уитни).

Критерий Вилкоксона - один из самых известных инструментов непараметрической статистики (наряду со статистиками типа Колмогорова-Смирнова и коэффициентами ранговой корреляции). Свойствам этого критерия и таблицам его критических значений уделяется место во многих монографиях по математической и прикладной статистике (см., например, [8, 9, 13]).

Однако в литературе имеются и неточные утверждения относительно возможностей критерия Вилкоксона. Так, одни полагают, что с его помощью можно обнаружить любое различие между функциями распределения $F(x)$ и $G(x)$. По мнению других, этот критерий нацелен на проверку равенства медиан распределений, соответствующих выборкам. И то, и другое, строго говоря, неверно. Это будет ясно из дальнейшего изложения.

Введем некоторые обозначения. Пусть $F^{-1}(t)$ - функция, обратная к функции распределения $F(x)$. Она определена на отрезке [0;1]. Положим $L(t) = G(F^{-1}(t))$. Поскольку $F(x)$ непрерывна и строго возрастает, то $F^{-1}(t)$ и $L(t)$ обладают теми же свойствами. Важную роль в дальнейшем изложении будет играть величина $a = P(X< Y)$. Как нетрудно показать,

$$a = P(X < Y) = \int_0^1 t\,dL(t).$$

Введем также параметры

$$b^2 = \int_0^1 L^2(t)dt - (1-a)^2, \quad g^2 = \int_0^1 t^2 dL(t) - a^2.$$

Тогда математические ожидания и дисперсии статистик Вилкоксона и Манна-Уитни согласно [13, с.160] выражаются через введенные величины:

$$M(U) = mna\ ,\ M(S) = mn + m(m+1)/2 - M(U) = mn(1- a) + m(m+1)/2,$$

$$D(S) = D(U) = mn\ [\ (n - 1)\ b^2 + (m - 1)\ g^2 + a(1 -a)\ ] \quad . (1)$$

Когда объемы обеих выборок безгранично растут, распределения статистик Вилкоксона и Манна-Уитни являются асимптотически нормальными (см., например, [13, гл.5 и 6]) с параметрами, задаваемыми формулами (1) .

Если выборки полностью однородны, т.е. их функции распределения совпадают, справедлива гипотеза

$$H_0: \quad F(x) = G(x) \text{ при всех } x, \qquad (2)$$

то $L(t) = t$ и $a= 1/2$. Подставляя в формулы (1), получаем, что

$$M(S) = m(m+n+1)/2, \quad D(S) = mn(m+n+1)/ 12 \quad (3) .$$

Следовательно, распределение нормированной и центрированной статистики Вилкоксона

$$T = (\ S - m(m+n+1)/2)\ (mn(m+n+1)/ 12\ )^{-1/2} \quad (4)$$

при росте объемов выборок приближается к стандартному нормальному распределению (с математическим ожиданием 0 и дисперсией 1).

Из асимптотической нормальности статистики T следует, что правило принятия решения для критерия Вилкоксона выглядит так:

- если $|T| \leq \Phi(1-\frac{\alpha}{2})$, то гипотеза (2) однородности (тождества) функций распределений принимается на уровне значимости α,
- если же $|T| > \Phi(1-\frac{\alpha}{2})$, то гипотеза (2) однородности (тождества) функций распределений отклоняется на уровне значимости α.

В эконометрике наиболее часто применяется уровень значимости $\alpha = 0{,}05$. Тогда значение модуля статистики T Вилкоксона надо сравнивать с граничным значением $\Phi(1-\frac{\alpha}{2}) = 1{,}96.$

Пример 1. Пусть даны две выборки. Первая содержит *m= 12* элементов 17; 22; 3; 5; 15; 2; 0; 7; 13; 97; 66; 14. Вторая содержит *n=14* элементов 47; 30; 2; 15; 1; 21; 25; 7; 44; 29; 33; 11; 6; 15. Проведем проверку однородности функций распределения двух выборок с помощью только что сформулированного правила принятия решений на основе критерия Вилкоксона.

Первым шагом является построение общего вариационного ряда для элементов двух выборок (табл.1).

Табл.1. Общий вариационный ряд для элементов двух выборок

Ранги	1	2	3,5	3,5	5	6	7	8,5	8,5	10	11	12	14
Элементы выборок	0	1	2	2	3	5	6	7	7	11	13	14	15
Номера выборок	1	2	1	2	1	1	2	1	2	2	1	1	1
Ранги	14	14	16	17	18	19	20	21	22	23	24	25	26
Элементы выборок	15	15	17	21	22	25	29	30	33	44	47	66	97
Номера выборок	2	2	1	2	1	2	2	2	2	2	2	1	1

Хотя с точки зрения теории математической статистики вероятность совпадения двух элементов выборок равна 0, в реальных выборках экономических данных совпадения встречаются. Так, в рассматриваемых выборках, как видно из табл.1, два раза повторяется величина 2, два раза - величина 7 и три раза - величина 15. В таких случаях говорят о наличии "связанных рангов", а соответствующим совпадающим величинам приписывают среднее арифметическое тех рангов которые они занимают. Так, величины 2 и 2 занимают в объединенной выборке места 3 и 4, поэтому им приписывается ранг (3+4)/2=3,5. Величины 7 и 7 занимают в объединенной выборке места 8 и 9, поэтому им приписывается ранг (8+9)/2=8,5. Величины 15, 15 и 15 занимают в объединенной выборке места 13, 14 и 15, поэтому им приписывается ранг (13+14+15)/3=14.

Следующий шаг - подсчет значения статистики Вилкоксона, т.е. суммы рангов элементов первой выборки

$$S = R_1 + R_2 + \ldots + R_m = 1+3{,}5+5+6+8{,}5+11+12+14+16+18+25+26=146.$$

Подсчитаем также сумму рангов элементов второй выборки

$$S_1 = 2+3{,}5+7+8{,}5+10+14+14+17+19+20+21+22+23+24= 205.$$

Величина S_1 может быть использована для контроля вычислений. Дело в том, что суммы рангов элементов первой выборки S и второй выборки S_1 вместе составляют сумму рангов объединенной выборки, т.е. сумму всех натуральных чисел от 1 до $m+n$. Следовательно,

$$S+ S_1 = (m+n)(m+n+1)/2= (12+14)(12+14+1)/2= 351.$$

В соответствии с ранее проведенными расчетами $S+S_1 = 146+205=351$. Необходимое условие правильности расчетов выполнено. Ясно, что справедливость этого условия не гарантирует правильности расчетов.

Перейдем к расчету статистики T. Согласно формуле (3)

$$M(S) = 12(12+14+1)/ 2 = 162,\ D(S) = 12 \cdot 14(12+14+1)/ 12= 378 .$$

Следовательно,

$$T = (S - 162) (378)^{- 1/2} = (146-162) / 19{,}44 = - 0.82.$$

Поскольку $|T| \leq 1{,}96$, то гипотеза однородности принимается на уровне значимости0,05.

Что будет, если поменять выборки местами, вторую назвать первой? Тогда вместо S надо рассматривать S_1 . Имеем

$$M(S_1) = 14(12+14+1)/ 2 = 189,\ D(S) = D(S_1) = 378 ,$$

$$T_1 = (S_1 - 189) (378)^{- 1/2} = (205-162) / 19{,}44 = 0.82.$$

Таким образом, значения статистики критерия отличаются только знаком (можно показать, что это утверждение верно всегда). Поскольку в правиле принятия решения используется только абсолютная величина статистики, то принимаемое решение не зависит от того, какую выборку считаем первой, а какую второй. Для уменьшения объема таблиц принято считать первой выборку меньшего объема.

Продолжим обсуждение критерия Вилкоксона. Правила принятия решений и таблица критических значений для критерия Вилкоксона строятся в предположении справедливости гипотезы полной однородности, описываемой формулой (2). А что будет, если эта гипотеза неверна? Другими словами, какова мощность критерия Вилкоксона?

Пусть объемы выборок достаточно велики, так что можно пользоваться асимптотической нормальностью статистики Вилкоксона. Тогда в соответствии с формулами (1) статистика Т будет асимптотически нормальна с параметрами

$$M(T) = (12mn)^{1/2} (1/2 - a) (m+n+1)^{- 1/2} ,$$

$$D(T) = 12 [(n - 1) b^2 + (m - 1) g^2 + a(1 -a)] (m+n+1)^{- 1} . \quad (5)$$

Из формул (5) видно большое значение гипотезы

$$H_{01}:\ a = P(X < Y) = 1/2 . \quad (6)$$

Если эта гипотеза неверна, то, поскольку $m \leq n$, справедлива оценка

$$|M(T)| \geq (12m\, n\, (2n+1)^{- 1})^{1/2} |1/2 - a| ,$$

а потому |E(T)| безгранично растет при росте объемов выборок. В то же время, поскольку

$$b^2 \leq \int_0^1 L^2(t)dt \leq 1,\quad g^2 \leq \int_0^1 t^2 dL(t) \leq 1,\quad \alpha(1-\alpha) \leq 1/4,$$

то

$$D(T) \leq 12 [(n - 1) + (m - 1) + 1/4] (m+n+1)^{- 1} \leq 12. \quad (7)$$

Следовательно, вероятность отклонения гипотезы H_{01} , когда она неверна, т.е. мощность критерия Вилкоксона как критерия проверки гипотезы (6), стремится к 1 при возрастании объемов выборок, т.е. критерий Вилкоксона является состоятельным для этой гипотезы при альтернативе

$$AH_{01}:\ a = P(X < Y) \neq 1/2 . \quad (8) .$$

Если же гипотеза (6) верна, то статистика Т асимптотически нормальна с математическим ожиданием 0 и дисперсией, определяемой формулой

$$D(T) = 12\ [(n - 1)\ b^2 + (m - 1)\ g^2 + 1/4\ ]\ (m+n+1)^{-1} . \quad (9)$$

Гипотеза (6) является сложной, дисперсия (9), как показывают приводимые ниже примеры, в зависимости от значений b^2 и g^2 может быть как больше 1, так и меньше 1, но согласно неравенству (7) никогда не превосходит 12.

Приведем пример двух функций распределения $F(x)$ и $G(x)$ таких, что гипотеза (6) выполнена, а гипотеза (2) - нет. Поскольку

$$a = P(X < Y) = \int_{-\infty}^{+\infty} F(x)dG(x)\ ,\ \ 1 - a = P(Y < X) = \int_{-\infty}^{+\infty} G(x)dF(x) \quad (10)$$

и $a = 1/2$ в случае справедливости гипотезы (2), то для выполнения условия (6) необходимо и достаточно, чтобы

$$\int_{-\infty}^{+\infty} (F(x) - G(x))dF(x) = 0 \quad (11)\ ,$$

а потому естественно в качестве $F(x)$ рассмотреть функцию равномерного распределения на интервале (-1 ; 1). Тогда формула (11) переходит в условие

$$\int_{-\infty}^{+\infty} (F(x) - G(x))dF(x) = -\frac{1}{2}\int_{-1}^{+1}\left(G(x) - \frac{(x+1)}{2} \right)dx = 0. \quad (11)\ .$$

Это условие выполняется, если функция $(G(x) - (x + 1)/2\ )$ является нечетной.

Пример 2. Пусть функции распределения $F(x)$ и $G(x)$ сосредоточены на интервале (-1 ; 1), на котором

$$F(x) = (x + 1)/2\ ,\ \ G(x) = (\ x + 1 + 1/\pi\ \sin\ \pi x\ )\ /\ 2\ .$$

Тогда

$$x=F^{-1}(t)=2\ -1,\ L(t)=G(F^{-1}(t))=(2t+1/\pi \sin\pi\ (2t-1))/2=t+1/2\,\pi \sin\pi\ (2t-1)\ .$$

Условие (11) выполнено, поскольку функция $(G(x) - (x + 1)/2)$ является нечетной. Следовательно, $a = 1/2$. Начнем с вычисления

$$g^2 = \int_0^1 t^2 dL(t)\ - 1/4 = \int_0^1 t^2 d(t + \frac{1}{2\pi}\sin\pi(2t-1)) - \frac{1}{4}.$$

Поскольку

$$d(t + \frac{1}{2\pi}\sin\pi(2t-1)) = (1+\cos\pi(2t-1))dt,$$

то

$$g^2 = \int_0^1 t^2(1+\cos\pi(2t-1))dt - \frac{1}{4} = \frac{1}{12} + \int_0^1 t^2\cos\pi(2t-1)dt.$$

С помощью замены переменных $t = (x +1)\ /\ 2$ получаем, что

$$\int_0^1 t^2\cos\pi(2t-1)dt = \frac{1}{8}\left(\int_{-1}^1 x^2\cos\pi x dx + 2\int_{-1}^1 x\cos\pi x dx + \int_{-1}^1 \cos\pi x dx \right).$$

В правой части последнего равенства стоят табличные интегралы (см., например, справочник [14, с.71]. Проведя соответствующие вычисления, получаем, что в правой части стоит $1/8\ (-4/\pi^2) = -1/(2\pi^2)$. Следовательно,

$$g^2 = 1/12 - 1/(2\pi^2) = 0{,}032672733...$$

Перейдем к вычислению b^2. Поскольку

$$b^2 = \int_0^1 L^2(t)dt - \frac{1}{4} = \int_0^1 \left(t + \frac{1}{2}\pi \sin \pi(2t-1)\right)^2 dt - \frac{1}{4},$$

то

$$b^2 = \frac{1}{12} + \frac{1}{\pi}\int_0^1 (t \sin \pi(2t-1))dt + \left(\frac{\pi}{2}\right)^2 \int_0^1 \sin^2 \pi(2t-1)dt.$$

С помощью замены переменных $t = (x+1)/2$ переходим к табличным интегралам (см., например, справочник [14, с.65]):

$$b^2 = \frac{1}{12} + \frac{1}{4\pi}\int_{-1}^1 x \sin \pi x dx + \frac{1}{4\pi}\int_{-1}^1 \sin \pi x dx + \frac{1}{8\pi^2}\int_{-1}^1 \sin^2 \pi x dx.$$

Проведя необходимые вычисления, получим, что

$$b^2 = \frac{1}{12} + \frac{1}{4\pi}\left(-\frac{2}{\pi}\right) + 0 + \frac{1}{8\pi^2} = \frac{1}{12} - \frac{3}{8\pi^2} = 0{,}045337893...$$

Следовательно, для рассматриваемых функций распределения нормированная и центрированная статистика Вилкоксона (см. формулу (4)) асимптотически нормальна с математическим ожиданием 0 и дисперсией (см. формулу (9))

$$D(T) = (0{,}544\,n + 0{,}392\,m + 2{,}064)\,(m+n+1)^{-1}\ .$$

Как легко видеть, дисперсия всегда меньше 1. Это значит, что в рассматриваемом случае гипотеза полной однородности (2) при проверке с помощью критерия Вилкоксона будет приниматься чаще, чем если она на самом деле верна.

На наш взгляд, это означает, что критерий Вилкоксона нельзя считать критерием для проверки гипотезы (2) при альтернативе общего вида. Он не всегда позволяет проверить однородность - не при всех альтернативах. Точно так же критерии типа хи-квадрат нельзя считать критериями проверки гипотез согласия и однородности - они позволяют обнаружить не все различия, поскольку некоторые из них "скрадывает" группировка.

Обсудим теперь, действительно ли критерий Вилкоксона нацелен на проверку равенства медиан распределений, соответствующих выборкам.

Пример 3. Построим семейство пар функций распределения $F(x)$ и $G(x)$ таких, что их медианы различны, но для $F(x)$ и $G(x)$ выполнена гипотеза (6). Пусть распределения сосредоточены на интервале (0 ; 1), и на нем $G(x) = x$, а $F(x)$ имеет кусочно-линейный график с вершинами в точках (0 ; 0), (λ , 1/2), (δ , 3/4), (1 ; 1). Следовательно,

$F(x) = 0$ при $x < 0$;

$F(x) = x / (2\lambda)$ на $[0 ; \lambda)$;

$F(x) = 1/2 + (x - \lambda) / (4\delta - 4\lambda)$ на $[\lambda ; \delta)$;

$F(x) = 3/4 + (x - \delta) / (4 - 4\delta)$ на $[\delta ; 1]$;

$F(x) = 1$ при $x > 1$.

Очевидно, что медиана $F(x)$ равна λ , а медиана $G(x)$ равна 1/2 .

Согласно соотношению (9) для выполнения гипотезы (6) достаточно определить δ как функцию λ , $\delta = \delta(\lambda)$, из условия

$$\int_0^1 F(x)dx = \frac{1}{2}.$$

Вычисления дают

$$\delta = \delta(\lambda) = 3(1 - \lambda)/2 .$$

Учитывая, что δ лежит между λ и 1, не совпадая ни с тем, ни с другим, получаем ограничения на λ, а именно, $1/3 < \lambda < 3/5$. Итак, построено искомое семейство пар функций распределения.

Пример 4. Пусть, как и в примере 3, распределения сосредоточены на интервале (0; 1), и на нем $F(x)=x$, а $G(x)$ - функция распределения, сосредоточенного в двух точках - β и 1, т.е. G(x) = 0 при x, не превосходящем β; $G(x) = h$ на (β ; 1] ; $G(x) = 1$ при $x > 1$. С такой функцией $G(x)$ легко проводить расчеты. Однако она не удовлетворяет принятым выше условиям непрерывности и строгого возрастания. Вместе с тем легко видеть, что она является предельной (сходимость в каждой точке отрезка [0 ; 1]) для последовательности функций распределения, удовлетворяющих этим условиям, а распределение статистики Вилкоксона для пары функций распределения примера 4 является предельным для последовательности соответствующих распределений статистики Вилкоксона, полученных в рассматриваемых условиях непрерывности и строгого возрастания.

Условие $P(X < Y) = 1/2$ выполнено, если $h = (1 - \beta)^{-1} / 2$ (при β из отрезка [0 ; 1/2]). Поскольку $h > 1/2$ при положительном β, то очевидно, что медиана $G(x)$ равна β, в то время как медиана $F(x)$ равна 1/2 . Значит, при $\beta = 1/2$ медианы совпадают, при всех иных положительных β - различны. При $\beta = 0$ медианой $G(x)$ является любая точка из отрезка [0 ; 1].

Легко подсчитать, что в условиях примера 4 параметры предельного распределения имеют вид

$$b^2 = \beta(1- \beta)^{-1} / 4 , \; g^2 = (1- 2\beta) / 4 .$$

Следовательно, распределение нормированной и центрированной статистики Вилкоксона будет асимптотически нормальным с математическим ожиданием 0 и дисперсией

$$D(T) = 3 [(n-1) \beta(1- \beta)^{-1} + (m-1) (1-2\beta) + 1] (m+n+1)^{-1} .$$

Проанализируем величину $D(T)$ в зависимости от параметра β и объемов выборок m и n. При достаточно больших m и n

$$D(T) = 3 w \beta (1 - \beta)^{-1} + 3 (1 - w) (1 - 2 \beta) ,$$

с точностью до величин порядка $(m+n)^{-1}$, где $w= n/(m+n)$. Значит, $D(T)$ - линейная функция от w, а потому достигает экстремальных значений на границах интервала изменения w, т.е. при $w = 0$ и $w = 1$. Легко видеть, что при $\beta(1-\beta)^{-1} < 1-2\beta$ минимум равен $3\beta(1-\beta)^{-1}$ (при $w = 1$), а максимум равен $3(1 - 2\beta)$ (при $w = 0$). В случае $\beta(1-\beta)^{-1} > 1-2\beta$ максимум равен $3\beta(1-\beta)^{-1}$ (при $w = 1$), а минимум равен $3(1 - 2\beta)$ (при $w = 0$). Если же $\beta(1-\beta)^{-1} = 1-2\beta$ (это равенство справедливо при $\beta=\beta_0 = 1 - 2^{-1/2} = 0{,}293$), то $D(T)=3 (2^{1/2}-1)=1{,}2426...$ при всех w из отрезка [0; 1].

Первый из описанных выше случаев имеет быть при $\beta < \beta_0$, при этом минимум $D(T)$ возрастает от 0 (при $\beta=0$, $w=1$ - предельный случай) до $3(2^{1/2} - 1)$ (при $\beta=\beta_0$, w - любом), а максимум уменьшается от 3 (при $\beta=0$, $w=0$ - предельный случай) до $3 (2^{1/2} - 1)$ (при $\beta=\beta_0$, w - любом). Второй случай относится к β из интервала (β_0 ; 1/2]. При этом минимум убывает от приведенного выше значения для $\beta = \beta_0$ до 0 (при β=1/2 , *w*=0 - предельный случай) , а максимум возрастает от того же значения при $\beta = \beta_0$ до 3 (при β=1/2 , *w*=0).

Таким образом, *D(T)* может принимать все значения из интервала (0 ; 3) в зависимости от значений β и *w*. Если $D(T) < 1$, то при применении критерия Вилкоксона к выборкам с рассматриваемыми функциями распределения гипотеза однородности (2) будет приниматься чаще (при соответствующих значениях β и *w* - с вероятностью, сколь угодно близкой к 1), чем если бы она самом деле была верна. Если $1<D(T)<3$, то гипотеза (2) также принимается достаточно часто. Так, если уровень значимости критерия Вилкоксона равен 0,05, то (асимптотическая) критическая область этого критерия, как показано выше, имеет вид $\{T: |T| \geq 1,96\}$. Если - самый плохой случай - $D(T)=3$, то гипотеза (2) принимается с вероятностью 0,7422.

Гипотеза сдвига. При проверке гипотезы однородности мы рассмотрели различные виды нулевых и альтернативных гипотез - гипотезу (2) и ее отрицание в качестве альтернативы, гипотезу (6) и ее отрицание, гипотезы о равенстве или различии медиан. В теоретических работах по математической статистике часто рассматривают гипотезу сдвига, в которой альтернативой гипотезе (2) является гипотеза

$$H_1: \quad F(x) = G(x + r) \quad (12)$$

при всех *x* и некотором сдвиге *r*, отличным от 0. Если верна альтернативная гипотеза H_1, то вероятность $P(X < Y)$ отлична от 1/2, а потому при альтернативе (12) критерий Вилкоксона является состоятельным.

В некоторых прикладных постановках гипотеза (12) представляется естественной. Например, если одним и тем же прибором проводятся две серии измерений двух значений некоторой величины (физической, химической и т.п.). При этом функция распределения *G(x)* описывает погрешности измерения одного значения, а *G(x+r)* - другого. Вопреки распространенному заблуждению, хорошо известно, что распределение погрешностей измерений, как правило, не является нормальным - см. об этом начало главы. Однако при анализе конкретных экономических данных как правило, нет никаких оснований считать, что отсутствие однородности всегда выражается столь однозначным образом, как следует из формулы (12). Поэтому эконометрику для проверки однородности необходимо использовать статистические критерии, состоятельные против любого отклонения от гипотезы однородности (2).

Почему же математики так любят гипотезу сдвига (12)? Да потому, что она дает возможность доказывать глубокие математические результаты, например, об асимптотической оптимальности критериев. К сожалению, с точки зрения эконометрики это напоминает поиск ключей под фонарем, где светло, а не там, где они потеряны.

Отметим еще одно обстоятельство. Часто говорят (в соответствии с классическим подходом математической статистики), что нельзя проверять нулевые гипотезы без рассмотрения альтернативных. Однако при эконометрическом анализе данных зачастую полностью ясна формулировка той гипотезы, которую желательно проверить (например, гипотезы полной однородности - см. формулу (2)), в то время как формулировка альтернативной гипотезы не очевидна (то ли это гипотеза о неверности равенства (2) хотя бы для одного значения *x*, то ли это альтернатива (8), то ли - альтернатива сдвига (12), и т.д.). В таких случаях целесообразно "обернуть" задачу - исходя из статистического критерия найти альтернативы, относительно которых он состоятелен. Именно это и проделано в настоящей пункте для критерия Вилкоксона.

Подведем итоги рассмотрения критерия Вилкоксона.

1. Критерий Вилкоксона (Манна-Уитни) является одним из самых распространенных непараметрических ранговых критериев, используемых для проверки однородности двух

выборок. Его значение не меняется при любом монотонном преобразовании шкалы измерения (т.е. он пригоден для эконометрического анализа данных, измеренных в порядковой шкале).

2. Распределение статистики критерия Вилкоксона определяется функциями распределения $F(x)$ и $G(x)$ и объемами m и n двух выборок. При больших объемах выборок распределение статистики Вилкоксона является асимптотически нормальным с параметрами, выписанными выше (см. формулы (1), (3) и (5)).

3. При альтернативной гипотезе, когда функции распределения выборок $F(x)$ и $G(x)$ не совпадают, распределение статистики Вилкоксона зависит от величины $a = P(X < Y)$. Если a отличается от 1/2, то мощность критерия Вилкоксона стремится к 1, и отличает нулевую гипотезу $F = G$ от альтернативной. Если же $a = 1/2$, то это не всегда имеет место. В примере 2 приведены две *различные* функции распределения выборок $F(x)$ и $G(x)$ такие, что гипотеза однородности $F = G$ при проверке с помощью критерия Вилкоксона будет приниматься *чаще,* чем если она на самом деле верна.

4. Следовательно, в случае общей альтернативы критерий Вилкоксона не является состоятельным, т.е. не всегда позволяет обнаружить различие функций распределения. Однако это не лишает его практической ценности, точно так же, как несостоятельность критериев типа хи-квадрат при проверке согласия, независимости или однородности не мешает отклонять нулевую гипотезу во многих практически важных случаях. Однако принятие нулевой гипотезы с помощью критерия Вилкоксона может означать не совпадение F и G, а лишь выполнение равенства $a = 1/2$.

5. Иногда утверждают, что с помощью критерия Вилкоксона можно проверять равенство медиан функций распределения F и G. Это не так. В примерах 3 и 4 указаны F и G с $a = 1/2$, но с различными медианами. Во многих случаях это различие нельзя обнаружить с помощью критерия Вилкоксона, как это показано при численном анализе асимптотической дисперсии в примере 4.

6. Указанные выше недостатки критерия Вилкоксона исчезают для специального вида альтернативы - т.н. "альтернативы сдвига" H_1: $F(x) = G(x + r)$. В этом частном случае при справедливости альтернативной гипотезы мощность стремится к 1, различие медиан также всегда обнаруживается. Однако альтернатива сдвига не всегда естественна. Ее целесообразно принять, если одним и тем же прибором проводятся две серии измерений двух значений некоторой величины (физической, химической и т.п.). При этом функция распределения $G(x)$ описывает результаты измерений с погрешностями одного значения, а $F(x) = G(x+r)$ - другого. Другими словами, меняется лишь измеряемое значение, а собственно распределение погрешностей - одно и то же, присущее используемому средству измерения (и обычно описанное в его техническом паспорте). Однако в большинстве эконометрических исследований нет никаких оснований считать, что при альтернативе функция распределения второй выборки лишь сдвигается, но не меняется каким-либо иным образом.

7. При всех своих недостатках критерий Вилкоксона прост в применении и часто позволяет обнаруживать различие групп (поскольку оно часто сводится к отличию $a = P(X < Y)$ от 1/2). Приведенные здесь критические замечания не следует понимать как призыв к полному отказу от использования критерия Вилкоксона. Однако для проверки гипотезы однородности в случае альтернативы общего вида можно порекомендовать состоятельные критерии, в частности, рассматриваемые в следующем пункте критерии Смирнова и типа омега-квадрат (Лемана-Розенблатта).

8. В литературе по прикладным статистическим методам соседствуют два стиля изложения. Один из них исходит из формулировок нулевой и альтернативных гипотез (или описания набора гипотез, из которого надо выбрать наиболее адекватную), для проверки которых строятся те или иные критерии. При другом стиле изложения упор делается на

алгоритмическое описание критериев для проверки тех или иных гипотез, а об альтернативах даже не упоминается.

Например, в литературе по математической статистике часто говорится, что для проверки нормальности используются критерии асимметрии и эксцесса (они описаны, например, в лучшем справочнике 1960-1980-х годов [8, табл. 4.7]). Однако эти критерии позволяют проверять некоторые соотношения между моментами распределения, но отнюдь не являются состоятельными критериями нормальности (не все отклонения от нормальности обнаруживают). Впрочем, для эконометрики эти критерии практического значения не имеют, поскольку заранее известно, что распределения конкретных экономических данных отличны от нормальных.

Так что недостатки критерия Вилкоксона не является исключением, мощность ряда иных популярных в математической статистике критериев заслуживает тщательного изучения, при этом заранее можно сказать, что зачастую они не позволяют проверять те гипотезы, с которыми традиционно связаны. При применении подобных критериев к анализу реальных данных необходимо тщательно взвешивать их достоинства и недостатки.

4.6. Состоятельные критерии проверки однородности для независимых выборок

В соответствии с эконометрической теорией естественно потребовать, чтобы рекомендуемый для массового использования в экономических и технико-экономических исследованиях критерий однородности был состоятельным. Напомним: это значит, что для любых отличных друг от друга функций распределения $F(x)$ и $G(x)$ (другими словами, при справедливости альтернативной гипотезы H_1) вероятность отклонения гипотезы H_0 должна стремиться к 1 при увеличении объемов выборок m и n. Из перечисленных выше в конце п.4 критериев состоятельными являются только критерии Смирнова и типа омега-квадрат.

Проведенное исследование мощности (методом статистических испытаний) первых четырех из перечисленных выше критериев (при различных вариантах функций распределения $F(x)$ и $G(x)$) подтвердило преимущество критериев Смирнова и омега-квадрат и при объемах выборок 6-12.

Критерий Смирнова однородности двух выборок. Он предложен членом-корреспондентом АН СССР Н.В. Смирновым в 1939 г. (см. справочник [8]). Единственное ограничение - функции распределения $F(x)$ и $G(x)$ должны быть непрерывными. Напомним, что согласно Л.Н. Большеву и Н.В. Смирнову [8] значение эмпирической функции распределения в точке x равно доле результатов наблюдений в выборке, меньших x. Критерий Смирнова основан на использовании эмпирических функций распределения $F_m(x)$ и $G_n(x)$, построенных по первой и второй выборкам соответственно. Значение статистики Смирнова

$$D_{m,n} = \sup_x | F_m(x) - G_n(x) |$$

сравнивают с соответствующим критическим значением (см., например, [8]) и по результатам сравнения принимают или отклоняют гипотезу H_0 о совпадении (однородности) функций распределения. Практически значение статистики $D_{m,n}$ рекомендуется согласно монографии [8] вычислять по формулам

$$D^{+}_{m,n} = \max_{1\le r\le n}\left[\frac{r}{n} - F_m(y'_r)\right] = \max_{1\le r\le m}\left[G_n(x'_s) - \frac{s-1}{m}\right],$$

$$D^{-}_{m,n} = \max_{1\le r\le n}\left[F_m(y'_r) - \frac{r-1}{n}\right] = \max_{1\le r\le m}\left[\frac{s}{m} - G_n(x'_s)\right],$$

$$D_{m,n} = \max(D^+_{m,n}, D^-_{m,n}),$$

где $x'_1<x'_2<\ldots<x'_m$ - элементы первой выборки $x_1,x_2,\ldots,x_m$, переставленные в порядке возрастания, а $y'_1<y'_2<\ldots<y'_n$ - элементы второй выборки $y_1,y_2,\ldots,y_n$, также переставленные в порядке возрастания.

Разработаны алгоритмы и программы для ЭВМ, позволяющие рассчитывать точные распределения, процентные точки и достигаемый уровень значимости для двухвыборочной статистики Смирнова $D_{m,n}$, разработаны подробные таблицы (см., например, методику [15], содержащую тексты программ и подробные таблицы).

Однако у критерия Смирнова есть и недостатки. Его распределение сосредоточено в сравнительно небольшом числе точек, поэтому функция распределения растет большими скачками. В результате не удается выдержать заданный уровень значимости, реальный уровень значимости может в несколько раз отличаться от номинального (подробному обсуждению неклассического феномена существенного отличия реального уровня значимости от номинального посвящена работа [16]).

Критерий типа омега-квадрат (Лемана-Розенблатта). Статистика критерия типа омега-квадрат для проверки однородности двух независимых выборок имеет вид:

$$A = \frac{mn}{m+n}\int_{-\infty}^{\infty}(F_m(x) - G_n(x))^2\, dH_{m+n}(x),$$

где $H_{m+n}(x)$ – эмпирическая функция распределения, построенная по объединенной выборке,

$$H_{m+n}(x) = \frac{m}{m+n} F_m(x) + \frac{n}{m+n} G_n(x).$$

Статистика A типа омега-квадрат была предложена Э. Леманом в 1951 г., изучена М. Розенблаттом в 1952 г., а затем и другими исследователями. Она зависит лишь от рангов элементов двух выборок в объединенной выборке. Пусть $x_1, x_2, \ldots, x_m$ - первая выборка, $x'_1 < x'_2 < \ldots x'_m$ - соответствующий вариационный ряд, $y_1, y_2, \ldots, y_n$ -вторая выборка, $y'_1 < y'_2 < \ldots < y'_n$ - вариационный ряд, соответствующий второй выборке. Поскольку функции распределения независимых выборок непрерывны, то с вероятностью 1 все выборочные значения различны, совпадения отсутствуют. Статистика А представляется в виде (см., например, [8]):

$$A = \frac{1}{mn(m+n)}[m\sum_{i=1}^{m}(r_i - i)^2 + n\sum_{j=1}^{n}(s_j - j)^2] - \frac{4mn-1}{6(m+n)},$$

где r_i - ранг x'_i и s_j - ранг y'_j в общем вариационном ряду, построенном по объединенной выборке.

Правила принятия решений при проверке однородности двух выборок на основе статистик Смирнова и типа омега-квадрат, т.е. таблицы критических значений в зависимости от уровней значимости и объемов значимости приведены, например, в таблицах [8].

Рекомендации по выбору критерия однородности. Для критерия типа омега-квадрат нет выраженного эффекта различия между номинальными и реальными уровнями значимости. Поэтому мы *рекомендуем для проверки однородности функций распределения (гипотеза H_0) применять статистику A типа омега-квадрат. Если методическое, табличное или программное обеспечение для статистики Лемана-Розенблатта отсутствует, рекомендуем использовать критерий Смирнова. Для проверки однородности математических ожиданий (гипотеза H'_0) целесообразно применять критерий Крамера-Уэлча.* По нашему мнению, статистики Стьюдента, Вилкоксона и др. допустимо использовать лишь в отдельных частных случаях, рассмотренных выше.

Некоторые соображения о внедрении современных методов прикладной статистики в практику технических и технико-экономических исследований. Даже из проведенного выше разбора лишь одной из типичных статистических задач - задачи проверки однородности двух выборок - можно сделать вывод о целесообразности широкого развертывания в организациях различных форм собственности работ по критическому анализу сложившейся в технических и технико-экономических исследованиях практики статистической обработки данных и по внедрению накопленного арсенала современных методов прикладной статистики. По нашему мнению, широкого внедрения заслуживают, в частности, методы многомерного статистического анализа, планирования эксперимента, статистики объектов нечисловой природы. Очевидно, рассматриваемые работы должны быть плановыми, организационно оформленными, проводиться мощными самостоятельными организациями и подразделениями. Целесообразно создание службы статистических консультаций в системе научно-исследовательских учреждений и вузов технического и технико-экономического профиля.

4.7. Методы проверки однородности для связанных выборок

Начнем с практического примера. Приведем письмо главного инженера подмосковного химического комбината (некоторые названия изменены).

"Директору Института высоких статистических технологий
и эконометрики (Фамилия, имя, отчество)

Наш комбинат выпускает мастику по ГОСТ (следует номер) и является разработчиком указанного стандарта.

В результате исследовательских работ по подбору стандартного метода определения вязкости мастики на комбинате накоплен большой опыт сравнительных данных определения вязкости по двум методам:

- неразбавленной мастики - на нестандартном приборе фабрики им. Петрова;
- раствора мастики - на стандартном вискозиметре ВЗ-4.

Учитывая высокую компетентность сотрудников Вашего института, прошу Вас, в порядке оказания технической помощи нашему предприятию, поручить соответствующей лаборатории провести обработку представленной статистики современными эконометрическими методами и выдать заключение о наличии (или отсутствии) зависимости между указанными выше методами определения вязкости мастики. Ваш5е заключение необходимо для решения спорного вопроса о целесообразности вновь ввести в ГОСТ (следует номер) метода определения вязкости мастики по вискозиметру ВЗ-4, который, по мнению некоторых потребителей, был необоснованно исключен из этого ГОСТ по изменению № 1.

Заранее благодарю Вас за оказанную помощь.

Приложение: статистика на 3 листах.

Главный инженер (Подпись) (Фамилия, имя, отчество)"

Комментарий. Вязкость мастики - один из показателей качества мастики. Измерять этот показатель можно по-разному. И, как оказалось, разные способы измерения дают разные результаты. Ничего необычного в этом нет. Однако поставщику и потребителю следует согласовать способы измерения показателей качества. Иначе достаточно часто поставщик (производитель) будет утверждать, что он выполнил условия контракта, а потребитель заявлять, что нет. Такая конфликтная ситуация иногда называется арбитражной, поскольку для ее решения стороны могут обращаться в арбитражный суд. Простейший метод согласования способов измерения показателей состоит в том, чтобы выбрать один из них и внести в государственный стандарт, который тем самым будет содержать не только описание продукции, перечень ее показателей качества и требований к ним, но и способы измерения этих показателей.

Заключение по статистическим данным, представленным химическим комбинатом. Для каждой из 213 партий мастики представлены два числа - результат измерения вязкости на нестандартном приборе фабрики им. Петрова и результат измерения вязкости на стандартном вискозиметре ВЗ-4. Требуется установить, дают ли два указанных метода сходные результаты. Если они дают сходные результаты, то нет необходимости вводить в соответствующий ГОСТ указание о методе определения вязкости. Если же методы дают существенно различные результаты, то подобное указание ввести необходимо.

Для применения эконометрических методов в рассматриваемой задаче необходимо описать вероятностную модель. Считаем, что статистические данные имеют вид $(x_i, y_i), i = 1,2,...,213$, где x_i -результат измерения на нестандартном приборе фабрики им. Петрова в i-ой партии, а y_i - результат измерения вязкости на стандартном вискозиметре ВЗ-4 в той же i-ой партии. Пусть a_i - истинное значение показателя качества в i-ой партии. Естественно считать, что указанные выше случайные вектора независимы в совокупности. При этом они не являются одинаково распределенными, поскольку отличаются истинными значениями показателей качества a_i. Принимаем, что *при каждом i случайные величины x_i - a_i и y_i - a_i независимы и одинаково распределены*. Это условие и означает *однородность в связанных выборках.* Параметры связи - величины a_i . Их наличие не позволяет объединить первые координаты в одну выборку, вторую - во вторую, как делалось в случае проверки однородности двух независимых выборок.

В предположении непрерывности функций распределения из условия однородности в связанных выборках вытекает, что

$$P(x_i < y_i) = P(x_i \geq y_i) = \frac{1}{2}.$$

Рассмотрим случайные величины $Z_i = x_i - y_i, \quad i = 1,2,...,213$. Из последнего соотношения вытекает, что при справедливости гипотезы однородности для связанных выборок эти случайные величины имеют нулевые медианы. Другими словами, проверка того, что метода измерения вязкости дают схожие результаты, эквивалентна проверке равенства 0 медиан величин Z_i.

Для проверки гипотезы о том, что медианы величин Z_i нулевые, применим широко известный критерий знаков (см., например, справочник [8, с.89-91]). Согласно этому критерию необходимо подсчитать, в скольких партиях $x_i < y_i$ и в скольких $x_i \geq y_i$. Для представленных химическим комбинатом данных $x_i < y_i$ в 187 случаях из 213 и $x_i \geq y_i$ в 26 случаях из 213.

Если рассматриваемая гипотеза верна, то число W осуществлений события $\{x_i < y_i\}$ имеет биномиальное распределение с параметрами p = 1/2 и n = 213. Математическое ожидание $M(W)$=106,5, а среднее квадратическое отклонение $\sigma = \sqrt{np(1-p)} = 7{,}3$. Следовательно, интервал $M(W) \pm 3\sigma$ - это интервал 84≤W≤129. Найденное по данным химического комбината значение W=187 лежит далеко вне этого интервала. Поэтому рассматриваемую гипотезу необходимо отвергнуть (на любом используемом в прикладных работах уровне значимости, в частности, на уровне значимости 1%).

Таким образом, статистический анализ показывает, что два метода дают существенно различные результаты - по прибору фабрики им. Петрова результаты измерений, как правило, меньше, чем по вискозиметру ВЗ-4. Это означает, что в соответствующий ГОСТ целесообразно ввести указание на метод определения вязкости.

Система вероятностных моделей при проверке гипотезы однородности для связанных выборок. Как и в случае проверки однородности для независимых выборок, система вероятностных моделей состоит из трех уровней. Наиболее простая модель - на уровне однородности альтернативного признака - уже рассмотрена. Она сводится к проверке гипотезы для биномиального распределения:

$$H_0 : p = \frac{1}{2}.$$

Речь идет о "критерии знаков". При справедливости гипотезы однородности число W осуществлений события $\{x_i < y_i\}$ имеет биномиальное распределение с вероятностью успеха p = 1/2 и числом испытаний n. Альтернативная гипотеза состоит в том, что вероятность успеха отличается от 1/2:

$$H_1 : p \neq \frac{1}{2}.$$

Гипотезу p = 1/2 можно проверять как непосредственно с помощью биномиального распределения (используя таблицы или программное обеспечение), так и опираясь на теорему Муавра-Лапласа. Согласно этой теореме

$$\lim_{n\to\infty} P\{\frac{2W-n}{\sqrt{n}} \leq x\} = \Phi(x)$$

при всех x, где $\Phi(x)$ - функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. Из теоремы Муавра-Лапласа вытекает правило принятия решений на уровне значимости 5%: если

$$\left|\frac{2W-n}{\sqrt{n}}\right| \leq 1{,}96,$$

то гипотезу однородности связанных выборок принимают, в противном случае отклоняют. Как обычно, при желании использовать другой уровень значимости применяют в качестве критического значения иной квантиль нормального распределения. Использование предельных теорем допустимо при достаточно больших объемах выборки. По поводу придания точного смысла термину "достаточно большой" продолжаются дискуссии. Обычно считается, что несколько десятков (два-три десятка) - это уже "достаточно много". Более правильно сказать, что ответ зависит от задачи, от ее сложности и практической значимости.

Второй уровень моделей проверки однородности связанных выборок - это уровень проверки однородности характеристик, прежде всего однородности математических ожиданий. Исходные данные - количественные результаты измерений (наблюдений, испытаний, анализов, опытов) двух признаков x_j и y_j , j = 1,2,…,n, а непосредственно анализируются их разности $Z_j = x_j - y_j$, j = 1,2,…,n. Предполагается, что эти разности независимы в совокупности и одинаково распределены, однако функция распределения неизвестна эконометрику. Необходимо проверить непараметрическую гипотезу

$$H_{01} : M(Z_j) = 0.$$

Альтернативная гипотеза также является непараметрической и имеет вид:

$$H_{11} : M(Z_j) \neq 0.$$

Как и в случае проверки гипотезы согласованности для независимых выборок с помощью критерия Крамера-Уэлча, в рассматриваемой ситуации естественно использовать статистику

$$Q = \frac{\overline{Z}}{s(Z)},$$

где

$$\overline{Z} = \frac{Z_1 + Z_2 + ... + Z_n}{n} -$$

среднее арифметическое разностей, а

$$s(Z) = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(Z_j - \overline{Z})^2} -$$

выборочное среднее квадратическое отклонение. Из центральной Предельной Теоремы теории вероятностей и теорем о наследовании сходимости, полученных в монографии [11], вытекает, что

$$\lim_{n\to\infty} P\{Q \le x\} = \Phi(x)$$

при всех x, где $\Phi(x)$ - функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. Отсюда вытекает правило принятия решений на уровне значимости 5%: если

$$|Q| \le 1{,}96,$$

то гипотезу однородности математических ожиданий связанных выборок принимают, в противном случае отклоняют. Как обычно, при желании использовать другой уровень значимости применяют в качестве критического значения иной квантиль нормального распределения. Повторим, что использование предельных теорем допустимо при достаточно больших объемах выборки.

Третий уровень моделей проверки однородности связанных выборок - это уровень проверки однородности (совпадения) функций распределения. Необходимо проверить непараметрическую гипотезу наиболее всеохватного вида:

$$H_{03} : F(x) = G(x), x \in R^1,$$

где

$$F(x) = P(x_i \le x), G(x) = P(y_i \le x).$$

При этом предполагается, что все участвующие в вероятностной модели случайные величины независимы (в совокупности) между собой.

Отметим одно важное свойство функции распределения случайной величины Z. Если случайные величины x и y независимы и одинаково распределены, то для $H(x)=P(Z\le x)$ выполнено, как нетрудно видеть, соотношение

$$H(-x)=1-H(x).$$

Это соотношение означает симметрию функции распределения относительно 0. Плотность такой функции распределения является четной функцией, ее значения в точках x и $(-x)$ совпадают.

Какого типа отклонения от гипотезы симметрии можно ожидать при альтернативных гипотезах?

Как и в случае проверки однородности независимых выборок, в зависимости от вида альтернативной гипотезы выделяют два подуровня моделей. Рассмотрим сначала альтернативу сдвига

$$H_{13} : G(x) = F(x+a).$$

В этом случае распределение Z при альтернативе отличается сдвигом от симметричного относительно 0. Для проверки гипотезы однородности может быть использован критерий знаковых рангов, разработанный Вилкоксоном (см., например, справочник [9, с.46-53]).

Он строится следующим образом. Пусть $R(Z_j)$ является рангом $|Z_j|$ в ранжировке от меньшего к большему абсолютных значений разностей $|Z_1|, |Z_2|,...,|Z_n|$, $j=1,2,...,n$. Положим для $j=1,2,...,n$

$$Q(Z_j) = \begin{cases} 1, Z_j > 0, \\ 0, Z_j < 0. \end{cases}$$

Статистика критерия знаковых рангов имеет вид

$$W^+ = \sum_{j=1}^{n} R(Z_j) Q(Z_j).$$

Таким образом, нужно просуммировать ранги положительных разностей в вариационном ряду, построенном стандартным образом по абсолютным величинам всех разностей.

Для практического использования статистики критерия знаковых рангов Вилкоксона либо обращаются к соответствующим таблицам и программному обеспечению, либо применяют асимптотические соотношения. При выполнении нулевой гипотезы статистика

$$W^{++} = \frac{W^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

имеет асимптотическое (при $n \to \infty$) стандартное нормальное распределение с математическим ожиданием 0 и дисперсией 1. Следовательно, правило принятия решений на уровне значимости 5%: имеет обычный вид если

$$\left| W^{++} \right| \le 1{,}96,$$

то гипотезу однородности связанных выборок по критерию знаковых рангов Вилкоксона принимают, в противном случае отклоняют. Как обычно, при желании использовать другой уровень значимости применяют в качестве критического значения иной квантиль нормального распределения. Повторим еще раз, что использование предельных теорем допустимо при достаточно больших объемах выборки.

Альтернативная гипотеза общего вида записывается как

$$H_{14} : H(-x_0) \ne 1 - H(x_0)$$

при некотором x_0. Таким образом, проверке подлежит гипотеза симметрии относительно 0, которую можно переписать в виде

$$H(x) + H(-x) - 1 = 0\,.$$

Для построенной по выборке $Z_j = x_j - y_j$, $j = 1,2,\ldots,n$, эмпирической функции распределения $H_n(x)$ последнее соотношение выполнено лишь приближенно:

$$H_n(x) + H_n(-x) - 1 \approx 0.$$

Как измерять отличие от 0? По тем же соображениям, что и в предыдущем пункте, целесообразно использовать статистику типа омега-квадрат. Соответствующий критерий был предложен в работе [17]. Он имеет вид

$$\omega_n^2 = \sum_{j=1}^{n} (H_n(Z_j) + H_n(-Z_j) - 1)^2.$$

В работе [17] найдено предельное распределение этой статистики:

$$\lim_{n \to \infty} P(\omega_n^2 < x) = S_0(x).$$

В табл.1 приведены критические значения статистики типа омега-квадрат для проверки симметрии распределения (и тем самым для проверки однородности связанных выборок), соответствующие наиболее распространенным значениям уровней значимости (расчеты проведены Г.В. Мартыновым).

Табл.1. Критические значения статистики ω_n^2

для проверки симметрии распределения

Значение функции распределения $S_0(x)$	Уровень значимости $\alpha = 1 - S_0(x)$	Критическое значение x статистики ω_n^2
0,90	0,10	1,20
0,95	0,05	1,66
0,99	0,01	2,80

Как следует из табл.1, правило принятия решений при проверке однородности связанных выборок в наиболее общей постановке и при уровне значимости 5% формулируется так. Вычислить статистику ω_n^2. Если $\omega_n^2 \leq 1{,}66$, то принять гипотезу однородности. В противном случае - отвергнуть.

Пример. Пусть величины Z_j, j=1,2,…,20, таковы:

20, 18, (-2), 34, 25, (-17), 24, 42, 16, 26, 13, (-23), 35, 21, 19, 8, 27, 11, (-5), 7.

Соответствующий вариационный ряд $Z(1) < Z(2) < ... < Z(20)$ имеет вид:

(-23)<(-17)<(-5)<(-2)<7<8<11<13<16<18<19<20<21<24<25<26<27<34<35<42.

Для расчета значения статистики ω_n^2 построим табл.2 из 7 столбцов и 20 строк, не считая заголовков столбцов (сказуемого таблицы). В первом столбце указаны номера (ранги) членов вариационного ряда, во втором - сами эти члены, в третьем - значения эмпирической функции распределения при значениях аргумента, совпадающих с членами вариационного ряда. В следующем столбце приведены члены вариационного ряда с обратным знаком, а затем указываются соответствующие значения эмпирической функции распределения. Например, поскольку минимальное наблюдаемое значение равно (-23), то $H_n(x)=0$ при x<-23, а потому для членов вариационного ряда с 14-го по 20-й в пятом столбце стоит 0. В качестве другого примера рассмотрим минимальный член вариационного ряда, т.е. (-23). Меняя знак, получаем 23. Это число стоит между 13-м и 14-м членами вариационного ряда, 21<23<24. На этом интервале эмпирическая функция распределения совпадает со своим значением в левом конце, поэтому следует записать в пятом столбце значение 0,65. Остальные ячейки пятого столбца заполняются аналогично. На основе третьего и пятого столбцов элементарно заполняется шестой столбец, а затем и седьмой. Остается найти сумму значенийб стоящих в седьмом столбце. Подобная таблица удобна как для ручного счета, так и при использовании электронных таблиц типа Excel.

Табл.2. Расчет значения статистики ω_n^2

для проверки симметрии распределения

j	$Z(j)$	$H_n(Z(j))$	$-Z(j)$	$H_n(-Z(j))$	$H_n(Z(j))+ H_n(-Z(j))-1$	$(H_n(Z(j))+ H_n(-Z(j))-1)^2$
1	-23	0,05	23	0,65	-0,30	0,09
2	-17	0,10	17	0,45	-0,45	0,2025
3	-5	0,15	5	0,20	-0,65	0,4225
4	-2	0,20	2	0,20	-0,60	0,36
5	7	0,25	-7	0,10	-0,65	0,4225
6	8	0,30	-8	0,10	-0,60	0,36
7	11	0.35	-11	0,10	-0,55	0,3025
8	13	0,40	-13	0,10	-0,50	0,25
9	16	0,45	-16	0,10	-0,45	0,2025
10	18	0,50	-18	0,05	-0,45	0,2025

11	19	0,55	-19	0,05	-0,40	0,16
12	20	0,60	-20	0,05	-0,35	0,1225
13	21	0,65	-21	0,05	-0,30	0,09
14	24	0,70	-24	0	-0,30	0,09
15	25	0,75	-25	0	-0,25	0,0625
16	26	0,80	-26	0	-0,20	0,04
17	27	0,85	-27	0	-0,15	0,0225
18	34	0,90	-34	0	-0,10	0,01
19	35	0,95	-35	0	-0,05	0,0025
20	42	1,00	-42	0	0	0

Результаты расчетов (суммирование значений по седьмому столбцу табл.2) показывают, что значение статистики ω_n^2 =3,055. В соответствии с табл.1 это означает, что на любом используемом в прикладных эконометрических исследованиях уровнях значимости отклоняется гипотеза симметрии распределения относительно 0 (а потому и гипотеза однородности в связанных выборках).

В настоящей главе затронута лишь небольшая часть непараметрических методов анализа числовых эконометрических данных. Обратим вн6имание на непараметрические оценки плотности, которые используются для описания данных, проверки однородности, в задачах восстановления зависимостей и других областях эконометрики. Эконометрические оценки плотности в общем виде рассмотрены в главе 8.

Цитированная литература

1. Новицкий П.В., Зограф И.А. Оценка погрешностей результатов измерений. - Л.: Энергоатомиздат, 1985. - 248 с.
2. Новицкий П.В. Основы информационной теории измерительных устройств. -Л.: энергия, 1968. - 248 с.
3. Боровков А.А. Теория вероятностей. - М.: Наука, 1976. - 352 с.
4. Петров В.В. Суммы независимых случайных величин. - М.: Наука, 1972. - 416 с.
5. Золотарев В.М. Современная теория суммирования независимых случайных величин. - М.: Наука, 1986. - 416 с.
6. Егорова Л.А., Харитонов Ю.С., Соколовская Л.В.//Заводская лаборатория. - 1976. Т.42, №10. С. 1237.
7. Артемьев Б.Г., Голубов С.М. Справочное пособие для работников метрологических служб.- М.: Изд-во стандартов, 1982. - 280 с.
8. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. – М.: Наука, 1983. - 416 с.
9. Холлендер М., Вульф Д. Непараметрические методы статистики. – М.: Финансы и статистика, 1983. - 518 с.
10. Боровков А.А. Математическая статистика. – М.: Наука, 1984. - 472 с.
11. Орлов А.И. Устойчивость в социально-экономических моделях. - М.:Наука,1979. – 296 с.
12. Крамер Г. Математические методы статистики / Пер. с англ. / 2-е изд. - М.: Мир, 1975. – 648 с.
13. Гаек Я., Шидак З. Теория ранговых критериев / Пер. с англ. - М.: Наука, 1971. – 376 с.
14. Смолянский М.Л. Таблицы неопределенных интегралов. - М.: ГИФМЛ, 1961. - 108 с.
15. Методика. Проверка однородности двух выборок параметров продукции при оценке ее технического уровня и качества. – М.: ВНИИ стандартизации, 1987. – 116 с.

16. Камень Ю.Э., Камень Я.Э., Орлов А.И. Реальные и номинальные уровни значимости в задачах проверки статистических гипотез / Заводская лаборатория. 1986. Т.52. № 12. С.55-57.
17. Орлов А.И. О проверке симметрии распределения. – Журнал «Теория вероятностей и ее применения». 1972. Т.17. No.2. С.372-377.

Глава 5. Многомерный статистический анализ

В многомерном статистическом анализе выборка состоит из элементов многомерного пространства. Отсюда и название этого раздела эконометрических методов. Из многих задач многомерного статистического анализа рассмотрим две - восстановления зависимости и классификации.

5.1. Оценивание линейной прогностической функции

Начнем с задачи точечного и доверительного оценивания линейной прогностической функции одной переменной.

Исходные данные – набор n пар чисел $(t_k\ ,\ x_k),\ k = 1,2,\dots,n,$ где t_k – независимая переменная (например, время), а x_k – зависимая (например, индекс инфляции, курс доллара США, объем месячного производства или размер дневной выручки торговой точки). Предполагается, что переменные связаны зависимостью

$$x_k = a\ (t_k - t_{cp}) + b + e_k\,,\ k = 1,2,\dots,n,$$

где a и b – параметры, неизвестные статистику и подлежащие оцениванию, а e_k – погрешности, искажающие зависимость. Среднее арифметическое моментов времени

$$t_{cp} = (t_1 + t_2 + \dots + t_n\ )\ /\ n$$

введено в модель для облегчения дальнейших выкладок.

Обычно оценивают параметры a и b линейной зависимости методом наименьших квадратов. Затем восстановленную зависимость используют для точечного и интервального прогнозирования.

Как известно, метод наименьших квадратов был разработан великим немецким математиком К. Гауссом в 1794 г. Согласно этому методу для расчета наилучшей функции, приближающей линейным образом зависимость x от t, следует рассмотреть функцию двух переменных

$$f(a,b) = \sum_{i=1}^{n} (x_i - a(t_i - t_{cp}) - b)^2 .$$

Оценки метода наименьших квадратов - это такие значения $a*$ и $b*$, при которых функция $f(a,b)$ достигает минимума по всем значениям аргументов. Чтобы найти эти оценки, надо вычислить частные производные от функции $f(a,b)$ по аргументам a и $b,$ приравнять их 0, затем из полученных уравнений найти оценки: Имеем:

$$\frac{\partial f(a,b)}{\partial a} = \sum_{i=1}^{n} 2(x_i - a(t_i - t_{cp}) - b)(-(t_i - t_{cp})),$$

$$\frac{\partial f(a,b)}{\partial b} = \sum_{i=1}^{n} 2(x_i - a(t_i - t_{cp}) - b)(-1).$$

Преобразуем правые части полученных соотношений. Вынесем за знак суммы общие множители 2 и (-1). Затем рассмотрим слагаемые. Раскроем скобки в первом выражении, получим, что каждое слагаемое разбивается на три. Во втором выражении также каждое слагаемое есть сумма трех. Значит, каждая из сумм разбивается на три суммы. Имеем:

$$\frac{\partial f(a,b)}{\partial a} = (-2)(\sum_{i=1}^{n} x_i (t_i - t_{cp}) - a\sum_{i=1}^{n} (t_i - t_{cp})^2 - b\sum_{i=1}^{n} (t_i - t_{cp})),$$

$$\frac{\partial f(a,b)}{\partial b} = (-2)(\sum_{i=1}^{n} x_i - a\sum_{i=1}^{n} (t_i - t_{cp})\ \ - bn).$$

Приравняем частные производные 0. Тогда в полученных уравнениях можно сократить множитель (-2). Поскольку

$$\sum_{i=1}^{n}(t_i - t_{cp}) = 0, \qquad (1)$$

уравнения приобретают вид

$$\sum_{i=1}^{n} x_i (t_i - t_{cp}) - a\sum_{i=1}^{n}(t_i - t_{cp})^2 = 0,$$

$$\sum_{i=1}^{n} x_i - bn = 0.$$

Следовательно, оценки метода наименьших квадратов имеют вид

$$a^* = \frac{\sum_{i=1}^{n} x_i (t_i - t_{cp})}{\sum_{i=1}^{n}(t_i - t_{cp})^2}, \quad b^* = x_{cp} = \frac{x_1 + x_2 + \ldots + x_n}{n} \quad . \quad (2)$$

В силу соотношения (1) оценку а* можно записать в более симметричном виде:

$$a^* = \frac{\sum_{i=1}^{n} (x_i - x_{cp})(t_i - t_{cp})}{\sum_{i=1}^{n}(t_i - t_{cp})^2}. \qquad (3)$$

Эту оценку нетрудно преобразовать и к виду

$$a^* = \frac{\sum_{i=1}^{n} x_i t_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} t_i}{\sum_{i=1}^{n} t_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} t_i\right)^2}. \qquad (4)$$

Следовательно, восстановленная функция, с помощью которой можно прогнозировать и интерполировать, имеет вид

$$x^*(t) = a^*(t - t_{cp}) + b^*.$$

Обратим внимание на то, что использование t_{cp} в последней формуле ничуть не ограничивает ее общность. Сравним с моделью вида

$$x_k = c\, t_k + d + e_k,\ k = 1,2,\ldots,n.$$

Ясно, что

$$c = a,\ \ d = b - at_{cp}.$$

Аналогичным образом связаны оценки параметров:

$$c^* = a^*,\ \ d^* = b^* - a^* t_{cp}.$$

Для получения оценок параметров и прогностической формулы нет необходимости обращаться к какой-либо вероятностной модели. Однако для того, чтобы изучать погрешности оценок параметров и восстановленной функции, т.е. строить доверительные интервалы для a^*, b^* и $x^*(t)$, подобная модель необходима.

Непараметрическая вероятностная модель. Пусть значения независимой переменной t детерминированы, а погрешности e_k , $k = 1,2,\ldots,n,$ - независимые одинаково распределенные случайные величины с нулевым математическим ожиданием и дисперсией σ^2, неизвестной статистику.

В дальнейшем неоднократно будем использовать Центральную Предельную Теорему (ЦПТ) теории вероятностей для величин e_k, $k = 1,2,\ldots,n$ (с весами), поэтому для выполнения ее условий необходимо предположить, например, что погрешности e_k, $k = 1,2,\ldots,n$, финитны или имеют конечный третий абсолютный момент. Однако заострять внимание на этих внутриматематических "условиях регулярности" нет необходимости.

Асимптотические распределения оценок параметров. Из формулы (2) следует, что

$$b^* = \frac{a}{n}\sum_{i=1}^{n}(t_i - t_{cp}) + b + \frac{1}{n}\sum_{i=1}^{n} e_i = b + \frac{1}{n}\sum_{i=1}^{n} e_i \quad . \quad (5)$$

Согласно ЦПТ оценка b^* имеет асимптотически нормальное распределение с математическим ожиданием b и дисперсией σ^2/n, оценка которой приводится ниже.

Из формул (2) и (5) вытекает, что

$$x_i - x_{cp} = a(t_i - t_{cp}) + b + e_i - b - \frac{1}{n}\sum_{i=1}^{n} e_i,$$

$$(x_i - x_{cp})(t_i - t_{cp}) = a(t_i - t_{cp})^2 + e_i(t_i - t_{cp}) - \frac{(t_i - t_{cp})}{n}\sum_{i=1}^{n} e_i.$$

Последнее слагаемое во втором соотношении при суммировании по i обращается в 0, поэтому из формул (2-4) следует, что

$$a^* = a + \sum_{i=1}^{n} c_i e_i, \quad c_i = \frac{(t_i - t_{cp})}{\sum_{i=1}^{n}(t_i - t_{cp})^2}. \quad (6)$$

Формула (6) показывает, что оценка a^* является асимптотически нормальной с математическим ожиданием a и дисперсией

$$D(a^*) = \sum_{i=1}^{n} c_i^2 D(e_i) = \frac{\sigma^2}{\sum_{i=1}^{n}(t_i - t_{cp})^2} \quad .$$

Отметим, что многомерная нормальность имеет быть, когда каждое слагаемое в формуле (6) мало сравнительно со всей суммой, т.е.

$$\lim_{n\to\infty} \max |t_i - t_{cp}| / \{\sum_{i=1}^{n}(t_i - t_{cp})^2\}^{1/2} = 0 \quad .$$

Из формул (5) и (6) и исходных предположений о погрешностях вытекает также несмещенность оценок параметров.

Несмещенность и асимптотическая нормальность оценок метода наименьших квадратов позволяют легко указывать для них асимптотические доверительные границы (аналогично границам в предыдущей главе) и проверять статистические гипотезы, например, о равенстве определенным значениям, прежде всего 0. Предоставляем читателю возможность выписать формулы для расчета доверительных границ и сформулировать правила проверки упомянутых гипотез.

Асимптотическое распределение прогностической функции. Из формул (5) и (6) следует, что

$$M(x^*(t)) = M\{a^*(t - t_{cp}) + b^*\} = M(a^*)(t - t_{cp}) + M(b^*) = a(t - t_{cp}) + b = x(t),$$

т.е. рассматриваемая оценка прогностической функции является несмещенной. Поэтому

$$D(x^*(t)) = D(a^*)(t - t_{cp})^2 + 2M\{(a^* - a)(b^* - b)(t - t_{cp})\} + D(b^*).$$

При этом, поскольку погрешности независимы в совокупности и $M(e_i) = 0$, то

$$M\{(a^* - a)(b^* - b)(t - t_{cp})\} = \frac{1}{n}\sum_{i=1}^{n} c_i (t - t_{cp}) M(e_i^2) = \frac{1}{n}(t - t_{cp})\sigma^2 \sum_{i=1}^{n} c_i = 0 .$$

Таким образом,

$$D(x^*(t)) = \sigma^2 \{\frac{1}{n} + \frac{(t - t_{cp})^2}{\sum_{i=1}^{n}(t_i - t_{cp})^2}\} .$$

Итак, оценка $x^*(t)$ является несмещенной и асимптотически нормальной. Для ее практического использования необходимо уметь оценивать остаточную дисперсию $M(e_i^2) = \sigma^2$.

Оценивание остаточной дисперсии. В точках t_k, $k = 1,2,\dots,n$, имеются исходные значения зависимой переменной x_k и восстановленные значения $x^*(t_k)$. Рассмотрим остаточную сумму квадратов

$$SS = \sum_{i=1}^{n} (x^*(t_i) - x(t_i))^2 = \sum_{i=1}^{n} \{(a^* - a)(t_i - t_{cp}) + (b^* - b) - e_i)^2.$$

В соответствии с формулами (5) и (6)

$$SS = \sum_{i=1}^{n} \{(t_i - t_{cp}) \sum_{j=1}^{n} c_j e_j + \tfrac{1}{n}\sum_{j=1}^{n} e_j - e_i\}^2 = \sum_{i=1}^{n} \{\sum_{j=1}^{n} \{c_j (t_i - t_{cp}) + \frac{1}{n}\} e_j - e_i\}^2 = \sum_{i=1}^{n} SS_i.$$

Найдем математическое ожидание каждого из слагаемых:

$$M(SS_i) = \sum_{j=1}^{n} \{c_j (t_i - t_{cp}) + \frac{1}{n}\}^2 \sigma^2 - 2\{c_i (t_i - t_{cp}) + \frac{1}{n})\sigma^2 + \sigma^2.$$

Из сделанных ранее предположений вытекает, что при $n \to \infty$ имеем $M(SS_i) \to \sigma^2, i = 1,2,\dots,n,$ следовательно, по закону больших чисел статистика SS/n является состоятельной оценкой остаточной дисперсии σ^2.

Получением состоятельной оценкой остаточной дисперсии завершается последовательность задач, связанных с рассматриваемым простейшим вариантом метода наименьших квадратов. Не представляет труда выписывание верхней и нижней границ для прогностической функции:

$$x_{верх}(t) = a^*(t - t_{cp}) + b^* + \delta(t), \quad x_{нижн}(t) = a^*(t - t_{cp}) + b^* - \delta(t),$$

где погрешность $\delta(t)$ имеет вид

$$\delta(t) = U(p)\sigma^* \{\frac{1}{n} + \frac{(t - t_{cp})^2}{\sum_{i=1}^{n}(t_i - t_{cp})^2}\}^{1/2}, \quad \sigma^* = \left(\frac{SS}{n}\right)^{1/2}.$$

Здесь p - доверительная вероятность, *U(p),* как и в главе 4 - квантиль нормального распределения порядка *(1+p)/2*, т.е.

$$\Phi(U(p)) = \frac{1+p}{2}.$$

При *p= 0,95* (наиболее применяемое значение) имеем *U(p) = 1,96.* Для других доверительных вероятностей соответствующие значения квантилей можно найти в статистических таблицах (см., например, наилучшее в этой сфере издание [1]).

Сравнение параметрического и непараметрического подходов. Во многих литературных источниках рассматривается параметрическая вероятностная модель метода наименьших квадратов. В ней предполагается, что погрешности имеют нормальное распределение. Это предположение позволяет математически строго получить ряд выводов. Так, распределения статистик вычисляются точно, а не в асимптотике, соответственно вместо квантилей нормального распределения используются квантили распределения Стьюдента, а остаточная сумма квадратов *SS* делится не на *n,* а на *(n-2).* Ясно, что при росте объема данных различия стираются.

Рассмотренный выше непараметрический подход не использует нереалистическое предположение о нормальности погрешностей (см. начало главы 4).. Платой за это является асимптотический характер результатов. В случае простейшей модели метода наименьших квадратов оба подхода дают практически совпадающие рекомендации. Это не всегда так, не всегда два подхода бают близкие результаты. Напомним, что в задаче обнаружения выбросов методы, опирающиеся на нормальное распределение, нельзя считать обоснованными, и обнаружено это было с помощью непараметрического подхода (см. главу 4).

Общие принципы. Кратко сформулируем несколько общих принципов построения, описания и использования эконометрических методов анализа данных. Во-первых, должны быть четко сформулированы исходные предпосылки, т.е. полностью описана используемая вероятностно-статистическая модель. Во-вторых, не следует принимать предпосылки, которые редко выполняются на практике. В-третьих, алгоритмы расчетов должны быть корректны с точки зрения математико-статистической теории. В-четвертых, алгоритмы должны давать полезные для практики выводы.

Применительно к задаче восстановления зависимостей это означает, что целесообразно применять непараметрический подход, что и сделано выше. Однако предположение нормальности, хотя и очень сильно сужает возможности применения, с чисто математической точки зрения позволяет продвинуться дальше. Поэтому для первоначального изучения ситуации, так сказать, "в лабораторных условиях", нормальная модель может оказаться полезной.

Пример оценивания по методу наименьших квадратов. Пусть даны *n*=6 пар чисел $(t_k\ ,\ x_k),\ k\ =\ 1,2,\ldots,6,$ представленных во втором и третьем столбцах табл.1. В соответствии с формулами (2) и (4) выше для вычисления оценок метода наименьших квадратов достаточно найти суммы выражений, представленных в четвертом и пятом столбцах табл.1.

Табл.1. Расчет по методу наименьших квадратов при построении линейной прогностической функции одной переменной

i	t_i	x_i	t_i^2	$t_i x_i$	$a*t_i$	$\hat{x}_i$	$x_i-\hat{x}_i$	$(x_i-\hat{x}_i)^2$
1	1	12	1	12	3,14	12,17	-0,17	0,03
2	3	20	9	60	9,42	18,45	1,55	2,40
3	4	20	16	80	12,56	21,59	-1,59	2,53
4	7	32	49	224	21,98	31,01	0,99	0,98
5	9	35	81	315	28,26	37,29	-2,29	5,24
6	10	42	100	420	31,40	40,43	1,57	2,46
Σ	34	161	256	1111			0,06	13,64

$\frac{\Sigma}{n}$	5,67	26,83	42,67	185,17				

В соответствии с формулой (2) b^* =26,83, а согласно формуле (4)

$$a^* = \frac{1111 - \frac{1}{6}161 \times 34}{256 - \frac{1}{6}(34)^2} = \frac{1111 - 912{,}33}{256 - 192{,}67} = \frac{198{,}67}{63{,}33} = 3{,}14.$$

Следовательно, прогностическая формула имеет вид

$$x^*(t) = 3{,}14(t - 5{,}67) + 26{,}83 = 3{,}14t - 3{,}14 \times 5{,}67 + 26{,}83 =$$
$$= 3{,}14t - 17{,}80 + 26{,}83 = 3{,}14t + 9{,}03.$$

Следующий этап анализа данных - оценка точности приближения функции методом наименьших квадратов. Сначала рассматриваются т.н. восстановленные значения

$$\hat{x}_i = x^*(t_i), \quad i = 1, 2, ..., n.$$

Это те значения, которые полученная в результате расчетов прогностическая функция принимает в тех точках, в которых известны истинные значения зависимой переменной x_i.

Вполне естественно сравнить восстановленные и истинные значения. Это и сделано в шестом - восьмом столбцах табл. 1. Для простоты расчетов в шестом столбце представлены произведения $a^* t_i$, седьмой отличается от шестого добавлением константы 9,03 и содержит восстановленные значения. Восьмой столбец - это разность третьего и седьмого.

Непосредственный анализ восьмого столбца табл.1 показывает, что содержащиеся в нем числа сравнительно невелики по величине по сравнению с третьим столбцом (на порядок меньше по величине). Кроме того, знаки "+" и "-" чередуются. Эти два признака свидетельствуют о правильности расчетов. При использовании метода наименьших квадратов знаки не всегда чередуются. Однако если сначала идут только плюсы, а потом только минусы (или наоборот, сначала только минусы, а потом только плюсы), то это верный показатель того, что в вычислениях допущена ошибка.

Верно следующее утверждение.

Теорема.

$$\sum_{i=1}^{n}(x_i - \hat{x}_i) = 0.$$

Доказательство этой теоремы оставляем читателю в качестве упражнения.

Однако сумма по восьмому столбцу дает 0,06, а не 0. Незначительное отличие от 0 связано с ошибками округления при вычислениях. Близость суммы значений зависимой переменной и суммы восстановленных значений - практический критерий правильности расчетов.

В последнем девятом столбце табл.1 приведены квадраты значений из восьмого столбца. Их сумма - это остаточная сумма квадратов SS = 13,64. В соответствии со сказанным выше оценками дисперсии погрешностей и их среднего квадратического отклонения являются

$$(\sigma^2)^* = \frac{SS}{n} = \frac{13{,}4}{6} = 2{,}27; \quad \sigma^* = \sqrt{\frac{SS}{n}} = \sqrt{\frac{13{,}4}{6}} = 1{,}49.$$

Рассмотрим распределения оценок параметров. Оценка b^* имеет асимптотически нормальное распределение с математическим ожиданием b и дисперсией, которая оценивается как 2,27/6=0,38 (здесь считаем, что 6 - "достаточно большое" число, что, конечно, можно оспаривать). Оценкой среднего квадратического отклонения является 0,615. Следовательно, при доверительной вероятности 0,95 доверительный интервал для параметра b имеет вид (26,83 - 1,96·0,615; 26,83 + 1,96·0,615) = (25,625; 28,035).

В формулах для дисперсий участвует величина

$$\sum_{i=1}^{n}(t_i - t_{cp})^2 = \sum_{i=1}^{n}(t_i^2 - 2t_i t_{cp} + t_{cp}^2) = \sum_{i=1}^{n} t_i^2 - 2t_{cp}\sum_{ш=1}^{n} t_i + nt_{cp}^2 = \sum_{i=1}^{n} t_i^2 - nt_{cp}^2.$$

Подставив численные значения, получаем, что

$$\sum_{i=1}^{n} t_i^2 - nt_{cp}^2 = 256 - 6(5,67)^2 = 63,1.$$

Дисперсия для оценки a^* коэффициента при линейном члене прогностической функции оценивается как 2,27/63,1=0,036, а среднее квадратическое отклонение - как 0,19. Следовательно, при доверительной вероятности 0,95 доверительный интервал для параметра a имеет вид (3,14 - 1,96·0,19; 3,14 + 1,96·0,19) = (2,77; 3,51).

Прогностическая формула с учетом погрешности имеет вид (при доверительной вероятности 0,95)

$$x^*(t) = 3{,}14t + 9{,}03 \pm 1{,}96 \times 1{,}49\sqrt{\frac{1}{6} + \frac{(t - 5{,}67)^2}{63{,}1}}.$$

В этой записи сохранено происхождение различных составляющих. Упростим:

$$x^*(t) = 3{,}14t + 9{,}03 \pm 2{,}92\sqrt{\frac{1}{6} + \frac{(t - 5{,}67)^2}{63{,}1}}.$$

Например, при t = 12 эта формула дает

$$x^*(12) = 46{,}71 \pm 2{,}615.$$

Следовательно, нижняя доверительная граница - это 44,095, а верхняя доверительная граница - это 49,325.

Насколько далеко можно прогнозировать? Обычный ответ таков - до тех пор, пока сохраняется тот стабильный комплекс условий, при котором справедлива рассматриваемая зависимость. Изобретатель метода наименьших квадратов Карл Гаусс исходил из задачи восстановления орбиты астероида (малой планеты) Церера. Движение подобных небесных тел может быть рассчитано на сотни лет. А вот параметры комет (например, срок возвращения) не поддаются столь точному расчету, поскольку за время пребывания в окрестности Солнца сильно меняется масса кометы. В социально-экономической области горизонты надежного прогнозирования еще менее определены. В частности, они сильно зависят от решений центральной власти.

Чтобы выявить роль погрешностей в прогностической формуле, рассмотрим формальный предельный переход $t \to \infty$. Тогда слагаемые 9,03; 1/6; 5,67 становятся бесконечно малыми, и

$$x^*(t) \approx 3{,}14t \pm \frac{2{,}92}{\sqrt{63{,}1}}t = (3{,}14 \pm 0{,}37)t.$$

Таким образом, погрешности составляют около

$$\frac{100 \times 0{,}37}{3{,}14}\% = 11{,}8\%$$

от тренда (математического ожидания) прогностической функции. В социально-экономических исследованиях подобные погрешности считаются вполне приемлемыми.

5.2. Основы линейного регрессионного анализа

В предыдущем пункте метод наименьших квадратов описан в простейшем случае. Он допускает различные обобщения. Например, метод наименьших квадратов дает алгоритм расчетов в случае, если исходные данные – по-прежнему набор n пар чисел (t_k, x_k), $k = 1,2,\ldots,n,$ где t_k – независимая переменная (например, время), а x_k – зависимая (например, индекс инфляции - см. главу 7), а восстанавливать надо не линейную зависимость, а квадратическую:

$$x(t) = at^2 + bt + c.$$

Следует рассмотреть функцию трех переменных

$$f(a,b,c) = \sum_{k=1}^{n} (x_k - at_k^2 - bt_k - c)^2.$$

Оценки метода наименьших квадратов - это такие значения параметров a^*, b^* и c^*, при которых функция $f(a,b,c)$ достигает минимума по всем значениям аргументов. Чтобы найти эти оценки, надо вычислить частные производные от функции $f(a,b,c)$ по аргументам a, b и c, приравнять их 0, затем из полученных уравнений найти оценки: Имеем:

$$\frac{\partial f(a,b,c)}{\partial a} = \sum_{k=1}^{n} \frac{\partial}{\partial a}(x_k - at_k^2 - bt_k - c)^2 = \sum_{k=1}^{n} 2(-t_k^2)(x_k - at_k^2 - bt_k - c)^2.$$

Приравнивая частную производную к 0, получаем линейное уравнение относительно трех неизвестных параметров a,b,c:

$$a\sum_{k=1}^{n} t_k^4 + b\sum_{k=1}^{n} t_k^3 + c\sum_{k=1}^{n} t_k^2 = \sum_{k=1}^{n} t_k^2 x_k.$$

Приравнивая частную производную по параметру b к 0, аналогичным образом получаем уравнение

$$a\sum_{k=1}^{n} t_k^3 + b\sum_{k=1}^{n} t_k^2 + c\sum_{k=1}^{n} t_k = \sum_{k=1}^{n} t_k x_k.$$

Наконец, приравнивая частную производную по параметру c к 0, получаем уравнение

$$a\sum_{k=1}^{n} t_k^2 + b\sum_{k=1}^{n} t_k + cn = \sum_{k=1}^{n} x_k.$$

Решая систему трех уравнений с тремя неизвестными, находим оценки метода наименьших квадратов.

Другие задачи, рассмотренные в предыдущем пункте (доверительные границы для параметров и прогностической функции и др.), также могут быть решены. Соответствующие алгоритмы более громоздки. Для их записи полезен аппарат матричной алгебры (см., например, одну из лучших в этой области монографий [2]). Для реальных расчетов используют соответствующие компьютерные программы.

Раздел многомерного статистического анализа, посвященный восстановлению зависимостей, называется регрессионным анализом. Термин "линейный регрессионный анализ" используют, когда рассматриваемая функция линейно зависит от оцениваемых параметров (от независимых переменных зависимость может быть произвольной). Теория оценивания неизвестных параметров хорошо развита именно в случае линейного регрессионного анализа. Если же линейности нет и нельзя перейти к линейной задаче, то, как правило, хороших свойств от оценок ожидать не приходится.

Продемонстрируем подходы в случае зависимостей различного вида. Если зависимость имеет вид многочлена (полинома)

$$x(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + ... + a_m t^m,$$

то коэффициенты многочлена могут быть найдены путем минимизации функции

$$f(a_0, a_1, a_2, a_3, ..., a_m) = \sum_{k=1}^{n} (x_k - a_0 - a_1 t_k - a_2 t_k^2 - a_3 t_k^3 - ... - a_m t_k^m)^2.$$

Функция от t не обязательно должна быть многочленом. Можно, например, добавить периодическую составляющую, соответствующую сезонным колебаниям. Хорошо известно, например, что инфляция (рост потребительских цен) имеет четко выраженный годовой цикл - в среднем цены быстрее всего растут зимой, в декабре - январе, а медленнее всего (иногда в среднем даже падают) летом, в июле - августе. Пусть для определенности

$$x(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + ... + a_m t^m + A \sin Bt,$$

тогда неизвестные параметры могут быть найдены путем минимизации функции

$$f(a_0, a_1, a_2, a_3, ..., a_m, A, B) = \sum_{k=1}^{n} (x_k - a_0 - a_1 t_k - a_2 t_k^2 - a_3 t_k^3 - ... - a_m t_k^m - A \sin Bt_k)^2.$$

Пусть *I(t)* -индекс инфляции в момент *t*. Принцип стабильности условий приводит к гипотезе о постоянстве темпов роста средних цен, т.е. индекса инфляции. Таким образом, естественная модель для индекса инфляции - это

$$I(t) = Ae^{Bt}.$$

Эта модель не является линейной, метод наименьших квадратов непосредственно применять нельзя. Однако если прологарифмировать обе части предыдущего равенства:

$$\ln I(t) = \ln A + Bt,$$

то получим линейную зависимость, рассмотренную в первом пункте настоящей главы.

Независимых переменных может быть не одна, а несколько. Пусть, например, по исходным данным $(x_k, y_k, z_k), k = 1,2,...,n,$ требуется оценить неизвестные параметры a и b в зависимости

$$z = ax + by + \varepsilon,$$

где ε - погрешность. Это можно сделать, минимизировав функцию

$$f(a,b) = \sum_{k=1}^{n} (z_k - ax_k - by_k)^2.$$

Зависимость от x и y не обязательно должна быть линейной. Предположим, что из каких-то соображений известно, что зависимость должна иметь вид

$$z = ax + by + cx^2 y + dxy + ey^3 + \varepsilon,$$

тогда для оценки пяти параметров необходимо минимизировать функцию

$$f(a,b,c,d,e) = \sum_{k=1}^{n} (z_k - ax_k - by_k - cx_k^2 y_k - dx_k y_k - ey_k^3)^2.$$

Более подробно рассмотрим пример из микроэкономики. В одной из оптимизационных моделей поведения фирмы используется т.н. производственная функция *f(K,L),* задающая объем выпуска в зависимости от затрат капитала *K* и труда *L.* В качестве конкретного вида производственной функции часто используется так называемая функция Кобба-Дугласа

$$f(K,L) = K^{\alpha} L^{\beta}.$$

Однако откуда взять значения параметров α и β? Естественно предположить, что они - одни и те же для предприятий отрасли. Поэтому целесообразно собрать информацию

$(f_k, K_k, L_k), k=1,2,...,n,$ где f_k - объем выпуска на k-ом предприятии, K_k- объем затрат капитала на k-ом предприятии, L_k - объем затрат труда на k-ом предприятии (в кратком изложении здесь не пытаемся дать точных определений используемым понятиям из экономики предприятия). По собранной информации естественно попытаться оценить параметры α и β. Но они входят в зависимость нелинейно, поэтому сразу применить метод наименьших квадратов нельзя. Помогает логарифмирование:

$$\ln f(K,L) = \alpha \ln K + \beta \ln L.$$

Следовательно, целесообразно сделать *замену переменных*

$$x_k = \ln K_k, y_k = \ln L_k, z_k = \ln f_k, k=1,2,3,...,n,$$

а затем находить оценки параметров α и β, минимизируя функцию

$$g(\alpha, \beta) = \sum_{k=1}^{n} (z_k - \alpha x_k - \beta y_k)^2.$$

Найдем частные производные:

$$\frac{\partial g(\alpha,\beta)}{\partial \alpha} = \sum_{k=1}^{n} 2(z_k - \alpha x_k - \beta y_k)(-x_k),$$

$$\frac{\partial g(\alpha,\beta)}{\partial \beta} = \sum_{k=1}^{n} 2(z_k - \alpha x_k - \beta y_k)(-y_k).$$

Приравняем частные производные к 0, сократим на 2, раскроем скобки, перенесем свободные члены вправо. Получим систему двух линейных уравнений с двумя неизвестными:

$$\alpha \sum_{k=1}^{n} x_k^2 + \beta \sum_{k=1}^{n} x_k y_k = \sum_{k=1}^{n} x_k z_k,$$

$$\alpha \sum_{k=1}^{n} x_k y_k + \beta \sum_{k=1}^{n} y_л^2 = \sum_{k=1}^{n} y_k z_k.$$

Таким образом, для вычисления оценок метода наименьших квадратов необходимо найти пять сумм

$$\sum_{k=1}^{n} x_k^2, \quad \sum_{k=1}^{n} x_k y_k, \quad \sum_{k=1}^{n} y_k^2, \quad \sum_{k=1}^{n} x_k z_k, \quad \sum_{k=1}^{n} y_k z_k.$$

Для упорядочения расчета этих сумм может быть использована таблица типа той, что применялась в первом пункте настоящей главы. Отметим, что рассмотренная там постановка переходит в разбираемую сейчас при $y_k = 1, \quad k = 1,2,...,n.$

Подходящая замена переменных во многих случаях позволяет перейти к линейной зависимости. Например, если

$$y = \frac{1}{a+bx},$$

то замена $z=1/y$ приводит к линейной зависимости $z = a + bx$. Если $y=(a+bx)^2$, то замена $z = \sqrt{y}$ приводит к линейной зависимости $z = a + bx$.

Основной показатель качества регрессионной модели. Одни и те же данные можно обрабатывать различными способами. Показателем отклонений данных от модели служит остаточная сумма квадратов *SS*. Чем этот показатель меньше, тем приближение лучше, значит, и модель лучше описывает реальные данные. Однако это рассуждение годится только для моделей с одинаковым числом параметров. Ведь если добавляется новый параметр, по которому можно минимизировать, то и минимум, как правило, оказывается меньше.

В качестве основного показателя качества регрессионной модели используют оценку остаточной дисперсии

$$\widehat{\sigma}^2(m) = \frac{SS}{n-m},$$

скорректированную на число m параметров, оцениваемых по наблюдаемым данным. В случае линейной прогностической модели, рассмотренной в первом пункте настоящей главы, оценка остаточной дисперсии имеет вид

$$\widehat{\sigma}^2 = \frac{SS}{n-2},$$

поскольку число оцениваемых параметров m=2.

Почему эта формула отличается от приведенной в первом пункте? Там в знаменателе n, а здесь - (n-2). Дело в том, что в первом пункте рассмотрена непараметрическая теория при большом объеме данных (при $n \to \infty$), а при безграничном возрастании n разница между n и (n-2) сходит на нет.

А вот при подборе вида модели знаменатель дроби, оценивающей остаточную дисперсию, приходится корректировать на число параметров. Если этого не делать, то придется заключить, что многочлен второй степени лучше соответствует данным, чем линейная функция, многочлен третьей степени лучше приближает исходные данные, чем многочлен второй степени, и т.д. В конце концов доходим до многочлена степени (n-1) с n коэффициентами, который проходит через все заданные точки. Но его прогностические возможности, скорее всего, существенно меньше, чем у линейной функции. *Излишнее усложнение эконометрических моделей вредно.*

Типовое поведение скорректированной оценки остаточной дисперсии

$$v(m) = \widehat{\sigma}^2(m)$$

в зависимости от параметра m в случае расширяющейся системы эконометрических моделей выглядит так. Сначала наблюдаем заметное убывание. Затем оценка остаточной дисперсии колеблется около некоторой константы (теоретического значения дисперсии погрешности).

Поясним ситуацию на примере эконометрической модели в виде многочлена

$$x(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + ... + a_m t^m.$$

Пусть эта модель справедлива при $m = m_0$. При $m < m_0$ в скорректированной оценке остаточной дисперсии учитываются не только погрешности измерений, но и соответствующие (старшие) члены многочлена (предполагаем, что коэффициенты при них отличны от 0). При $m \geq m_0$ имеем

$$\lim_{n\to\infty} v(m) = \sigma^2.$$

Следовательно, скорректированная оценка остаточной дисперсии будет колебаться около указанного предела. Поэтому в качестве оценки неизвестной эконометрику степени многочлена (полинома) можно использовать первый локальный минимум скорректированной оценки остаточной дисперсии, т.е.

$$m^* = \min\{m : v(m-1) > v(m), \quad v(m) \leq v(m+1)\}.$$

В работе [3] найдено предельное распределение этой оценки степени многочлена.

Теорема. При справедливости некоторых условий регулярности

$$\lim_{n\to\infty} P(m^* < m_0) = 0, \quad \lim_{n\to\infty} P(m^* = m_0 + u) = \lambda(1-\lambda)^u, \; u = 0,1,2,....,$$

где

$$\lambda = \Phi(1) - \Phi(-1) = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} \exp\left\{-\frac{x^2}{2}\right\} dx \approx 0{,}68268.$$

Таким образом, предельное распределение оценки m^* степени многочлена (полинома) является геометрическим. Это означает, в частности, что оценка не является состоятельной. При этом вероятность получить меньшее значение, чем истинное, исчезающе мала. Далее имеем:

$$P(m^* = m_0) \to 0{,}68268, \quad P(m^* = m_0 + 1) \to 0{,}68268(1 - 0{,}68268) = 0{,}21663,$$

$$P(m^* = m_0 + 2) \to 0{,}68268(1 - 0{,}68268)^2 = 0{,}068744,$$

$$P(m^* = m_0 + 3) \to 0{,}68268(1 - 0{,}68268)^3 = 0{,}021814...$$

Разработаны и иные методы оценивания неизвестной степени многочлена, например, с помощью многократного применения процедуры проверки адекватности регрессионной зависимости с помощью статистики Фишера (см. работу [3]). Предельное поведение оценок - таково же, как в приведенной выше теореме, только значение параметра λ иное.

Линейный и непараметрические парные коэффициенты корреляции. Термин "корреляция" означает "связь". В эконометрике этот термин обычно используется в сочетании "коэффициенты корреляции".

Рассмотрим способы измерения связи между двумя случайными переменными. Пусть исходными данными является набор случайных векторов $(x_i, y_i) = (x_i(\omega), y_i(\omega)),\ i = 1,2,...n.$ Коэффициентом корреляции, более подробно, линейным парным коэффициентом корреляции К. Пирсона называется (см. приложение 1 в конце настоящей книги)

$$r_n = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}.$$

Если $r_n = 1$, то $y_i = ax_i + b$, причем $a>0$. Если же $r_n = -1$, то $y_i = ax_i + b$, причем $a<0$. Таким образом, близость коэффициента корреляции к 1 (по абсолютной величине) говорит о достаточно тесной линейной связи.

Коэффициенты корреляции типа r_n используются во многих алгоритмах многомерного статистического анализа эконометрических данных. В теоретических рассмотрениях часто считают, что случайный вектор имеет многомерное нормальное распределение. Распределения реальных данных, как правило, отличны от нормальных (см. главу 4). Почему же распространено представление о многомерном нормальном распределении? Дело в том, что теория в этом случае проще. В частности, равенство 0 теоретического коэффициента корреляции (см. приложение 1) эквивалентно независимости случайных величин. Поэтому проверка независимости сводится к проверке статистической гипотезы о равенстве 0 теоретического коэффициента корреляции. Эта гипотеза принимается, если $|r_n| < C(n,\alpha)$, где $C(n,\alpha)$ - некоторое граничное значение, зависящее от объема выборки n и уровня значимости α.

Если случайные вектора $(x_i, y_i) = (x_i(\omega), y_i(\omega)),\ i = 1,2,...n,$ независимы и одинаково распределены, то выборочный коэффициент корреляции сходится к теоретическому при безграничном возрастании объема выборки:

$$r_n \to \rho = \frac{M(x_1 - M(x_1))(y_1 - M(y_1))}{\sqrt{D(x_1)}\sqrt{D(y_1)}}$$

(сходимость по вероятности).

Более того, выборочный коэффициент корреляции является асимптотически нормальным. Это означает, что

$$\lim_{n\to\infty} P\left(\frac{r_n - \rho}{\sqrt{D_0(r_n)}} < x\right) = \Phi(x),$$

где $\Phi(x)$ - функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1, а $D_0(r_n)$ - асимптотическая дисперсия выборочного коэффициента корреляции. Она имеет довольно сложное выражение, приведенное в монографии [4, с.393]:

$$D_0(r_n) = \frac{\rho^2}{4n}\left(\frac{\mu_{40}}{\mu_{20}^2} + \frac{\mu_{04}}{\mu_{02}^2} + \frac{2\mu_{22}}{\mu_{20}\mu_{02}} + \frac{4\mu_{22}}{\mu_{11}^2} - \frac{4\mu_{31}}{\mu_{11}\mu_{20}} - \frac{4\mu_{13}}{\mu_{11}\mu_{02}}\right).$$

Здесь под μ_{km} понимаются теоретические центральные моменты порядка *k* и *m*, а именно,

$$\mu_{km} = M(x_1 - M(x_1))^k (y_1 - M(y_1))^m$$

(см. приложение 1 в конце книги).

Для расчета непараметрического *коэффициента ранговой корреляции Спирмена* необходимо сделать следующее. Для каждого x_i рассчитать его ранг r_i в вариационном ряду, построенном по выборке $x_1, x_2, ...x_n$. Для каждого y_i рассчитать его ранг q_i в вариационном ряду, построенном по выборке $y_1, y_2, ...y_n$. Для набора из n пар (r_i, q_i), $i = 1,2,...,n$, вычислить (линейный) коэффициент корреляции. Он называется коэффициентом ранговой корреляции, поскольку определяется через ранги. В качестве примера рассмотрим данные из табл.2 (см. монографию [5]).

Табл.2. Данные для расчета коэффициентов корреляции

i	1	2	3	4	5
x_i	5	10	15	20	25
y_i	6	7	30	81	300
r_i	1	2	3	4	5
q_i	1	2	3	4	5

Для данных табл.2 коэффициент линейной корреляции равен 0,83, непосредственной линейной связи нет. А вот коэффициент ранговой корреляции равен 1, поскольку увеличение одной переменной однозначно соответствует увеличению другой переменной. Во многих экономических задачах, например, при выборе инвестиционных проектов для осуществления, достаточно именно монотонной зависимости одной переменной от другой.

Поскольку суммы рангов и их квадратов нетрудно подсчитать, то *коэффициент ранговой корреляции Спирмена* равен

$$\rho_n = 1 - \frac{6\sum_{i=1}^{n}(r_i - q_i)^2}{n^3 - n}.$$

Отметим, что *коэффициент ранговой корреляции Спирмена* остается постоянным при любом строго возрастающем преобразовании шкалы измерения результатов наблюдений. Другими словами, он является адекватным в порядковой шкале (см. главу 3), как и другие ранговые статистики (см. статистики Вилкоксона, Смирнова, типа омега-квадрат для проверки однородности независимых выборок в главе 4 и общее обсуждение в главе 8).

Широко используется также коэффициент ранговой корреляции τ Кендалла, коэффициент ранговой конкордации Кендалла и Б. Смита и др. Наиболее подробное обсуждение этой тематики содержится в монографии [6], необходимые для практических расчетов таблицы имеются в справочнике [1]. Дискуссия о выборе вида коэффициентов корреляции продолжается до настоящего времени [5].

Непараметрическая регрессия. Рассмотрим общее понятие регрессии как условного математического ожидания. Пусть случайный вектор $(x(\omega), y(\omega))$ имеет плотность *p(x,y).* Как известно из любого курса теории вероятностей, плотность условного распределения $y(\omega)$ при условии $x(\omega) = x_0$ имеет вид

$$p(y \mid x) = p(y \mid x(\omega) = x_0) = \frac{p(x, y)}{\int\limits_{-\infty}^{+\infty} p(x, y) dy}.$$

Условное математическое ожидание, т.е. регрессионная зависимость, имеет вид

$$f(x) = \int\limits_{-\infty}^{+\infty} y p(y \mid x) dy = \frac{\int\limits_{-\infty}^{+\infty} y p(x, y) dy}{\int\limits_{-\infty}^{+\infty} p(x, y) dy}.$$

Таким образом, для нахождения оценок регрессионной зависимости достаточно найти оценки совместной плотности распределения вероятности $p_n(x, y)$ такие, что

$$p_n(x, y) \to p(x, y)$$

при $n \to \infty$. Тогда непараметрическая оценка регрессионной зависимости

$$f_n(x) = \frac{\int\limits_{-\infty}^{+\infty} y p_n(x, y) dy}{\int\limits_{-\infty}^{+\infty} p_n(x, y) dy}$$

при $n \to \infty$ будет состоятельной оценкой регрессии как условного математического ожидания

$$f_n(x) \to f(x).$$

Общий подход к построению непараметрических оценок плотности распределения вероятностей развит в главе 8 ниже.

5.3. Основные понятия теории классификации

При внедрении современных эконометрических и статистических методов в практику технико-экономических исследований, при разработке соответствующих программных продуктов невозможно обойтись без классификации этих методов. Естественно исходить из вида обрабатываемых данных. В соответствии с современными воззрениями делим эконометрику и прикладную статистику на четыре области:

- статистика случайных величин (одномерная статистика);
- многомерный статистический анализ;
- статистика временных рядов и случайных величин;

- статистика объектов нечисловой природы.

В первой области элемент выборки - число, во второй - вектор, в третьей - функция, в четвертой - объект нечисловой природы. Термин "объект нечисловой природы" относится к элементам математического пространства, не являющегося векторным (линейным). Их нельзя складывать, умножать на числа, в отличие от чисел, векторов и функций. Примерами являются бинарные отношения (упорядочения, разбиения на классы, толерантности); множества, нечеткие множества; результаты измерений в номинальной и порядковой шкалах (т.е. по качественным признакам), в частности булевы вектора; вектора разнотипных признаков; тексты и т.д. (подробнее см., например, главу 8).

Математический аппарат статистики объектов нечисловой природы базируется на использовании расстояний (мер близости, показателей различия) в пространствах таких объектов. Это вызвано отсутствием в таких пространствах операций суммирования, на которых основано большинство методов других областей статистики. Любые методы, использующие только расстояния (меры близости, показатели различия) между объектами, следует относить к статистике объектов нечисловой природы, поскольку такие методы могут работать с объектами произвольного пространства, если в нем задана метрика или ее аналоги.

Таким образом, весьма многие математические методы классификации объектов или признаков следует включать в статистику объектов нечисловой природы. Она является уже весьма развитой области прикладной математики. Ей посвящено несколько тысяч статей и книг.

В настоящем пункте рассматривается важное направление эконометрики и прикладной статистики – математические методы классификации. Основная их часть относится к статистике объектов нечисловой природы, а именно, методы классификации, основанные на расстояниях между объектами.

Основные направления в математической теории классификации. Какие научные исследования относить к этой теории? Исходя из потребностей специалиста, применяющего математические методы классификации, целесообразно принять, что сюда входят исследования, во-первых, отнесенные самими авторами к этой теории; во вторых, связанные с ней общностью тематики, хотя бы их авторы и не упоминали термин «классификация». Это предполагает ее сложную внутреннюю структуру.

В литературных источниках наряду с термином «классификация» в близких смыслах используются термины «группировка», «распознавание образов», «диагностика», «дискриминация», «сортировка» и др. Терминологический разнобой связан прежде всего с традициями научных кланов, к которым относятся авторы публикаций, а также с внутренним делением самой теории классификации.

В научных исследованиях по современной теории классификации можно выделить два относительно самостоятельных направления. Одно из них опирается на опыт таких наук, как биология, география, геология, и таких прикладных областей, как ведение классификаторов продукции и библиотечное дело. Типичные объекты рассмотрения - классификация химических элементов (таблица Д.И. Менделеева), биологическая систематика, универсальная десятичная классификация публикаций (УДК), классификатор товаров на основе штрих-кодов.

Другое направление опирается на опыт технических исследований, экономики, маркетинговых исследований, социологии, медицины. Типичные задачи - техническая и медицинская диагностика, а также, например, разбиение на группы отраслей промышленности, тесно связанных между собой, выделение групп однородной продукции. Обычно используются такие термины, как «распознавание образов» или «дискриминантный анализ». Это направление обычно опирается на математические модели; для проведения расчетов интенсивно используется ЭВМ. Однако относить его к математике столь же нецелесообразно, как астрономию или квантовую механику. Рассматриваемые математические модели можно и нужно изучать на формальном уровне, и такие исследования проводятся. Но направление в целом сконцентрировано на решении конкретных задач прикладных областей и вносит вклад в технические или экономические науки, медицину, социологию, но, как правило, не в математику. Использование математических методов как инструмента исследования нельзя относить к чистой математике.

В 60-х годах XX века внутри прикладной статистики достаточно четко оформилась область, посвященная методам классификации. Несколько модифицируя формулировки М. Дж. Кендалла и А. Стьюарта 1966 г. (см. русский перевод [7, с.437]), в теории классификации выделим три подобласти: дискриминация (дискриминантный анализ), кластеризация (кластер-анализ), группировка. Опишем эти подобласти.

В дискриминантном анализе классы предполагаются заданными - плотностями вероятностей или обучающими выборками. Задача состоит в том, чтобы вновь поступающий объект отнести в один из этих классов. У понятия «дискриминация» имеется много синонимов: диагностика, распознавание образов с учителем, автоматическая классификация с учителем, статистическая классификация и т.д.

При кластеризации и группировке целью является выявление и выделение классов. Синонимы: построение классификации, распознавание образов без учителя, автоматическая классификация без учителя, таксономия и др. Задача кластер-анализа состоит в выяснении по эмпирическим данным, насколько элементы "группируются" или распадаются на изолированные "скопления", "кластеры"(от *cluster* (англ.) - гроздь, скопление). Иными словами, задача - выявление естественного разбиения на классы, свободного от субъективизма исследователя, а цель - выделение групп однородных объектов, сходных между собой, при резком отличии этих групп друг от друга.

При группировке, наоборот, «мы хотим разбить элементы на группы независимо от того, естественны ли границы разбиения или нет» [7, с.437]. Цель по-прежнему состоит в выявлении групп однородных объектов, сходных между собой (как в кластер-анализе), однако «соседние» группы могут не иметь резких различий (в отличие от кластер-анализа). Границы между группами условны, не являются естественными, зависят от субъективизма исследователя. Аналогично при лесоустройстве проведение просек (границ участков) зависит от специалистов лесного ведомства, а не от свойств леса.

Задачи кластеризации и группировки принципиально различны, хотя для их решения могут применяться одни и те же алгоритмы. Важная для практической деятельности проблема состоит в том, чтобы понять, разрешима ли задача кластер-анализа для конкретных данных или возможна только их группировка, поскольку они достаточно однородны и не разбиваются на резко разделяющиеся между собой кластеры.

Как правило, в математических задачах кластеризации и группировки основное - выбор метрики, расстояния между объектами, меры близости, сходства, различия. Хорошо известно, что для любого заданного разбиения объектов на группы и любого $\varepsilon > 0$ можно указать метрику такую, что расстояния между объектами из одной группы будут меньше ε, а между объектами из разных групп - больше $1/\varepsilon$. Тогда любой разумный алгоритм кластеризации даст именно заданное разбиение.

Ситуация осложняется использованием одного и того же термина в разных смыслах. Термином "классификация" (и термином "диагностика") обозначают, по крайней мере, три разные вещи: процедуру построения классификации (и выделение классов, используемых при диагностике), построенную классификацию (систему выделенных классов) и процедуру ее использования (правила отнесения вновь поступающего объекта к одному из ранее выделенных классов). Другими словами, имеем естественную триаду: построение – изучение – использование классификации.

Как уже отмечалось, для построения системы диагностических классов используют разнообразные методы кластерного анализа и группировки объектов. Наименее известен второй член триады – изучение отношений эквивалентности, полученных в результате построения системы диагностических классов. Статистический анализ полученных, в частности экспертами, отношений эквивалентности - часть статистики бинарных отношений и тем самым - статистики объектов нечисловой природы. Помимо общих результатов этой области эконометрики и прикладной статистики, представляют интерес частные результаты, полученные специально для отношений эквивалентности (см. главу 8)).

Диагностика в узком смысле слова (процедура использования классификации, т.е. отнесения вновь поступающего объекта к одному из выделенных ранее классов) - предмет дискриминантного анализа. Отметим, что с точки зрения статистики объектов нечисловой природы дискриминантный анализ является частным случаем общей схемы регрессионного анализа, соответствующим ситуации, когда зависимая переменная принимает конечное число значений, а именно - номера классов, а вместо квадрата разности стоит функция потерь от неправильной классификации. Однако есть ряд специфических постановок, выделяющих задачи диагностики среди всех регрессионных задач.

О построении диагностических правил. Начнем с обсуждения одного распространенного заблуждения. Иногда рекомендуют сначала построить систему диагностических классов, а потом в каждом диагностическом классе отдельно проводить регрессионный анализ (в классическом смысле) или применять иные методы многомерного статистического анализа. Однако обычно забывают, что при этом нельзя опираться на вероятностную модель многомерного нормального распределения, так как распределение результатов наблюдений, попавших в определенный кластер, будет отнюдь не нормальным, а усеченным нормальным (усечение определяется границами кластера).

Процедуры построения диагностических правил делятся на вероятностные и детерминированные. К первым относятся так называемые задачи расщепления смесей. В них предполагается, что распределение вновь поступающего случайного элемента является смесью вероятностных законов, соответствующих диагностическим классам. Как и при выборе степени полинома в регрессии (см. предыдущий пункт настоящей главы), при анализе реальных социально-экономических данных встает вопрос об оценке числа элементов смеси, т.е. числа диагностических классов. Были изучены результаты применения обычно рекомендуемого критерия Уилкса для оценки числа элементов смеси. Оказалось (см. статью [8]), что оценка с помощью критерия Уилкса не является состоятельной, асимптотическое распределение этой оценки – геометрическое, как и в случае задачи восстановления зависимости в регрессионном анализе (см. выше). Итак, продемонстрирована несостоятельность обычно используемых оценок. Для получения состоятельных оценок достаточно связать уровень значимости в критерии Уилкса с объемом выборки, как это было предложено и для задач регрессии.

Как уже отмечалось, задачи построения системы диагностических классов целесообразно разбить на два типа: с четко разделенными кластерами (задачи кластер-анализа) и с условными границами, непрерывно переходящими друг в друга классами (задачи группировки). Такое деление полезно, хотя в обоих случаях могут применяться одинаковые алгоритмы. Сколько же существует алгоритмов построения системы диагностических правил? Иногда называют то или иное число. На самом же деле их бесконечно много, в чем нетрудно убедиться.

Действительно, рассмотрим один определенный алгоритм - алгоритм средней связи. Он основан на использовании некоторой меры близости *d(x,y)* между объектами *x* и *y*. Как он работает? На первом шаге каждый объект рассматривается как отдельный кластер. На каждом следующем шаге объединяются две ближайших кластера. Расстояние между объектами рассчитывается как средняя связь (отсюда и название алгоритма), т.е. как среднее арифметическое расстояний между парами объектов, один из которых входит в первый кластер, а другой - во второй. В конце концов все объекты объединяются вместе, и результат работы алгоритма представляет собой дерево последовательных объединений (в терминах теории графов), или "Дендрограмму". Из нее можно выделить кластеры разными способами. Один подход - исходя из заданного числа кластеров. Другой - из соображений предметной области. Третий - исходя из устойчивости (если разбиение долго не менялось при возрастании порога объединения - значит оно отражает реальность). И т.д.

К алгоритму средней связи естественно сразу добавить алгоритм ближайшего соседа (когда расстоянием между кластерами называется минимальное из расстояний между парами объектов, один из которых входит в первый кластер, а другой - во второй) и алгоритм дальнего соседа (когда расстоянием между кластерами называется максимальное из расстояний между парами объектов, один из которых входит в первый кластер, а другой - во второй).

Каждый из трех описанных алгоритмов (средней связи, ближайшего соседа, дальнего соседа), как легко проверить, порождает бесконечное (континуальное) семейство алгоритмов кластер-анализа. Дело в том, что величина $d^a(x,y)$, $a>0$, также является мерой близости между x и y и порождает новый алгоритм. Если параметр a пробегает отрезок, то получается бесконечно много алгоритмов классификации.

Каким из них пользоваться при обработке данных? Дело осложняется тем, что практически в любом пространстве данных мер близости различных видов существует весьма много. Именно в связи с обсуждаемой проблемой следует указать на принципиальное различие между кластер-анализом и задачами группировки.

Если классы реальны, естественны, существуют на самом деле, четко отделены друг от друга, то любой алгоритм кластер-анализа их выделит. Следовательно, *в качестве критерия естественности классификации следует рассматривать устойчивость относительно выбора алгоритма кластер-анализа.*

Проверить устойчивость можно, применив к данным несколько подходов, например, столь непохожие алгоритмы, как «ближнего соседа» и «дальнего соседа». Если полученные результаты содержательно близки, то они адекватны действительности. В противном случае следует предположить, что естественной классификации не существует, задача кластер-анализа не имеет решения, и можно проводить только группировку.

Как уже отмечалось, часто применяется т.н. агломеративный иерархический алгоритм "Дендрограмма", в котором вначале все элементы рассматриваются как отдельные кластеры, а затем на каждом шагу объединяются два наиболее близких кластера. Для работы «Дендрограммы» необходимо задать правило вычисления расстояния между кластерами. Оно вычисляется через расстояние $d(x,y)$ между элементами x и y. Поскольку $d^a(x,y)$ при $0<a<1$ также расстояние, то, как правило, существует бесконечно много различных вариантов этого алгоритма. Представим себе, что они применяются для обработки одних и тех же реальных данных. Если при всех a получается одинаковое разбиение элементов на кластеры, т.е. результат работы алгоритма устойчив по отношению к изменению a (в смысле общей схемы устойчивости, рассмотренной в главе 10 ниже), то имеем «естественную» классификацию. В противном случае результат зависит от субъективно выбранного исследователем параметра a, т.е. задача кластер-анализа неразрешима (предполагаем, что выбор a нельзя специально обосновать). Задача группировки в этой ситуации имеет много решений. Из них можно выбрать одно по дополнительным критериям.

Следовательно, получаем эвристический критерий: если решение задачи кластер-анализа существует, то оно находится с помощью любого алгоритма. Целесообразно использовать наиболее простой.

Проблема поиска естественной классификации. Существуют различные точки зрения на эту проблему. На Всесоюзной школе-семинаре «Использование математических методов в задачах классификации» (г. Пущино, 1986 г.), в частности, были высказаны мнения, что естественная классификация:

- закон природы;
- основана на глубоких закономерностях, тогда как искусственная классификация - на неглубоких;
- для конкретного индивида та, которая наиболее быстро вытекает из его тезауруса;
- удовлетворяет многим целям; цель искусственной классификации задает человек;
- классификация с точки зрения потребителя продукции;
- классификация, позволяющая делать прогнозы;
- имеет критерием устойчивость.

Приведенные высказывания уже дают представление о больших расхождениях в понимании «естественной классификации». Этот термин следует признать нечетким, как, впрочем, и многие другие термины, как социально-экономические, научно-технические, так и используемые в обыденном языке. Нетрудно подробно обоснована нечеткость естественного языка и тот факт, что "мы мыслим нечетко", что однако не слишком мешает нам решать производственные и жизненные проблемы. Кажущееся рациональным требование выработать сначала строгие определения, а потом развивать науку - невыполнимо. Следовать ему - значит отвлекать силы от реальных задач. При системном подходе к теории классификации становится ясно, что строгие определения можно надеяться получить на последних этапах построения теории. Мы же сейчас находимся чаще всего на первых этапах. Поэтому, не давая определения понятиям «естественная классификация»и «естественная диагностика», обсудим, как проверить на «естественность» классификацию (набор диагностических классов), полученную расчетным путем.

Можно выделить два критерия «естественности», по поводу которых имеется относительное согласие:

А. Естественная классификация должна быть реальной, соответствующей действительному миру, лишенной внесенного исследователем субъективизма;

Б. Естественная классификация должна быть важной или с научной точки зрения (давать возможность прогноза, предсказания новых свойств, сжатия информации и т.д.), или с практической.

Пусть классификация проводится на основе информации об объектах, представленной в виде матрицы «объект-признак» или матрицы попарных расстояний (мер близости). Пусть алгоритм классификации дал разбиение на кластеры. Как можно получить доводы в пользу естественности этой классификации? Например, уверенность в том, что она - закон природы, может появиться только в результате длительного ее изучения и практического применения. Это соображение относится и к другим из перечисленных выше критериев, в частности к Б (важности). Сосредоточимся на критерии А (реальности).

Понятие «реальности» кластера требует специального обсуждения. (оно начато в работе [8]). Рассмотрим существо различий между понятиями «классификация» и «группировка». Пусть, к примеру, необходимо деревья, растущие в определенной местности, разбить на группы находящихся рядом друг с другом. Ясна интуитивная разница между несколькими отдельными рощами, далеко отстоящими друг от друга и разделенными полями, и сплошным лесом, разбитым просеками на квадраты с целью лесоустройства. Однако формально определить эту разницу столь же сложно, как определить понятие «куча зерен», чем занимались еще в Древней Греции (одно зерно не составляет кучи, два зерна не составляют кучи,…, если к тому, что не составляет кучи, добавить еще одно зерно, то куча не получится; значит - по принципу математической индукции - никакое количество зерен не составляет кучи; но ясно, что миллиард зерен - большая куча зерен - подсчитайте объем!).

Переформулируем сказанное в терминах "кластер-анализа" и "методов группировки". Выделенные с помощью первого подхода кластеры реальны, а потому могут рассматриваться как кандидаты в "естественные". Группировка дает "искусственные" классы, которые не могут быть "естественными".

Выборку из унимодального распределения можно, видимо, рассматривать как "естественный", "реальный" кластер. Применим к ней какой-либо алгоритм классификации ("средней связи", "ближайшего соседа" и т.п.). Он даст какое-то разбиение на классы, которые, разумеется, не являются "реальными", поскольку отражают прежде всего свойства алгоритма, а не исходных данных. Как отличить такую ситуацию от противоположной, когда имеются реальные кластеры и алгоритм классификации более или менее точно их выделяет? Как известно, "критерий истины – практика", но слишком много времени необходимо для применения подобного критерия. Поэтому представляет интерес критерий, оценивающий "реальность" выделяемых с помощью алгоритма классификации кластеров одновременно с его применением.

Такой показатель существует - это критерий устойчивости. Устойчивость - понятие широкое. Общая схема формулирования и изучения проблем устойчивости рассмотрена в главе 10. В частности, поскольку значения признаков всегда измеряются с погрешностями, то "реальное" разбиение должно быть устойчиво (т.е. не меняться или меняться слабо) при малых отклонениях исходных данных. Алгоритмов классификации существует бесконечно много, и "реальное" разбиение должно быть устойчиво по отношению к переходу к другому алгоритму. Другими словами, если "реальное" разбиение на диагностические классы возможно, то оно находится с помощью любого алгоритма автоматической классификация. Следовательно, критерием естественности классификации может служить совпадение результатов работы двух достаточно различающихся алгоритмов, например "ближайшего соседа" и "дальнего соседа".

Выше рассмотрены два типа "глобальных" критериев "естественности классификации", касающихся разбиения в целом. "Локальны"» критерии относятся к отдельным кластерам. Простейшая постановка такова: достаточно ли однородны два кластера (две совокупности) для их объединения:? Если объединение возможно, то кластеры не являются "естественными". Преимущество этой постановки в том, что она допускает применение статистических критериев однородности двух выборок. В одномерном случае (классификация по одному признаку) разработано большое число подобных критериев — Крамера-Уэлча, Смирнова, омега-квадрат (Лемана-Розенблатта), Вилкоксона, Ван-дер-Вардена, Лорда, Стьюдента и др. (см. главу 4 и справочник [1]). Имеются критерии и для многомерных данных. Для одного из видов объектов нечисловой природы - люсианов - статистические методы выделения "реальных" кластеров развиты в работе [9].

Что касается глобальных критериев, то для изучения устойчивости по отношению к малым отклонениям исходных данных естественно использовать метод статистических испытаний и проводить расчеты по "возмущенным" данным. Некоторые теоретические утверждения, касающиеся влияния «возмущений» на кластеры различных типов, получены в работе [8].

Опишем практический опыт реализации анализа устойчивости. Несколько алгоритмов классификации были применены к данным, полученным при проведении маркетинга образовательных услуг и приведенным в работе [10]. Для анализа данных были использованы широко известные алгоритмы "ближайшего соседа", "дальнего соседа" и алгоритм кластер-анализа из работы [11]. С содержательной точки зрения полученные разбиения отличались мало. Поэтому есть основания считать, что с помощью этих алгоритмов действительно выявлена «реальная» структура данных.

Идея устойчивости как критерия "реальности" иногда реализуется неадекватно. Так, для однопараметрических алгоритмов один из специалистов предлагал выделять разбиения, которым соответствуют наибольшие интервалы устойчивости по параметру, т.е. наибольшие приращения параметра между очередными объединениями кластеров. Для данных работы [10] это предложение не дало полезных результатов - были получены различные разбиения: три алгоритма - три разбиения. И с теоретической точки зрения предложение этого специалиста несостоятельно. Покажем это.

Действительно, рассмотрим алгоритм "ближайшего соседа", использующий меру близости $d(x,y)$, и однопараметрическое семейство алгоритмов с мерой близости $d^a(x,y)$, $a>0$, также являющихся алгоритмами "ближайшего соседа". Тогда дендрограммы, полученные с помощью этих алгоритмов, совпадают при всех a, поскольку при их реализации происходит лишь сравнение мер близости между объектами. Другими словами, дендрограмма, полученная с помощью алгоритма «ближайшего соседа», является адекватной в порядковой шкале (измерения меры близости $d(x,y)$), т.е. сохраняется при любом строго возрастающем преобразовании этой меры (см. главу 3). Однако выделенные по обсуждаемому методу "устойчивые разбиения" меняются. В частности, при достаточно большом a "наиболее объективным" в соответствии с предложением этого специалиста будет, как нетрудно показать, разбиение на два кластера! Таким образом, разбиение, выдвинутое им как "устойчивое", на самом деле оказывается весьма неустойчивым.

5.4. Эконометрика классификации

Рассмотрим несколько конкретных эконометрических вопросов теории классификации.

Вероятностная теория кластер-анализа. Как и для прочих методов эконометрики и прикладной статистики, свойства алгоритмов кластер-анализа необходимо изучать на вероятностных моделях. Это касается, например, условий естественного объединения двух кластеров.

Вероятностные постановки нужно применять, в частности, при перенесении результатов, полученных по выборке, на генеральную совокупность. Вероятностная теория кластер-анализа и методов группировки различна для исходных данных типа таблиц «объект x признак» и матриц сходства. Для первых параметрическая вероятностно-статистическая теория называется "расщеплением смесей". Непараметрическая теория основана на непараметрических оценках плотностей вероятностей и их мод. Основные результаты, связанные с непараметрическими оценками плотности, обсуждаются ниже (глава 8).

Если исходные данные - матрица сходства $||d(x,y)||$, то необходимо признать, что развитой вероятностно-статистической теории пока нет. Подходы к ее построению обсуждались в работе [8]. Одна из основных проблем - проверка "реальности" кластера, его объективного существования независимо от расчетов исследователя. Проблема "реальности" кластера давно обсуждается специалистами различных областей. Типичное рассуждение таково. Предположим, что результаты наблюдений можно рассматривать как выборку из некоторого распределения с монотонно убывающей плотностью при увеличении расстояния от некоторого центра. Примененный к подобным данным какой-либо алгоритм кластер-анализа порождает некоторое разбиение. Ясно, что оно - чисто формальное, поскольку выделенным таксонам (кластерам) не соответствуют никакие "реальные" классы. Другими словами, задача кластер-анализа не имеет решения, а алгоритм дает лишь группировку. При обработке реальных данных мы не знаем вида плотности. Проблема состоит в том, чтобы определить, каков результат работы алгоритма (реальные кластеры или формальные группы).

Частный случай этой проблемы - проверка обоснованности объединения двух кластеров, которые мы рассматриваем как два множества объектов, а именно, множества $\{a_1, a_2,..., a_k\}$ и $\{b_1, b_2,..., b_m\}$. Пусть, например, используется алгоритм типа "Дендрограмма". Естественной представляется следующая идея. Пусть есть две совокупности мер близости: одна - меры близости между объектами, лежащими внутри одного кластера, т.е. $d(a_i,a_j)$, $1\leq i<j\leq k$, $d(b_\alpha,b_\beta)$, $1\leq\alpha<\beta\leq m$, и другая - меры близости между объектами, лежащими в разных кластерах, т.е. $d(a_i,b_\alpha)$, $1\leq i\leq k$, $1\leq\alpha\leq m$. Эти две совокупности мер близости предлагается рассматривать как независимые выборки и проверять гипотезу о совпадении их функций распределения. Если гипотеза не отвергается, объединение кластеров считается обоснованным; в противном случае - объединять нельзя, алгоритм прекращает работу.

В рассматриваемом подходе есть две некорректности (см. также работу [8, разд.4]). Во-первых, меры близости не являются независимыми случайными величинами. Во-вторых, не учитывается, что объединяются не заранее фиксированные кластеры (с детерминированным составом), а полученные в результате работы некоторого алгоритма, и их состав (в частности, количество элементов) оказывается случайным От первой из этих некорректностей можно частично избавиться. Справедливо следующее утверждение.

Теорема 1. *Пусть $a_1, a_2,..., a_k, b_1, b_2,..., b_m$ - независимые одинаково распределенные случайные величины (со значениями в произвольном пространстве). Пусть случайная величина $d(a_1,a_2)$ имеет все моменты. Тогда при $k,m\to\infty$ распределение статистики*

$$\frac{8\sqrt{3}U-3(k+m)(k+m-1)(k(k+1)+m(m+1))}{2(k+m)\sqrt{km(k^2+m^2)}}$$

(где U - сумма рангов элементов первой выборки в объединенной выборке; первая выборка составлена из внутрикластерных расстояний (мер близости) $d(a_i,a_j)$, $1\leq i<j\leq k$, и $d(b_\alpha,b_\beta)$, $1\leq\alpha<\beta\leq m$, а вторая - из межкластерных расстояний $d(a_i,b_\alpha)$, $1\leq i\leq k$, $1\leq\alpha\leq m$) сходится к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1.

На основе теоремы 1 очевидным образом формулируется правило проверки обоснованности объединения двух кластеров. Другими словами, мы проверяем статистическую гипотезу, согласно которой объединение двух кластеров образует однородную совокупность. Если величина U слишком мала, статистическая гипотеза однородности отклоняется (на заданном уровне значимости), и возможность объединения отбрасывается. Таким образом, хотя расстояния между объектами в кластерах зависимы, но эта зависимость слаба, и доказана математическая теорема о допустимости применения критерия Вилкоксона для проверки возможности объединения кластеров.

О вычислительной сходимости алгоритмов кластер-анализа. Алгоритмы кластер-анализа и группировки зачастую являются итерационными. Например, формулируется правило улучшения решения задачи кластер-анализа шаг за шагом, но момент остановки вычислений не обсуждается. Примером является известный алгоритм "Форель", в котором постепенно улучшается положение центра кластера. В этом алгоритме на каждом шагу строится шар определенного заранее радиуса, выделяются элементы кластеризуемой совокупности, попадающие в этот шар, и новый центр кластера строится как центр тяжести выделенных элементов. При анализе алгоритма «Форель» возникает проблема: завершится ли процесс улучшения положения центра кластера через конечное число шагов или же он может быть бесконечным. Она получила название «проблема остановки». Для широкого класса так называемых "эталонных алгоритмов" проблема остановки была решена в работе [8]: процесс улучшения остановится через конечное число шагов.

Отметим, что алгоритмы кластер-анализа могут быть модифицированы разнообразными способами. Например, описывая алгоритм "Форель" в стиле статистики объектов нечисловой природы, заметим, что вычисление центра тяжести для совокупности многомерных точек – это нахождение эмпирического среднего для меры близости, равной квадрату евклидова расстояния. Если взять более естественную меру близости – само евклидово расстояние, то получим алгоритм кластер-анализа "Медиана", отличающийся от "Форели" тем, что новый центр строится не с помощью средних арифметических координат элементов, попавших в кластер, а с помощью медиан.

Проблема остановки возникает не только при построении диагностических классов. Она принципиально важна, в частности, и при оценивании параметров вероятностных распределений методом максимального правдоподобия. Обычно не представляет большого труда выписать систему уравнений максимального правдоподобия и предложить решать ее каким-либо численным методом. Однако когда остановиться, сколько итераций сделать, какая точность оценивания будет при этом достигнута? Общий ответ, видимо, невозможно найти, но обычно нет ответа и для конкретных семейств распределения вероятностей. Именно поэтому мы нет оснований рекомендовать решать системы уравнений максимального правдоподобия, вместо них целесообразно использовать т.н. одношаговые оценки (подробнее см. об этих оценках работу [12]). Эти оценки задаются конечными формулами, но асимптотически столь же хороши (на профессиональном языке - эффективны), как и оценки максимального правдоподобия.

О сравнении алгоритмов диагностики по результатам обработки реальных данных. Перейдем к этапу применения диагностических правил, когда классы, к одному из которых нужно отнести вновь поступающий объект, уже выделены.

В прикладных эконометрических исследованиях применяют различные методы дискриминантного анализа, основанные на вероятностно-статистических моделях, а также с ними не связанные, т.е. эвристические, использующие детерминированные методы анализа данных. Независимо от "происхождения", каждый подобный алгоритм должен быть исследован как на параметрических и непараметрических вероятностно-статистических моделях порождения данных, так и на различных массивах реальных данных. Цель исследования - выбор наилучшего алгоритма в определенной области применения, включение его в стандартные программные продукты, методические материалы, учебные программы и пособия. Но для этого надо уметь сравнивать алгоритмы по качеству. Как это делать?

Часто используют такой показатель качества алгоритма диагностики, как "вероятность правильной классификации" (при обработке конкретных данных - "частота правильной классификации"). Чуть ниже мы покажем, что этот показатель качества некорректен, а потому пользоваться им не рекомендуется. Целесообразно применять другой показатель качества алгоритма диагностики - оценку специального вида т.н. "расстояния Махаланобиса" между классами. Изложение проведем на примере разработки программного продукта для специалистов по диагностике материалов. Прообразом является диалоговая система «АРМ материаловеда», разработанная Институтом высоких статистических технологий и эконометрики для ВНИИ эластомерных материалов.

При построении информационно-исследовательской системы диагностики материалов (ИИСДМ) возникает задача сравнения прогностических правил «по силе». Прогностическое правило - это алгоритм, позволяющий по характеристикам материала прогнозировать его свойства. Если прогноз дихотомичен («есть» или «нет»), то правило является алгоритмом диагностики, при котором материал относится к одному из двух классов. Ясно, что случай нескольких классов может быть сведен к конечной последовательности выбора между двумя классами.

Прогностические правила могут быть извлечены из научно-технической литературы и практики. Каждое из них обычно формулируется в терминах небольшого числа признаков, но наборы признаков сильно меняются от правила к правилу. Поскольку в ИИСДМ должно фиксироваться лишь ограниченное число признаков, то возникает проблема их отбора. Естественно отбирать лишь те их них, которые входят в наборы, дающие наиболее «надежные» прогнозы. Для придания точного смысла термину «надежный» необходимо иметь способ сравнения алгоритмов диагностики по прогностической "силе".

Результаты обработки реальных данных с помощью некоторого алгоритма диагностики в рассматриваемом случае двух классов описываются долями: правильной диагностики в первом классе κ; правильной диагностики во втором классе λ; долями классов в объединенной совокупности π_i, $i = 1.2$; $\pi_1 + \pi_2 = 1$.

Величины $\kappa, \lambda, \pi_1, \pi_2$ определяются ретроспективно.

Нередко как показатель качества алгоритма диагностики (прогностической «силы») используют долю правильной диагностики

$$\mu = \pi_1 \kappa + \pi_2 \lambda.$$

Однако показатель μ определяется, в частности, через характеристики π_1 и π_2, частично заданные исследователем (например, на них влияет тактика отбора образцов для изучения). В аналогичной медицинской задаче величина μ оказалась больше для тривиального прогноза (у всех больных течение заболевания будет благоприятно), чем для использованного в работе [13] группы под руководством академика АН СССР И.М. Гельфанда алгоритма выделения больных с прогнозируемым тяжелым течением заболевания, применение которого с медицинской точки зрения вполне оправдано. Другими словами, по доле правильной классификации алгоритм академика И.М. Гельфанда оказался хуже тривиального - объявить всех больных легкими, не требующими специального наблюдения. Этот вывод нелеп. И причина появления нелепости понятна. Хотя доля тяжелых больных невелика, но смертельные исходы сосредоточены именно в этой группе больных. Поэтому целесообразна гипердиагностика - рациональнее часть легких больных объявить тяжелыми, чем сделать ошибку в противоположную сторону. Применение теории статистических решений в рассматриваемой постановке вряд ли возможно, поскольку оценить количественно потери от смерти больного нельзя по этическим соображениям. Поэтому, на наш взгляд, долю правильной диагностики μ нецелесообразно использовать как показатель качества алгоритма диагностики.

Применение теории статистических решений требует знания потерь от ошибочной диагностики, а в большинстве научно-технических и экономических задач определить потери, как уже отмечалось, сложно. В частности, из-за необходимости оценивать человеческую жизнь в денежных единицах. По этическим соображениям это, на наш взгляд, недопустимо. Сказанное не означает отрицания пользы страхования, но, очевидно, страховые выплаты следует рассматривать лишь как способ первоначального смягчения потерь от утраты близких.

Для выявления информативного набора признаков целесообразно использовать метод пересчета на модель линейного дискриминантного анализа, согласно которому статистической оценкой прогностической "силы" является

$$\delta^* = \Phi(d^*/2), \ \ d^* = \Phi^{-1}(\kappa) + \Phi^{-1}(\lambda),$$

где $\Phi(x)$ - функция стандартного нормального распределения вероятностей с математическим ожиданием 0 и дисперсией 1, а $\Phi^{-1}(y)$ - обратная ей функция.

Если классы описываются выборками из многомерных нормальных совокупностей с одинаковыми матрицами ковариаций, а для классификации применяется классический линейный дискриминантный анализ Р.Фишера, то величина $d*$ представляет собой состоятельную статистическую оценку так называемого расстояния Махаланобиса между рассматриваемыми двумя совокупностями (конкретный вид этого расстояния сейчас не имеет значения), независимо от порогового значения, определяющего конкретное решающее правило. В общем случае показатель δ^* вводится как эвристический.

Пусть алгоритм классификации применяется к совокупности, состоящей из *m* объектов первого класса и *n* объектов второго класса.

Теорема 2. *Пусть m, n→∞. Тогда для всех x*

$$P\left\{\frac{\delta*-\delta}{A(\kappa,\lambda)}<x\right\}\to\Phi(x),$$

где δ *- истинная "прогностическая сила" алгоритма диагностики;* $\delta*$ *- ее эмпирическая оценка,*

$$A^2(\kappa,\lambda)=\frac{1}{4}\left\{\left[\frac{\varphi(d*/2)}{\varphi(\Phi^{-1}(\kappa))}\right]^2\frac{\kappa(1-\kappa)}{m}+\left[\frac{\varphi(d*/2)}{\varphi(\Phi^{-1}(\lambda))}\right]^2\frac{\lambda(1-\lambda)}{n}\right\};$$

$\varphi(x)=\Phi'(x)$*) - плотность стандартного нормального распределения вероятностей с математическим ожиданием 0 и дисперсией 1.*

С помощью теоремы 2 по κ и λ обычным образом определяют доверительные границы для "прогностической силы" δ.

Как проверить обоснованность пересчета на модель линейного дискриминантного анализа? Допустим, что классификация состоит в вычислении некоторого прогностического индекса *y* и сравнении его с заданным порогом *c*; объект относят к первому классу, если $y\le c$, ко второму, если $y>c$. Возьмем два значения порога c_1 и c_2. Если пересчет на модель линейного дискриминантного анализа обоснован, то "прогностические силы" для обоих правил совпадают: $\delta(c_1)=\delta(c_2)$. Эту статистическую гипотезу можно проверить.

Пусть κ_1 - доля объектов первого класса, для которых $y\le c_1$, а κ_2 - доля объектов первого класса, для которых $c_1<y\le c_2$. Аналогично пусть λ_2 - доля объектов второго класса, для которых $c_1<y\le c_2$, а λ_3 - доля объектов второго класса, для которых $y>c_2$. Тогда можно рассчитать две оценки одного и того же расстояния Махаланобиса. Они имеют вид:

$$d*(c_1)=\Phi^{-1}(\kappa_1)+\Phi^{-1}(\lambda_2+\lambda_3),\quad d*(c_2)=\Phi^{-1}(\kappa_1+\kappa_2)+\Phi^{-1}(\lambda_3).$$

Теорема 3. *Если истинные прогностические силы двух правил диагностики совпадают,* $\delta(c_1)=\delta(c_2)$, *то при* $m\to\infty, n\to\infty$ *при всех x*

$$P\left\{\frac{d*(c_1)-d*(c_2)}{B}<x\right\}\to\Phi(x),$$

где

$$B^2=\frac{1}{m}T(\kappa_1;\kappa_2)+\frac{1}{n}T(\lambda_3;\lambda_2);$$

$$T(x;y)=\frac{x(1-x)}{\varphi^2(\Phi^{-1}(x))}+\frac{(x+y)(1-x-y)}{\varphi^2(\Phi^{-1}(x+y))}-\frac{2x(1-x-y)}{\varphi(\Phi^{-1}(x))\varphi(\Phi^{-1}(x+y))}.$$

Из теоремы 3 вытекает метод проверки рассматриваемой гипотезы: при выполнении неравенства

$$\left|\frac{d*(c_1)-d*(c_2)}{B}\right|\le\Phi^{-1}(1-\frac{\alpha}{2})$$

она принимается на уровне значимости, асимптотически равном α, в противном случае - отвергается.

Подходы к построению прогностических правил. Для решения задач диагностики используют два подхода – параметрический и непараметрический. Первый из них обычно основан на использовании того или иного индекса и сравнения его с порогом. Индекс может быть построен по статистическим данным, например, как в уже упомянутом линейном дискриминантном анализе Фишера. Часто индекс представляет собой линейную функцию от характеристик, выбранных специалистами предметной области, коэффициенты которой подбирают эмпирически. Непараметрический подход связан с леммой Неймана-Пирсона в математической статистике и с теорией статистических решений. Он опирается на использование непараметрических оценок плотностей распределений вероятностей, описывающих диагностические классы.

Обсудим ситуацию подробнее. Математические методы диагностики, как и статистические методы в целом, делятся на параметрические и непараметрические. Первые основаны на предположении, что классы описываются распределениями из некоторых параметрических семейств. Обычно рассматривают многомерные нормальные распределения, при этом зачастую принимают гипотезу о том, что ковариационные матрицы для различных классов совпадают. Именно в таких предположениях сформулирован классический дискриминантный анализ Фишера. Как известно, обычно нет оснований считать, что наблюдения извлечены из нормального распределения.

Поэтому более корректными, чем параметрические, являются непараметрические методы диагностики. Исходная идея таких методов основана на лемме Неймана-Пирсона, входящей в стандартный курс математической статистики. Согласно этой лемме решение об отнесении вновь поступающего объекта (сигнала, наблюдения и др.) к одному из двух классов принимается на основе отношения плотностей $f(x)/g(x)$, где $f(x)$ - плотность распределения, соответствующая первому классу, а $g(x)$ - плотность распределения, соответствующая второму классу. Если плотности распределения неизвестны, то применяют их непараметрические оценки, построенные по обучающим выборкам. Пусть обучающая выборка объектов из первого класса состоит из n элементов, а обучающая выборка для второго класса - из m объектов. Тогда рассчитывают значения непараметрических оценок плотностей $f_n(x)$ и $g_m(x)$ для первого и второго классов соответственно, а диагностическое решение принимают по их отношению. Таким образом, для решения задачи диагностики достаточно научиться строить непараметрические оценки плотности для выборок объектов произвольной природы.

Методы построения непараметрических оценок плотности распределения вероятностей в пространствах произвольной природы рассмотрены в главе 8.

Цитированная литература

1. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. – М.: Наука, 1983. - 416 с.
2. Себер Дж. Линейный регрессионный анализ. - М.: Мир, 1980. - 456 с.
3. Орлов А.И. Оценка размерности модели в регрессии. – В сб.: Алгоритмическое и программное обеспечение прикладного статистического анализа. Ученые записки по статистике, т.36. - М.: Наука, 1980. - С.92-99.
4. Крамер Г. Математические методы статистики. - М.: Мир, 1975. - 648 с.
5. Красильников В.В. Статистика объектов нечисловой природы. - Наб. Челны: Изд-во Камского политехнического института, 2001. - 144 с.
6. Кендэл М. Ранговые корреляции. - М.: Статистика, 1975. - 216 с.

7. Кендалл М.Дж., Стьюарт А. Многомерный статистический анализ и временные ряды. - М.: Наука, 1976. – 736 с.
8. Орлов А.И. Некоторые вероятностные вопросы теории классификации. – В сб.: Прикладная статистика. Ученые записки по статистике, т.45. - М.: Наука, 1983. – С.166-179.
9. Орлов А.И. Парные сравнения в асимптотике Колмогорова. – В сб.: Экспертные оценки в задачах управления. - М.: Изд-во ИПУ, 1982. - С. 58-66.
10. Орлов А.И.; Гусейнов Г.А. Математические методы в изучении способных к математике школьников – В сб.: Исследования по вероятностно-статистическому моделированию реальных систем. - М.: ЦЭМИ АН СССР, 1977. - С.80-93.
11. Куперштох В.Л., Миркин Б.Г., Трофимов В.А. Сумма внутренних связей как показатель качества классификации // Автоматика и телемеханика. 1976. № 3. С.91-98.
12. Орлов А.И. О нецелесообразности использования итеративных процедур нахождения оценок максимального правдоподобия. // Заводская лаборатория. 1986. Т.52. № 5. С.67-69.
13. Гельфанд И.М., Алексеевская М.А., Губерман Ш.А. и др. Прогнозирование исхода инфаркта миокарда с помощью программы "Кора-3" // Кардиология. 1977. Т.17. № 6. С.19-23.

Глава 6. Эконометрика временных рядов

Под временными рядами понимают экономические величины, зависящие от времени. При этом время предполагается дискретным, в противном случае говорят о случайных процессах, а не о временных рядах.

6.1. Модели стационарных и нестационарных временных рядов, их идентификация

Пусть $t = 0, \pm 1, \pm 2, \pm 3, ...$ Рассмотрим временной ряд *X(t).* Пусть сначала временной ряд принимает числовые значения. Это могут быть, например, цены на батон хлеба в соседнем магазине или курс обмена доллара на рубли в ближайшем обменном пункте. Обычно в поведении временного ряда выявляют две основные тенденции - тренд и периодические колебания.

При этом под трендом понимают зависимость от времени линейного, квадратичного или иного типа, которую выявляют тем или иным способом сглаживания (например, экспоненциального сглаживания) либо расчетным путем, в частности, с помощью метода наименьших квадратов. Другими словами, тренд - это очищенная от случайностей основная тенденция временного ряда.

Временной ряд обычно колеблется вокруг тренда, причем отклонения от тренда часто обнаруживают правильность. Часто это связано с естественной или назначенной периодичностью, например, сезонной или недельной, месячной или квартальной (например, в соответствии с графиками выплаты заплаты и уплаты налогов). Иногда наличие периодичности и тем более ее причины неясны, и задача эконометрика - выяснить, действительно ли имеется периодичность.

Элементарные методы оценки характеристик временных рядов обычно достаточно подробно рассматриваются в курсах "Общей теории статистики" (см., например, учебники [1,2]), поэтому нет необходимости подробно разбирать их здесь. (Впрочем, о некоторых современных методах оценивания длины периода и самой периодической составляющей речь пойдет ниже.)

Характеристики временных рядов. Для более подробного изучения временных рядов используются вероятностно-статистические модели. При этом временной ряд *X(t)* рассматривается как случайный процесс (с дискретным временем) основными характеристиками являются математическое ожидание *X(t)*, т.е.

$$a(t) = MX(t)\ ,$$

дисперсия *X(t)*, т.е.

$$\sigma^2(t) = DX(t)$$

и *автокорреляционная функция* временного ряда *X(t)*

$$\rho(t,s) = \frac{M(X(t) - a(t))(X(s) - a(s))}{\sigma(t)\sigma(s)},$$

т.е. функция двух переменных, равная коэффициенту корреляции между двумя значениями временного ряда *X(t)* и *X(s).*

В теоретических и прикладных исследованиях рассматривают широкий спектр моделей временных рядов. Выделим сначала *стационарные* модели. В них совместные функции распределения $F(t_1, t_2, ..., t_k)$ для любого числа моментов времени *k*, а потому и все перечисленные выше характеристики временного ряда *не меняются со временем*. В частности, математическое ожидание и дисперсия являются постоянными величинами, автокорреляционная функция зависит только от разности *t-s.* Временные ряды, не являющиеся стационарными, называются *нестационарными.*

Линейные регрессионные модели с гомоскедастичными и гетероскедастичными, независимыми и автокоррелированными остатками. Как видно из сказанного выше, основное - это "очистка" временного ряда от случайных отклонений, т.е. оценивание математического ожидания. В отличие от простейших моделей регрессионного анализа, рассмотренных в главе 5, здесь естественным образом появляются более сложные модели. Например, дисперсия может зависеть от времени. Такие модели называют гетероскедастичными, а те, в которых нет зависимости от времени - гомоскедастичными. (Точнее говоря, эти термины могут относиться не только к переменной "время", но и к другим переменным.)

Далее, в главе 5 предполагалось, что погрешности независимы между собой. В терминах настоящей главы это означало бы, что автокорреляционная функция должна быть вырожденной - равняться 1 при равенстве аргументов и 0 при их неравенстве. Ясно, что для реальных временных рядов так бывает отнюдь не всегда. Если естественный ход изменений наблюдаемого процесса является достаточно быстрым по сравнению с интервалом между последовательными наблюдениями, то можно ожидать "затухания" автокорреляции" и получения практически независимых остатков, в противном случае остатки будут автокоррелированы.

Идентификация моделей. Под идентификацией моделей обычно понимают выявление их структуры и оценивание параметров. Поскольку структура - это тоже параметр, хотя и нечисловой (см. главу 8), то речь идет об одной из типовых задач эконометрики - оценивании параметров.

Проще всего задача оценивания решается для линейных (по параметрам) моделей с гомоскедастичными независимыми остатками. Восстановление зависимостей во временных рядах может быть проведено на основе методов наименьших квадратов и наименьших модулей, рассмотренных в главе 5 моделей линейной (по параметрам) регрессии. На случай временных рядов переносятся результаты, связанные с оцениванием необходимого набора регрессоров, в частности, легко получить предельное геометрическое распределение оценки степени тригонометрического полинома.

Однако на более общую ситуацию такого простого переноса сделать нельзя. Так, например, в случае временного ряда с гетероскедастичными и автокоррелированными остатками снова можно воспользоваться общим подходом метода наименьших квадратов, однако система уравнений метода наименьших квадратов и, естественно, ее решение будут иными. Формулы в терминах матричной алгебры, о которых упоминалось в главе 5, будут отличаться. Поэтому рассматриваемый метод называется "*обобщенный метод наименьших квадратов* (ОМНК)" (см., например, [3, с.212]).

Замечание. Как уже отмечалось в главе 5, простейшая модель метода наименьших квадратов допускает весьма далекие обобщения, особенно в области системам одновременных эконометрических уравнений для временных рядов. Для понимания соответствующей теории и алгоритмов необходимо профессиональное владение матричной алгеброй. Поэтому мы отсылаем тех, кому это интересно, к литературе по системам эконометрических уравнений [4-9] и непосредственно по временным рядам [10-25], в которой особенно много интересуются спектральной теорией, т.е. выделением сигнала из шума и разложением его на гармоники. Подчеркнем в очередной раз, что за каждой главой настоящей книги стоит большая область научных и прикладных исследований, вполне достойная того, чтобы посвятить ей много усилий. Однако из-за ограниченности объема книги мы вынуждены изложение сделать конспективным.

6.2. Системы эконометрических уравнений

Пример модели авторегрессии. В качестве первоначального примера рассмотрим эконометрическую модель временного ряда, описывающего рост индекса потребительских цен (индекса инфляции). Пусть $I(t)$ - рост цен в месяц t (подробнее об этой проблематике см. главу 7). Тогда по мнению некоторых экономистов естественно предположить, что

$$I(t) = cI(t-1) + a + b\,S(t-4) + e, \qquad (1)$$

где *I(t- 1)* - рост цен в предыдущий месяц (а *с* - некоторый коэффициент затухания, предполагающий, что при отсутствии внешний воздействий рост цен прекратится), *а* - константа (она соответствует линейному изменению величины *I(t)* со временем), *b S (t* - 4) - слагаемое, соответствующее влиянию эмиссии денег (т.е. увеличения объема денег в экономике страны, осуществленному Центральным Банком) в размере *S (t* - 4) и пропорциональное эмиссии с коэффициентом *b*, причем это влияние проявляется не сразу, а через 4 месяца; наконец, е - это неизбежная погрешность.

Модель (1), несмотря на свою простоту, демонстрирует многие характерные черты гораздо более сложных эконометрических моделей. Во-первых, обратим внимание на то, что некоторые переменные определяются (рассчитываются) внутри модели, как *I(t).* Их называют *эндогенными (внутренними).* Другие задаются извне (это *экзогенные* переменные). Иногда, как в теории управления, среди экзогенных переменных, выделяют *управляемые* переменные - те, с помощью которых менеджер может привести систему в нужное ему состояние.

Во-вторых, в соотношении (1) появляются переменные новых типов - с лагами, т.е. аргументы в переменных относятся не к текущему моменту времени, а к некоторым прошлым моментам.

В-третьих, составление эконометрической модели типа (1) - это отнюдь не рутинная операция. Например, запаздывание именно на 4 месяца в связанном с эмиссией денег слагаемом *b S (t* - 4) - это результат достаточно изощренной предварительной статистической обработки. Далее, требует изучения вопрос зависимости или независимости величин *S (t* - 4) и *I(t).* От решения этого вопроса зависит, как выше уже отмечалось, конкретная реализация процедуры метода наименьших квадратов.

С другой стороны, в модели (1) всего 3 неизвестных параметра, и постановку метода наименьших квадратов выписать нетрудно:

$$f(a,b,c) = \sum_{1 \le t \le k} (I(t) - cI(t-1) - a - bS(t-4))^2.$$

Проблема идентифицируемости. Представим теперь модель тапа (1) с большим числом эндогенных и экзогенных переменных, с лагами и сложной внутренней структурой. Вообще говоря, ниоткуда не следует, что существует хотя бы одно решение у такой системы. Поэтому возникает не одна, а две проблемы. Есть ли хоть одно решение (проблема идентифицируемости)? Если да, то как найти наилучшее решение из возможных? (Это - проблема статистической оценки параметров.)

И первая, и вторая задача достаточно сложны. Для решения обоих задач разработано множество методов, обычно достаточно сложных (см. список литературы), лишь часть из которых имеет научное обоснование. В частности, достаточно часто пользуются статистическими оценками, не являющимися состоятельными (строго говоря, их даже нельзя назвать оценками).

Коротко опишем некоторые распространенные приемы при работе с системами линейных эконометрических уравнений.

Система линейных одновременных эконометрических уравнений. Чисто формально можно все переменные выразить через переменные, зависящие только от текущего момента времени. Например, в случае уравнения (1) достаточно положить

$$H(t) = I(t-1),\ G(t) = S(t-4).$$

Тогда уравнение пример вид

$$I(t) = cH(t) + a + b\,G(t) + e. \qquad (2)$$

Отметим здесь же возможность использования регрессионных моделей с переменной структурой путем введения фиктивных переменных. Эти переменные при одних значениях времени (скажем, начальных) принимают заметные значения, а при других - сходят на нет (становятся фактически равными 0). В результате формально (математически) одна и та же модель описывает совсем разные зависимости.

Косвенный, двухшаговый и трехшаговый методы наименьших квадратов. Как уже отмечалось, разработана масса методов эвристического анализа систем эконометрических уравнений. Они предназначены для решения тех или иных проблем, возникающих при попытках найти численные решения систем уравнений.

Одна из проблем связана с наличием априорных ограничений на оцениваемые параметры. Например, доход домохозяйства может быть потрачен либо на потребление, либо на сбережение. Значит, сумма долей этих двух видов трат априори равна 1. А в системе эконометрических уравнений эти доли могут участвовать независимо. Возникает мысль оценить их методом наименьших квадратов, не обращая внимания на априорное ограничение, а потом подкорректировать. Такой подход называют косвенным методом наименьших квадратов.

Двухшаговый метод наименьших квадратов состоит в том, что оценивают параметры отдельного уравнения системы, а не рассматривают систему в целом. В то же время трехшаговый метод наименьших квадратов применяется для оценки параметров системы одновременных уравнений в целом. Сначала к каждому уравнению применяется двухшаговый метод с целью оценить коэффициенты и погрешности каждого уравнения, а затем построить оценку для ковариационной матрицы погрешностей, После этого для оценивания коэффициентов всей системы применяется обобщенный метод наименьших квадратов (см. выше).

Менеджеру и экономисту не следует становиться специалистом по составлению и решению систем эконометрических уравнений, даже с помощью тех или иных программных систем, но он должен быть осведомлен о возможностях этого направления эконометрики, чтобы в случае производственной необходимости квалифицированно сформулировать задание для специалистов-эконометриков.

От оценивания тренда (основной тенденции) перейдем ко второй основной задаче эконометрики временных рядов - оцениванию периода (цикла).

6.3. Оценивание длины периоды и периодической составляющей

В настоящем пункте рассмотрим достаточно широкий класс практически полезных непараметрических оценок длины периода и периодической составляющей во временных рядах. Из общих результатов статистики объектов нечисловой природы (см. главу 8) вытекает состоятельность этих оценок.

Начнем с того, что во многих прикладных задачах рассматривают временной ряд (или случайный процесс) $y(t)=x(t)+e(t)$, где $x(t)$ - детерминированная периодическая функция от времени t, т.е. $x(t)=x(t+T)$ при некотором T, где T - длина периода (минимальная из возможных, поскольку $2T$, $3T$, $4T$ - тоже, как легко видеть, длины периодов), а $e(t)$ - “шумы”, случайные погрешности, искажающие периодический сигнал. Требуется оценить (минимальную) длину периода T и периодическую составляющую $x(t)$. При этом не предполагается, что функция $x(t)$ входит в какое-либо параметрическое семейство, например, конечных сумм синусов и косинусов, т.е. рассматривается задача непараметрического оценивания (минимальной) длины периода и периодической составляющей сигнала.

Приведем примеры прикладных постановок.

1. По акустическим сигналам необходимо установить тип двигателя (и его национальную принадлежность). Предполагается, что двигатели различаются по длине периода и виду основного периодического сигнала. Процедура идентификации основана на оценивании длины периода и периодической составляющей регистрируемого сигнала. Очевидна важность такой задачи при быстрой технической диагностике. В частности, высокая производительность, а потому и высокая экономическая эффективность при ремонте впрямую зависят от умения решать поставленную задачу.

2. В предположении цикличности экономических процессов требуется по статистическим данным установить длину цикла и на основе вида периодической составляющей построить прогноз, например, прогноз урожайности, емкости рынка тех или иных товаров или экономической активности в целом. В экономической литературе часто говорят об экономических циклах, но почти никогда не дают строгого определения понятия цикла. (Под строгим определением понимаем такое, согласно которому можно отличить "цикл" от "не цикла", можно выделить начало и конец цикла, отделить один цикл от другого, короче, однозначно выделить цикл как самостоятельный объект экономического изучения.)
3. По мнению авторов работы [26], для среднесрочного прогнозирования развития социокультурной сферы (социально-политического “климата”, живописи, музыки, архитектуры, поэзии и т.д.) необходимо выявить ее цикличность с помощью объективных измерений на базе субъективных первичных данных (т.е. на базе оценок экспертов).
4. В исторических событиях, описываемых согласно распространенной в настоящее время т.н. скалигеровской хронологии, автор работы [27] обнаруживает цикличность. Эта цикличность полностью объясняется новой статистической хронологией (см., например, [28]), построенной с помощью специальных методов статистики объектов нечисловой природы (см. главу 8), предназначенных для анализа текстов исторических хроник, и одновременно служит еще одним подтверждением новой статистической хронологии.

Описание метода оценивания. Пусть рассматриваемые функции $y(t)$, $x(t)$, $e(t)$ определены на отрезке $[0; A]$. При фиксированном T рассмотрим “куски” сигнала $y(t)$ на последовательных отрезках длины T, т.е. на отрезках $[0;T]$, $[T;2T]$, $[2T;3T]$, ... Удобно ввести последовательность функций на отрезке $[0;T]$, полученную сдвигами этих кусков к началу координат:

$$y_1(t)=y(t),\ y_2(t)=y(t+T),\ y_3(t)=y(t+2T),\ \ldots$$

Все они определены на отрезке $[0;T]$. Число этих функций равно числу полных периодов длины T, укладывающихся на отрезке $[0;A]$, т.е. равно целой части числа A/T. Отметим еще раз, что если T - период, то $2T$, $3T$, $4T$, ... - тоже периоды. В дальнейшем из всех периодов будем рассматривать и оценивать, как правило, только наименьший.

Если $T=T_0$ - истинный период (или кратный ему) и погрешности $e(t)$ отсутствуют, то все введенные в предыдущем абзаце функции совпадают между собой и с периодической составляющей:

$$x(t)=y_1(t)=y_2(t)=y_3(t)=\ldots$$

при всех t из $[0;T]$. При наличии погрешностей полного совпадения не будет. Однако отклонения определяются лишь шумами в различные моменты времени. При этом в качестве оценки периодической составляющей $x(t)$ естественно взять среднее арифметическое $y_{cp}(t)$ функций $y_1(t)$, $y_2(t)$, $y_3(t)$, ... (могут быть использованы и другие виды средних величин).

Если же T отличается от истинного периода T_0 (и кратных ему величин), то различия функций $y_1(t)$, $y_2(t)$, $y_3(t)$, ... между собой определяются также и различием значений $x(t)$ в точках, отстоящих друг от друга на интервалы, длина которых кратна T.

В предположении отсутствия погрешностей (т.е. когда $e(t)$ тождественно равно 0) рассмотрим поведение функции $y_{cp}(t)$ на отрезке $[0;T]$ при росте длины интервала A наблюдения сигнала, а потому и при росте числа периодов - целой части числа A/T. Если $T = T_0$ или T кратно T_0, то, как уже сказано, $y_{cp}(t)$ совпадает с периодической составляющей $x(t)$. Если число T/T_0 иррационально, то можно показать, что значения $t+mT(modT_0)$, где m - натуральные числа такие, что $t+mT<A$, асимптотически (при росте A) равномерно заполняют отрезок $[0;T_0]$, а потому при выполнении соответствующих условий регулярности, например, непрерывности периодической составляющей сигнала, функция $y_{cp}(t)$ приближается к константе - среднему значению периодического сигнала $x(t)$, т.е. интегралу от $x(t)$ по отрезку $[0;T_0]$, деленному на T_0. При этом при конечных A функция $y_{cp}(t)$ отлична от константы. (Здесь запись $t+mT(modT_0)$ означает теоретико-числовое

сравнение по модулю T_0 , т.е. взятие дробной части от числа $(t+mT)/T_0$, что соответствует вычитанию соответствующего количества целых п. периодов T_0 .)

Если же число T/T_0 рационально, то наблюдаем промежуточный случай по сравнению с двумя описанными выше, в котором $y_{cp}(t)$, как можно показать, приближается к периодической функции с периодом $T=T_0/n$ при некотором натуральном n. Эта функция получена усреднением n последовательных участков длины T_0/n периодического сигнала $x(t)$. Она не является константой, хотя разброс ее значений меньше, чем для исходного периодического сигнала, поскольку T_0 - минимальная длина периода.

Из сказанного вытекает, что для оценивания T целесообразно ввести два показателя: показатель разброса $F(T;Y)=F(T;\ y_1(t),\ y_2(t),\ y_3(t),\ \ldots\ )$ множества функций $\{y_1(t),\ y_2(t),\ y_3(t),\ \ldots\ \}$ на отрезке $[0;T]$ и показатель размаха $G(T;Y)=G(T,\ y_{cp}(t))$ функции $y_{cp}(t)$ на отрезке $[0;T]$. (Символ Y означает здесь, что показатели разброса и размаха строятся по функции $y(t)$.) При этом показатель разброса нацелен на оценку различий в значениях семейства функций при одном и том же значении аргумента, а показатель размаха - на различие значений одной и той же функции при различных значениях аргумента. Ниже выписан ряд формул для этих показателей в случае непрерывного времени. Для дискретного времени их можно адаптировать двумя способами: либо заменив sup на max, а интеграл на сумму; либо расширив область определения используемых функций на весь отрезок, например, соединив соседние точки отрезками или использовав для заполнения пропусков сплайны более высокого порядка.

В качестве оценки длины периода по фиксированным показателям разброса $F(T;Y)$ и размаха $G(T;Y)$ представляется рациональным использовать то T, при котором отношение $F(T;Y)/G(T;Y)$ **впервые** (при росте T начиная с 0) достигает минимума (впервые - поскольку величины, кратные периоду, сами являются периодами). Поскольку показатели разброса $F(T;Y)$ и размаха $G(T;Y)$ могут быть выбраны многими разными способами, можно указанным выше способом построить целое семейство алгоритмов оценивания длины периода, с каждым из которых может быть связано семейство методов оценивания периодической составляющей путем того или иного способа усреднения функций $y_1(t),\ y_2(t),\ y_3(t),\ \ldots$

Показатели разброса и размаха. Ввести показатели разброса $F(T;Y)=F(T;\ y_1(t),\ y_2(t),\ y_3(t),\ \ldots\ )$ можно разными способами. Пусть $k=[A/T]$. Можно использовать различные функционалы супремумного типа (здесь и далее число слагаемых k не будем указывать в обозначении функционалов). Первым рассмотрим максимальный разброс непосредственно между значениями функций:

$$F_1(T,Y)=\sup\{|\ y_i(t)-y_j(t)\ |, i, j=1,2,...,k, 0\le t\le T\}.$$

Второй функционал супремумного типа будет учитывать не произвольные отклонения, а только отклонения от "средней функции", т.е. иметь вид

$$F_2(T,Y)=\sup\{|\ y_i(t)-y_{cp}(t)\ |, i=1,2,...,k, 0\le t\le T\}.$$

Третий функционал показывает, какую зону "заметают" значения функций:

$$F_3(T,Y)=\sup\{y_i(t), i=1,2,...,k, 0\le t\le T\}-\inf\{y_i(t), i=1,2,...,k, 0\le t\le T\}.$$

Для применения функционалов интегрального типа целесообразно сделать замену переменной $q=t/T$ и перейти к функциям $Y_i(q)=y_i(t)=y_i(qT),\ i=1,2,\ldots,k,\ Y_{cp}(q)=y_{cp}(t)=y_{cp}(qT)$, определенным на отрезке [0;1]. В качестве показателя разброса представляется полезным рассмотреть то или иное отклонение совокупности функций $Y_i(q)$, i=1,2,...,k, друг относительно друга. Можно сказать, что эти функции заполняют некую “трубку”, которая тоньше всего при истинном значении периода T, а внутри нее проходит периодическая составляющая $X(q)=x(t)=x(qT)$. Естественно рассмотреть различные функционалы интегрального типа. Например, можно проинтегрировать максимум модулей попарных разностей:

$$F_4(T,Y)=\int_0^1 \max\{|Y_i(q)-Y_j(q)|, i,j=1,2,...,k\}dq.$$

Вместо максимума можно проинтегрировать сумму:

$$F_5(T,Y)=\int_0^1 \sum_{i,j=1}^{k} |Y_i(q)-Y_j(q)|\,dq.$$

Как и для функционалов супремумного типа, естественно рассмотреть показатели разброса относительно "средней функции":

$$F_6(T,Y)=\int_0^1 \max\{|Y_i(q)-Y_{cp}(q)|, i=1,2,...,k\}dq,$$

$$F_7(T,Y)=\int_0^1 \sum_{i=1}^{k} |Y_i(q)-Y_{cp}(q)|\,dq.$$

Следующие четыре функционала, используемые как показатели разброса, аналогичны четырем предыдущим, но включают в себя расчет квадратов:

$$F_8(T,Y)=\int_0^1 [\max\{|Y_i(q)-Y_j(q)|, i,j=1,2,...,k\}]^2\,dq,$$

$$F_9(T,Y)=\int_0^1 \sum_{i,j=1}^{k} \{Y_i(q)-Y_j(q)\}^2\,dq,$$

$$F_6(T,Y)=\int_0^1 [\max\{|Y_i(q)-Y_{cp}(q)|, i=1,2,...,k\}]^2\,dq,$$

$$F_7(T,Y)=\int_0^1 \sum_{i=1}^{k} \{Y_i(q)-Y_{cp}(q)\}^2\,dq.$$

Список показателей разброса можно существенно расширить. В частности, естественно использовать также расстояния в функциональных пространствах L^p при произвольных $p\geq 1$, а для оценивания периодической составляющей применять не только среднее арифметическое, но и другие виды средних.

Показатели размаха также можно ввести самыми различными способами. Например, можно рассмотреть такой показатель:

$$G_1(T,Y)=\sup\{|y_{cp}(t)-y_{cp}(s)|,\ 0\le t\le T,\ 0\le s\le T\}=$$
$$=\sup\{y_{cp}(t),\ 0\le t\le T\}-\inf\{y_{cp}(t),\ 0\le t\le T\}.$$

Пусть сделана замена переменной $q=t/T$ и осуществлен переход к функции $Y_{cp}(q)=y_{cp}(t)=y_{cp}(qT)$. Возможными показателями размаха являются:

$$G_2(T,Y)=\int_0^1\int_0^1 |Y_{cp}(q)-Y_{cp}(r)|\,dqdr,$$

$$G_2(T,Y)=\int_0^1\int_0^1 (Y_{cp}(q)-Y_{cp}(r))^2\,dqdr.$$

Введем среднее значение оценки периодической составляющей:

$$Y_{cp}=\int_0^1 Y_{cp}(q)dq.$$

К естественным показателям размаха относятся, например, такие:

$$G_4(T,Y)=\sup\{|Y_{cp}(q)-Y_{cp}|,\ 0\le q\le 1\},$$

$$G_5(T,Y)=\int_0^1 |Y_{cp}(q)-Y_{cp}|\,dq,$$

$$G_6(T,Y)=\int_0^1 (Y_{cp}(q)-Y_{cp})^2\,dq.$$

Список показателей размаха, как и список показателей разброса, можно значительно расширить. В частности, естественно использовать расстояния в функциональных пространствах L^p при произвольном $p\geq 1$, а для оценивания периодической составляющей применять не только среднее арифметическое, но и другие виды средних - медиану, среднее геометрическое и др. (см. главу 3). Вопрос о выборе наилучших (в каком-либо смысле) показателей размаха и разброса в настоящем пункте не обсуждается. Некоторые из причин этого отказа от оптимизации системы показателей рассмотрены ниже.

Алгоритмы оценивания. С прикладной точки зрения остается численно минимизировать один или несколько из 66 описанных выше функционалов

$F_i(T;Y)/G_j(T;Y)$, i=1,2,...,11, j=1,2,...,6.

Численная минимизация по одному параметру (возможной длине периода) для современных ЭВМ не вызывает проблем, даже если попросту перебирать возможные значения периода с шагом 0,001. По нескольким реальным или смоделированным сигналам можно установить, какой из функционалов позволяет оценить период и периодическую составляющую реально встречающихся сигналов наиболее точно. Возможно и одновременное использование всех или части функционалов, что в соответствии с методологией устойчивости (см. главу 10) позволяет установить чувствительность оценок к выбору метода оценивания, найти интервал их разброса. Проведенные в Институте высоких статистических технологий и эконометрики расчеты по реальным и смоделированным данным о временных рядах показали, что описанные выше алгоритмы позволяют оценивать длину периода и восстанавливать периодическую составляющую временного ряда достаточно точно с практической точки зрения.

В обширной литературе по временным рядам (см., например, монографии [10-25], дающие представление обо всем массиве литературы по этой тематике) проблеме оценивания периода не уделяется большого внимания. Фактически рекомендуют пользоваться либо периодограммой, либо автокорреляционной функцией. С помощью периодограммы (несостоятельной оценки спектральной плотности) можно выделить лишь синусоидальные составляющие, в то время как в кратко рассмотренных выше прикладных задачах периодическая составляющая представляет интерес сама по себе, без разложения на гармоники. Вторая рекомендация более полезна. В качестве оценки периода можно взять наименьшее положительное число, в котором достигается локальный максимум автокорреляционной функции. Эмпирический коэффициент автокорреляции - еще один функционал типа тех, что перечислены выше.

При поверхностном взгляде на проблемы статистического оценивания, как и на иные проблемы прикладной математики, часто возникает желание обсудить “оптимальность” тех или иных процедур. При более глубоком анализе становятся очевидными два обстоятельства. Во-первых, оптимальность имеет быть лишь в рамках той или иной теоретической модели, при отклонениях от которой оптимальность оценки, как правило, пропадает. Например, выборочное среднее арифметическое как оценка математического ожидания случайной величины оптимальна тогда и только тогда, когда распределение результатов наблюдений - гауссово (доказательство этого утверждения приведено в монографии [30]). С другой стороны, для практически любой статистической процедуры можно подобрать свойство оптимальности так, чтобы эта процедура оказалась оптимальной (как подобрать - это уже дело профессионала). Так, например, метод наименьших модулей оптимален, если погрешности имеют распределение Лапласа, а метод наименьших квадратов - когда их распределение гауссово. Поскольку реальные распределения - не Лапласа и не Гаусса, то указанные математические результаты не могут иметь большого практического значения.

Однако представляется полезным получить доказательства состоятельности оценок изучаемых параметров в возможно более широких, например, непараметрических, постановках. Хотя на основе самого факта сходимости нельзя оценить близость оценок к интересующим исследователя параметрам, но получение доказательства состоятельности - первый шаг при изучении скорости сходимости (подробнее об этом см. главу 10).

Состоятельность оценок. Наиболее общий подход к установлению асимптотического поведения решений экстремальных статистических задач развит в статистике объектов нечисловой природы для случая пространств произвольной природы (см. главу 8, а также работу [31]). Согласно этому подходу сначала при фиксированном T доказывается сходимость (по вероятности) при $\mathrm{A} \to \infty$ значений функционала (показателя разброса) к некоторой предельной функции, а затем проверяются условия, обеспечивающие сходимость *Argmin* допредельного случайного процесса к *Argmin* этой детерминированной функции.

Свойства алгоритмов приходится изучать в рамках тех или иных вероятностно-статистических моделей. Моделей может быть много. Достаточно вспомнить историю Центральной Предельной Теоремы (ЦПТ) теории вероятностей, которая на протяжении более 200 лет доказывалась во все более и более широких условиях, вплоть до необходимых и достаточных условий Линдеберга-Феллера (после чего начались обобщения на зависимые слагаемые, на суммы случайных элементов гильбертовых пространств и др.). Отметим, что иногда математические модели далеко выходят за пределы, достаточные для обоснования алгоритмов анализа реальных данных. Так, почти всегда распределения реальных величин дискретны и финитны, а потому, в частности, существуют все моменты. Однако условия финитности и дискретности в вероятностно-статистических моделях часто необоснованно ослабляются. В результате возникают проблемы, не имеющие отношения к реальным данным, например, связанные с измеримостью относительно тех или иных сигма-алгебр. Поэтому в настоящем пункте ограничимся наиболее простыми моделями из адекватных реальным постановкам. Считаем, что читатель знаком с основными определениями, относящимися к теории случайных процессов.

Теорема 1. Пусть случайный процесс $e(t)$ имеет нулевое математическое ожидание, является стационарным и эргодическим (т.е. выполнена теорема Биркгофа-Хинчина) с непрерывными траекториями. Тогда при фиксированном T и $\mathrm{A} \to \infty$ имеем

$$\sup\{|\, E_{cp}(q)\,|,\ 0 \le q \le 1\} \to 0$$

(сходимость по вероятности), где $E_{cp}(q)=Y_{cp}(q)-X_{cp}(q)$, т.е. $E_{cp}(q)$ - среднее арифметическое погрешностей $e(qT)$, $e(qT+T)$, $e(qT+2T)$,...

Доказательство теоремы 1 проводится стандартными методами теории стационарных временных рядов (с шагом T) с использованием известного условия достаточно быстрого убывания элементов матрицы Лорана по мере удаления от ее главной диагонали (т.е. условия, необходимого и достаточного для справедливости теоремы Биркгофа-Хинчина). С помощью теоремы 1 можно найти асимптотику введенных выше показателей разброса и размаха.

Теорема 2. В предположениях теоремы 1 при фиксированном T и $\mathrm{A} \to \infty$ пронормированные показатели разброса $F_i(T;Y)$ для наблюдаемого сигнала $У$ сближаются по распределению с соответствующими положительными случайными величинами $\mathrm{W}_i(\mathrm{T}, \mathrm{X}, \omega)$, зависящими от T, характеристик случайного процесса $e(t)$ и периодической составляющей X, т.е. существуют числовые последовательности $s_i(k)$ такие, что

$$s_i(k) F_i(T, Y) \Rightarrow W_i(T, X, \omega), \quad i = 1,2,...,11.$$

Доказательство теоремы 2 проводится с помощью достаточно трудоемких (в частности, из-за числа функционалов), но стандартных рассуждений (они относятся к теории случайных процессов как части теории вероятностей), посвященных максимумам (не супремумам, т.к. траектории функции $x(t)$ и случайного процесса $e(t)$ непрерывны) случайных процессов и интегралам от них, с использованием принципа инвариантности (см., например, учебное пособие [32]) и ряда результатов теории стационарных случайных процессов (см., например, монографию [19]). Таким образом, пронормированные функционалы разброса асимптотически не зависят от числа слагаемых - в этом и состоит основной смысл теоремы 2.

Теорема 3. В предположениях теоремы 1 при фиксированном T и $А \to \infty$ показатели размаха для наблюдаемого сигнала $У$ сближаются с соответствующими показателями для периодической составляющей X, т.е.

$$G_j(T,Y) - G_j(T,X) \to 0, \quad j = 1,2,...,6.$$

Для доказательства используются стандартные оценки, основанные на виде конкретных функционалов, задающих показатели размаха. В отличие от теоремы 2 предельные показатели детерминированы.

Аналоги теорем 2 и 3 верны также и при использовании (в качестве показателей разброса и размаха) расстояний в функциональных пространствах L^p при произвольном $p \geq 1$, а для оценивания периодической составляющей - не только среднего арифметического, но и других видов средних - медианы, среднего квадратического, среднего геометрического, обобщенных средних по Колмогорову (см. главу 3) и др.

Теорема 4. Пусть выполнены условия теоремы 1, периодическая составляющая непрерывна и имеет период T_0 . Тогда при фиксированном T и $А \to \infty$ показатели разброса (пронормированные) и размаха стремятся к некоторым детерминированным пределам, зависящим только от T и T_0 , т.е.

$$s_i(k) F_i(T,Y) \to F_i(T,T_0), \quad i = 1,2,...,11,$$

$$G_j(T,Y) \to G_j(T,X), \quad j = 1,2,...,6.$$

(сходимость по вероятности), минимум каждой из функций $F_i(T;T_0)$, i=1,2,...,11, и максимум каждой из функций $G_j(T;T_0)$, j=1,2,...,6, достигается при T=T_0 и при T, кратных T_0 , причем у показателей разброса $F_i(T;T_0)$ возможны и иные минимумы, а у показателей размаха $G_j(T;T_0)$ других максимумов нет.

Доказательство вытекает из теорем 2 и 3 и свойств усреднения периодической составляющей при росте длины интервала наблюдения сигнала, описанных в начале настоящего пункта. Отметим, что предельные значения функционала разброса $F_i(T;T_0)$, вообще говоря, показывают разброс случайной погрешности, другими словами, не всегда зависят от периодической составляющей, а потому из-за нормировки на единичный отрезок в ряде случаев оказываются константами. Вместе с тем численные эксперименты показывают, что отмеченная сходимость к пределу является сравнительно медленной, и минимизация непосредственно функционалов разброса (без учета показателей размаха) при конкретной длине сигнала позволяет достаточно точно выделить периодическую составляющую из массива реальных данных. Однако описанные выше теоретические результаты заставили отказаться от первоначальной гипотезы о том, что достаточно использовать только показатели разброса, и привели к необходимости скорректировать алгоритмы, введя деление на показатели размаха.

Теорема 5. В предположениях теоремы 4 **о**ценки, являющиеся первыми локальными минимумами при минимизации по T отношений одного из 11 перечисленных выше показателей разброса к одному из 6 показателей размаха, являются состоятельными оценками истинного периода T_0, а функция $y_{cp}(t)$ является состоятельной оценкой периодической составляющей $x(t)$ на отрезке $[0;T_0]$.

Согласно теоремам 1-4 установлена сходимость (по вероятности) значений допредельных функционалов к предельным при каждом конкретном T. Для доказательства сходимости минимумов допредельных функционалов к минимумам предельных можно воспользоваться общей теорией асимптотического поведения решения экстремальных статистических задач (см. главу 8 или работу [31]). Условие асимптотической равномерной разбиваемости сформулированное в работе [31], выполнено, как можно показать, в силу непрерывности траекторий случайного процесса (непрерывного сглаживания для временного ряда) и его периодической составляющей, откуда и вытекает заключение теоремы 5, дающей теоретико-статистическое обоснование использованию системы описанных выше эвристических алгоритмов оценивания длины периода и периодической составляющей. При известной или достаточно точно оцененной длине периода сама периодическая составляющая естественным образом оценивается с помощью усреднения перенесенных к началу координат кусков временного ряда, и в силу теоремы 1 эта оценка является состоятельной. Затем для получения оценки математического ожидания сигнала на всей области его определения указанную оценку можно периодически продолжить.

Замечание. При практическом использовании описанных в настоящем пункте алгоритмов целесообразно учитывать дополнительные особенности реальных временных рядов. В частности, обратим внимание на неустойчивость супремумов (в смысле главы 10 настоящей книги) по отношению к выбросам (резко выделяющимся наблюдениям) сравнительно с функционалами интегрального типа. Бывают ситуации, когда методики или аппаратура, регистрирующие значения реальных временных рядов, могут допускать сбои в отдельные моменты времени. Например, если происходит валютный кризис типа "черного вторника", когда курс доллара по отношению к рублю, строго говоря, не определен, другими словами, с точки зрения экономических агентов одновременно существует масса сильно отличающихся курсов. Аналогичная ситуация бывает и в целом ряде других случаев. Набор подходящих ассоциаций вызывают решения руководства страны об обмене денежных знаков, особенно с дискриминационными составляющими. Во всех подобных ситуациях временные ряды дают резкие выбросы (всплески), которые затем, как правило, сглаживаются. Поэтому целесообразно в качестве показателей разброса и размаха использовать функционалы интегрального типа. Вопросам оценивания длины периода и периодической составляющей посвящены многие публикации, в том числе работа [33].

6.4. Метод ЖОК оценки результатов взаимовлияний факторов

Различные субъекты и факторы экономической жизни постоянно влияют друг на друга. Как правило, для каждого из рассматриваемых экономических субъектов (и факторов) можно выделить "непосредственное окружение", которое оказывает на него влияние на него в конкретный момент. Как правило, на него же этот субъект оказывает некоторое обратное влияние. Дальше начинается самое интересное - волны влияний, порожденные разными субъектами, распространяются по всей совокупности, частично усиливают друг друга, частично погашают, порождая в каждый момент времени новые волны.

Разработан компьютерный метод (см. работу [34]), называемый далее ЖОК, предназначенный для оценки результатов влияния описывающих ситуацию факторов на итоговые показатели и друг на друга. Метод ЖОК позволяет получать выводы, полезные для управления различными экономическими структурами на микро- и макроуровнях, от бригад и предприятий до государства в целом. Этот метод использует экономико-математическую модель многомерного временного ряда, в которой коэффициенты непосредственного влияния факторов друг на друга и начальные условия задаются экспертами, т.е. представляет собой синтез экспертных и экономико-математических методов. Опишем основные составляющие этого метода.

Сначала экспертным путем определяется список факторов, которые необходимо учитывать при анализе конкретной ситуации. В качестве примера рассмотрим здесь типовое промышленное предприятие. Для него такими факторами являются, видимо, устойчивость развития, уровень рентабельности, оценка состояния основных и оборотных фондов, положение на рынке, кадровый потенциал, финансовое положение, технологический уровень, технический уровень и качество продукции, степень учета экологических требований, уровень сертификации, научно-технический потенциал и степень его использования, положение в социальной сфере, развитость профсоюзного движения, оценка отношений с конкурентами и властями, и т.д. Основная часть перечисленных факторов носит качественный характер.

Далее определяются необходимые для работы модели начальные уровни факторов, соответствующие современному (т.е. начальному) состоянию изучаемого экономического объекта (проводится оцифровка нечисловых переменных). Они оцениваются экспертами на шкале от (-1) до (+1) с шагом 0,1. В методе ЖОК степень привлечения экспертов может быть различна - от использования одного эксперта, хорошо знающего ситуацию и на основе своих знаний и интуиции указывающего необходимые параметры и связи, до подключения к работе комиссии экспертов, коллективно оценивающих указанные параметры и связи, с использованием той или иной схемы сбора и анализа экспертных мнений (см. главу 12).

Затем экспертами составляется блок-схема непосредственных влияний факторов друг на друга и оценивается степень непосредственных влияний с помощью такой же шкалы от (-1) до (+1) с шагом 0,1. Получается экономико-математическая модель в виде взвешенного ориентированного графа с начальными данными в вершинах. Она несколько напоминает хорошо известную экономистам схему межотраслевого баланса В.Леонтьева, но в отличие от нее использует не только количественные, но - в основном - качественные факторы. Затем просчитываются итерации (опосредованные влияния второго, третьего и т.д. уровней, соответствующие второму, третьему и т.д. моментам времени) вплоть до получения стабильного состояния. Результат работы модели - конечные уровни факторов.

Модель позволяет просчитать развитие экономической структуры при различных сценариях. Обычно одновременно используют три типа сценариев - “Прогноз”, “Поиск” и “Оптимизация”.

Сценарий “Прогноз” показывает результат при отсутствии управляющих воздействий. Он демонстрирует, как будет развиваться ситуация, если в нее не вмешиваться. Исходные данные для сценария “Прогноз” - начальные значения факторов и матрица непосредственных взаимовлияний факторов.

В сценариях типа “Поиск” вводится новое понятие - управляющие факторы. В сценариях этого типа анализируются результаты изменений при наличии тех или иных конкретных воздействий на управляющие факторы. Обычно специалист, работающий с системой ЖОК, имеет целью увеличение значений тех или иных факторов при “удержании” некоторых иных в заданных пределах. В сценариях типа “Поиск” осуществляется эвристический процесс оптимизации, а также анализ поведения системы при тех или иных воздействиях на начальные значения факторов.

В сценариях типа “Оптимизация” кроме списка управляющих факторов задаются целевые факторы и условия на них, которых необходимо добиться. Обычно это - условия выхода на определенные уровни, например, рентабельность должна быть не менее 0.5, а социальная напряженность - не более 0.3. С помощью оптимизационных алгоритмов находится наилучшее управление, позволяющее достигнуть цели или максимально к ней приблизиться. Однако найденные компьютером рекомендации могут включать слишком резкие изменения тех или иных начальных параметров, поэтому результаты расчетов скорее указывают на перспективные варианты изменения управляющих параметров, чем непосредственно задают план действий. С помощью сценариев типа “Поиск” можно на основе этих результатов найти практически реализуемые рекомендации.

Система ЖОК позволяет проследить динамику изменения значений факторов вплоть до их стабилизации, которая обычно наступает через 15-25 итераций (интервалов времени). Такая быстрая сходимость вначале кажется неожиданной. Возможно, сам факт стабилизации является самым важным методологическим выводом из экспериментов с моделью ЖОК: "После первоначальных всплесков замкнутая экономическая система стабилизируется, хотя бы и на весьма низком уровне производства и потребления."

При этом с помощью оцененных экспертами коэффициентов важности факторов (с учетом знака) можно отслеживать общую оценку экономической ситуации.

Система ЖОК является человеко-машинной. Для эффективной работы специалиста желательно, чтобы общее число факторов, используемых в конкретной модели, не превышало 20, а число непосредственных взаимосвязей - 40, хотя эти ограничения несущественны для математического обеспечения компьютерной системы ЖОК. Они существенны для наглядности при построении, обсуждении и совершенствовании модели, для того, чтобы факторы и связи между ними можно было изобразить на листе бумаги или экране компьютера в виде блок-схемы.

Система ЖОК с успехом использовалась для анализа ряда конкретных экономических ситуаций. Так, по заказу Минфина РФ она применялась для анализа взаимовлияний факторов, определяющих динамику налогооблагаемой базы и сбора подоходного налога с физических лиц, налога на имущество, налогов и сборов за пользование природными ресурсами и др. Построенная серия эконометрических моделей обладала некоторыми общими чертами. Прогноз, исходящий из современного экономического положения, во всех случаях указывал на дальнейшее ухудшение ситуации. Активное вмешательство государства в экономику приводило к значительному улучшению показателей, в то время как управление с помощью чисто экономических (монетаристских) методов не позволяло улучшить исходное положение. Полученные результаты подтверждают известную концепцию пяти нобелевских лауреатов по экономике (К.Эрроу, В.Леонтьев и др.), разрабатываемую совместно с Отделением экономики Российской академии наук (Д.С.Львов, С.Ю.Глазьев и др.), о необходимости активного регулирования государством экономических процессов.

Другие примеры применения системы ЖОК касались оптимизации экономической стороны деятельности промышленного предприятия или организации в иной сфере, экономических взаимоотношений отраслей народного хозяйства, а также макроэкономического моделирования, в ходе которого удалось вскрыть две неточности в основной схеме известной монографии К.Р.Макконнелла и С.Л.Брю “Экономикс: Принципы, проблемы и политика” [35], а затем исправить их, включив дополнительные блоки в соответствующую модель.

Эконометрический метод ЖОК может найти широкое применение для анализа экономического состояния и перспектив промышленных предприятий, банков, различных государственных и коммерческих структур.

Подведем итоги главы. Рассмотрены методы анализа и моделирования временных рядов. Они используются прежде всего для прогнозирования экономических явлений и процессов (см. главу 14). Надо отметить, что как самим временным рядам, так и вопросам их прогнозирования посвящена огромная литература. Дополнительно к названным выше монографиям укажем книги [36-42]. Наряду с вероятностно-статистическими методами при прогнозировании активно применяют экспертные методы (см. главу 12).

В настоящей главе рассмотрены лишь основы и отдельные вопросы эконометрики временных рядов - одной из наиболее обширных и сложных (с математико-статистической точки зрения) областей эконометрики. Читатель, желающий глубже познакомиться с этой специфической частью эконометрики, должен обратиться к литературе, в частности, указанной в конце главы.

Цитированная литература

1. Елисеева И.И., Юзбашев М.М. Общая теория статистики. - М.: Финансы и статистика., 1998. - 368 с.

2. Общая теория статистики. Статистическая методология в изучении коммерческой деятельности. / Под ред.А.А. Спирина, О.Э.Башиной. - М,: Финансы и статистика, 1994. - 296 с.
3. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Исследование зависимостей. - М.: Финансы и статистика, 1985. - 488 с.
4. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. - М.: Юнити, 1998. - 1022 с.
5. Доугерти К. Введение в эконометрику. - М.: МГУ, 1999. - 402 с.
6. Катышев П.К., Пересецкий А.А. Сборник задач к начальному курсу эконометрики. - М.: Дело, 1999. - 72 с.
6. Кулинич Е.И. Эконометрия. - М.: Финансы и статистика, 1999. - 302 с.
7. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. - М.: Дело, 1997. - 248 с.
8. Нейлор Т. Машинные имитационные эксперименты с моделями экономических систем. - М.: Мир, 1975. - 500 с.
9. Харин Ю.С., Малюгин В.Н. и др. Основы имитационного и статистического моделирования. - Минск: ДизайнПро, 1997. - 218 с.
10. Андерсон Т. Статистический анализ временных рядов. - М.: Мир, 1976.
11. Бендат Дж., Пирсол А. Измерение и анализ случайных процессов. - М.: Мир, 1974. - 464 с.
12. Бриллинджер Д. Временные ряды. Обработка данных и теория. - М.: Мир, 1980.
13. Венсель В.В. Интегральная регрессия и корреляция: статистическое моделирование рядов динамики. - М.: Финансы и статистика, 1983.
14. Гренандер У. Случайные процессы и статистические выводы. -М.: ИЛ, 1961. - 168 с.
15. Журбенко И.Г. Спектральный анализ временных рядов. - М.: Изд-во Моск. ун-та, 1982.
16. Журбенко И.Г. Анализ стационарных и однородных случайных систем. - М.: Изд-во Моск. ун-та, 1987. 240 с.
17. Кендалл М.Дж., Стъюарт А. Многомерный статистический анализ и временные ряды. - М.: Наука, 1976.
18. Кендэл М. Временные ряды. - М.: Финансы и статистика, 1981.
19. Крамер Г., Лидбеттер М. Стационарные случайные процессы. - М.: Мир, 1969.
20. Ковалева Л.Н. Многофакторное прогнозирование на основе рядов динамики. - М.: Финансы и статистика, 1980.
21. Отнес Р., Эноксон Л. Прикладной анализ временных рядов. - М.: Мир, 1982.
22. Рабинер Р., Гоулд Б. Теория и применение цифровой обработки сигналов. - М.: Мир, 1978.
23. Статистический анализ экономических временных рядов и прогнозирование. (Серия "Ученые записки по статистике", тт.22-23.) - М.: Наука, 1973.
24. Хеннан Э. Многомерные временные ряды. - М.: Мир, 1974.
25. Цветков Э.И. Основы теории статистических измерений. - Л.: Энергоатомиздат, 1986. - 256 с.
26. Петров В.М., Мажуль Л.А. Цикличность социокультурной сферы и проблемы среднесрочного прогнозирования ее развития. // Математическое и компьютерное моделирование в науках о человеке и обществе. Тезисы докладов Всероссийской конференции. - М.: Госуд. ун-т управления, 1999. - С.63-66.
27. Николаев А.В. Структура исторического цикла. // Математическое и компьютерное моделирование в науках о человеке и обществе. Тезисы докладов Всероссийской конференции. - М.: Госуд. ун-т управления, 1999. - С.54-54.
28. Носовский Г.В., Фоменко А.Т. Введение в новую хронологию. (Какой сейчас век?). - М.: КРАФТ+ЛЕАН, 1999.
29. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979.

30. Каган А.М., Линник Ю.В., Рао С.Р. Характеризационные задачи математической статистики. - М.: Наука, 1972.
31. Орлов А.И. Асимптотика решений экстремальных статистических задач // Анализ нечисловых данных в системных исследованиях. Сб. трудов. Вып.10. - М.: ВНИИСИ, 1982. - С.4-12.
32. Биллингсли П. Сходимость вероятностных мер. - М.: Наука, 1977.
33. Орлов А.И. Метод оценивания длины периода и периодической составляющей сигнала. - В сб.: Статистические методы оценивания и проверки гипотез. Межвузовский сборник научных трудов. - Пермь: Изд-во Пермского государственного университета, 1999. С.38-49.
34. Жихарев В.Н., Кольцов В.Г., Орлов А.И. Новый эконометрический метод "ЖОК" оценки результатов взаимовлияний факторов в инженерном менеджменте . - В сб.: Проблемы технологии, управления и экономики / Под общей редакцией канд. экон. наук. Панкова В.А. Ч.1. Краматорск: Донбасская государственная машиностроительная академия, 1999. - С.87-89.
35. Макконнелл К.Р., Брю С.Л. Экономикс: Принципы, проблемы и политика. В 2 т.: Пер. с англ. 11-го изд. - М.: Республика, 1992.
36. Тейл Г. Эконометрические прогнозы и принятие решений. - М.: Статистика, 1971. - 488 с.
37. Френкель А.А. Математические методы анализа динамики и прогнозирования производительности труда. - М.: Экономика, 1972. - 190 с.
38. Четыркин Е.М. Статистические методы прогнозирования. -М.: Статистика, 1977.
39. Бокс Дж., Дженкинс Г. Анализ временных рядов: Прогноз и управление. - М.: Мир, 1974.
40. Гренджер К., Хатанака М., Спектральный анализ временных рядов в экономике. - - М.: Статистика, 1972.
41. Дженкинс Г., Ваттс Д. Спектральный анализ и его приложения. - М.: Мир, 1971.
42. Маленво Э. Статистические методы в эконометрии. Вып.1,2. - М.: Статистика, 1975, 1976.

Глава 7. Эконометрический анализ инфляции

Наиболее часто встречаются людям такие экономические характеристики, как цены на товары и услуги. Как правило, они изменяются с течением времени. Вполне естественно подвергнуть цены на товары и услуги эконометрическому анализу.

Под инфляцией в настоящей главе понимается рост цен (крайне редко бывает, чтобы цены устойчиво падали). При анализе экономических процессов, протяженных во времени, необходимо переходить к сопоставимым ценам. Это невозможно сделать без расчета индекса роста цен, т.е. индекса инфляции. Проблема состоит в том, что цены на разные товары растут с различной скоростью, и необходимо эти скорости усреднять. На примере разработанной в Институте высоких статистических технологий и эконометрики минимальной потребительской корзины продовольственных товаров, составленной на основе физиологических норм потребления, продемонстрируем свойства и алгоритмы расчета индекса инфляции.

7. 1. Определение индекса инфляции

Средняя цена и разброс цен. В конкретном акте купли-продажи цена товара или услуги полностью определена. Однако в современных условиях, когда в большинстве случаев продавец, а иногда и покупатель, могут влиять на цену товара или услуги, эта цена зачастую меняется от одного акта купли-продажи к другому.

Можно выделить несколько вариантов.

1. Конкретный продавец меняет цену в зависимости от конкретного покупателя. Пример: индивидуальный продавец на базаре.
2. В конкретном магазине цена фиксирована, но от магазина к магазину она меняется. Примеры: большинство товаров (продаваемых в магазинах и киосках), цены которых указаны для сведения покупателей.
3. Единые цены в регионе, например, на услуги почтовой связи.

Пусть в определенный день (момент времени) t осуществлено k актов купли-продажи определенного товара, причем в акте с номером j, где $j=1,2,...,k$, куплено количество товара S_j по цене ϵ_j. Тогда экономическая ситуация описывается двумя векторами размерности k, а именно, вектором объемов продаж:

$$(S_1,S_2,...,S_j,...,S_k),$$

и вектором соответствующих цен:

$$(\epsilon_1, \epsilon_2,..., \epsilon_j,..., \epsilon_k).$$

С целью проведения дальнейшего экономического анализа необходимо перейти к обобщенным характеристикам - к объему продаж-покупок:

$$q = \sum_{j=1}^{k} S_j$$

и средней цене за единицу товара:

$$p = \frac{\sum_{j=1}^{k} e_j S_j}{q}.$$

Разумеется, при переходе от 2k чисел к двум числам, т.е. при "сжатии информации", часть информации теряется. Подобный переход допустим, если все цены c_j сравнительно мало отличаются от средней цены p. В России 1993-1995 гг. с ежемесячным ростом цен не менее 5-10-20 % экономически значим именно общий рост цен. В ситуациях со стабильной средней ценой внимание экономистов привлекает разброс цен относительно средней p. Например, в современной Франции цены на определенный товар в фешенебельных центральных магазинах и в окраинных непрестижных супермаркетах могут отличаться в несколько раз. Это - одно из проявлений так называемой "ценовой диверсификации".

В Москве даже в период общей сильной инфляции также наблюдалось это явление, хотя и в гораздо меньших масштабах (табл.1).

Таблица 1

Цены на севере и в центре Москвы в апреле 1994 г.

Наименование	Север - цена в руб.	Центр- цена в руб.
Хлеб пшеничный	370	390
Огурцы	1800	2000
Яблоки	2000	2500
Молоко	570	615

Измерения цен и объемов продаж относится к определенному моменту времени t, поэтому q и p являются функциями от t, т.е. $q = q(t)$, $p = p(t)$.

При этом время t принимает дискретные значения. Организации, ведущие экономический мониторинг, с 1991 г. измеряют цены с интервалом в неделю.

Функция от дискретного (с постоянным шагом) временного параметра называется временным рядом. Анализом и прогнозированием временных рядов занимается специальный раздел эконометрики - статистика временных рядов (см. главу 6).

Потребительская корзина и доход. Рассмотрим конкретного покупателя, т.е. конкретного экономического субъекта: физическое лицо, домохозяйство или фирму. Он покупает не один товар, а много. Обозначим через n количество типов товаров или услуг (далее кратко - товаров), которые он хочет и может купить. Обозначим через

$$Q_i = Q_i(t),\ i=1,2,...,n,$$

объемы покупок этих товаров по соответствующим ценам:

$$r_i = r_i(t),\ i=1,2,...,n$$

(имеется в виду цена за единицу измерения соответствующего товара - штуку или килограмм...).

Обратим внимание, что выше индекс j относится к различным актам купли-продажи, в которых участвуют различные продавцы и покупатели. В последних же рассмотрениях речь идет о покупках одного экономического субъекта, а потому используется иной индекс суммирования - i.

Расходы на покупки рассматриваемого экономического субъекта равны

$$C = C(t) = \sum_{i=1}^{n} r_i(t) Q_i(t).$$

Эту величину следует сопоставлять с доходом D рассматриваемого экономического субъекта. Если $D - C > 0$, то экономическое положение субъекта благоприятно, его доход больше расходов. Если же $D - C < 0$, то его положение неблагоприятно, доход меньше расходов. Это означает, что он расходует ранее накопленные средства, делает долги (в частности, берет кредиты) и т.д.

Из сказанного вытекает, что величина расходов $C(t)$ регулируется экономическим субъектом в соответствии с его доходом. Изменение (рост) цен $r_i(t)$ с течением времени t делает невозможным сохранение прежней структуры потребления $(Q_1(t), Q_2(t), ..., Q_n(t))$, если рост дохода отстает от роста цен. Структура потребления изменяется, сокращается потребление относительно дорогих товаров и услуг, в порядке компенсации увеличивается потребление относительно дешевых. Например, уменьшается потребление мяса и увеличивается - хлеба и картофеля. При быстром росте цен возможен и другой эффект - "бегство от рубля". В связи с обесцениванием сбережений экономические субъекты направляют доход на текущее потребление, ценой отказа от накопления средств на приобретение дорогостоящих товаров длительного пользования.

Проблемы усреднения темпов роста цен. Обсудим проблемы измерения роста цен на товары и услуги. Рассмотрим сначала конкретного экономического субъекта. Ясно, что цены на различные товары меняются по-разному. Для конкретного товара рост цены описывается величиной (индексом)

$$I_i(t_1, t_2) = r_i(t_2)/r_i(t_1).$$

Эти индексы различны для различных товаров.

Одна из основных проблем в современной экономике - проблема агрегирования с целью сжатия информации (см., например, монографию [1]). Как свести к одной величине индексы цен для различных товаров и услуг?

Уровень цен выражается в виде индекса. Он является измерителем соотношения между совокупной ценой определенного набора товаров, называемого "рыночной корзиной" (или "потребительской корзиной"), для данного (текущего) момента времени, и совокупной ценой идентичной либо сходной группы товаров в базовый момент времени.

Первое, что приходит в голову - усреднить индексы для отдельных товаров и услуг. Но какое среднее взять? Среднее арифметическое? Среднее геометрическое? Среднее гармоническое? Среднее квадратическое? В экономике используется много различных видов средних (см., например, главу 3 выше).

На первый взгляд, представляется естественным использовать взвешенное среднее арифметическое индексов роста цен на отдельные товары и услуги, а в качестве весов использовать относительные объемы потребления этих товаров и услуг. А именно, средним (или индексом) роста цен за интервал времени $[t_1, t_2]$ представляется естественным назвать величину

$$I(t_1, t_2) = \frac{1}{C}\sum_{i=1}^{n} I_i(t_1, t_2) Q_i(t_1) = \frac{1}{C}\sum_{i=1}^{n} Q_i(t_1)\frac{r_i(t_2)}{r_i(t_1)}, \qquad (1)$$

где $C = Q_1(t_1) + Q_2(t_1) + Q_3(t_1) + ... + Q_n(t_1)$.

Обратите внимание на то, что индекс роста цен зависит от базовых значений цен и объемов потребления, определяемых в момент начала измерения t_1, и от текущих цен, измеренных в момент t_2.

Очевидно, при реальных расчетах в качестве базового момента целесообразно брать тот, в котором соотношение различных товаров и услуг в потреблении является типичным для рассматриваемого экономического субъекта, т.е. базовый момент должен отражать стабильное потребление. Например, нельзя в качестве базового момента брать январь 1992 г., когда были отпущены цены на основные продукты и товары народного потребления. В январе 1992 г. структура потребления была сильно искажена, поскольку существенная часть населения, испытывая недостаток наличных денег, резко сократила закупки дорогостоящих товаров и услуг, в частности, используя накопленные запасы продуктов.

Пример 1. Пусть n = 2; $r_1(t_1) = 1{,}0$; $r_2(t_1) = 10{,}0$; $Q_1(t_1) = 0{,}9$; $Q_2(t_1) = 0{,}1$; $r_1(t_2) = 3{,}0$; $r_2(t_2) = 15{,}0$.

Тогда индекс роста цен согласно формуле (1) равен

$$3{,}0/1{,}0 *0{,}9 + 15{,}0/10{,}0*0{,}1 = 2{,}7 + 0{,}15 = 2{,}85.$$

Введенный индекс роста цен меняется от одного экономического субъекта к другому. Он зависит как от структуры потребления конкретного субъекта, так и от цен, по которым тот покупает товары и услуги.

Например, резкий рост цен на сигареты отразится на индексе курящих, но не на индексе остальных физических лиц.

Цены могут зависеть от региона, а также от категории физических или юридических лиц. Так, в городе Иваново в июле 1993 г. сахарный песок отпускался по карточкам по цене 330 руб. за 1 кг. при рыночной цене 600 руб., а поллитровая бутылка водки - по цене 480 руб. при рыночной цене 800 руб.

Стоимостные весовые коэффициенты В чем очевидный экономический недостаток приведенной выше формулы (1) для индекса цен? В ней используется величина *C* - сумма объемов потребления $Q_i(t)$. При попытке ее расчета возникает необходимость складывать объемы потребления, выраженные в физических единицах измерения, например, килограммы картофеля складывать с буханками хлеба, бутылками молока, пачками сигарет и штуками холодильников. Тот читатель, кто этого не заметил, не справился с задачей, поставленной автором настоящей книги, на проверку соответствия экономико-математической модели здравому экономическому смыслу.

Ясно поэтому, что целесообразно измерять потребление не в физических единицах, а в стоимостных. Пусть

$$W_i(t) = \frac{r_i(t)Q_i(t)}{C} = \frac{r_i(t)Q_i(t)}{\sum\limits_{1\le k\le n} r_k(t)Q_k(t)} -$$

- доля потребления *i*-го продукта или услуги в общем потреблении *C* (в стоимостном выражении).

Определение 1. *Индексом инфляции называется*

$$I_{(1)}(t_1,t_2) = \sum_{1\le i\le n} I_i(t_1,t_2)W_i(t_1).$$

Пример 2. В условиях примера 1 объем потребления $C = C(t_1) = 1{,}0*0{,}9 + 10{,}0*0{,}1 = 1{,}9$, доля (в стоимостном выражении) первого продукта в потреблении равна 0,9/1,9, доля второго (в стоимостном выражении) - 1,0/1,9, индекс роста цен

$$I_{(1)}(t_1,t_2) = (3{,}0/1{,}0)*(0{,}9/1{,}9) + (15{,}0/10{,}0)*(1{,}0/1{,}9) = 2{,}21.$$

Результаты расчетов в примерах 1 и 2 различны. В этом нет ничего удивительного, поскольку используются разные весовые коэффициенты при вычислении взвешенного среднего. В этих примерах максимальный индекс роста цен для отдельных товаров есть $I_1(t_1,t_2) = 3{,}0$, минимальный $I_2(t_1,t_2) = 1{,}5$, а усредненные индексы $I_{(1)}(t_1,t_2)$ и $I_{(2)}(t_1,t_2)$ располагаются внутри интервала [1,5; 3,0]. Нетрудно показать, что любое число в этом интервале можно получить в качестве индекса роста цен, соответствующим образом подобрав объемы потребления.

Сравнение стоимостей потребительских корзин. Другой подход к измерению роста цен основан на сравнении стоимостей потребительской корзины $(Q_1(t), Q_2(t), ..., Q_n(t))$ в старых $r_i(t_1)$, $i=1,2,...,n$, и новых $r_i(t_2)$, $i=1,2,...,n$, ценах.

Определение 2. *Индексом инфляции называется*

$$I_{(2)}(t_1,t_2)=\frac{\sum_{1\le i\le n} r_i(t_2)Q_i(t_1)}{\sum_{1\le i\le n} r_i(t_1)Q_i(t_1)}.$$

Пример 3. В условиях примера 1 стоимость потребительской корзины в старых ценах равна 1,0*0,9 + 10,0*0,1 = 1,9, в новых ценах ее стоимость равна 3,0*0,9 + 15,0*0,1 = 4,2, и индекс роста цен (инфляции) равен

$$I_{(2)}(t_1,t_2) = 4{,}2/1{,}9 = 2{,}21.$$

Эквивалентность двух подходов. В примерах 2 и 3 получился один и тот же результат. Случайно ли это? Оказывается, нет.

Теорема 1. Индексы инфляции, введенные согласно определениям 1 и 2, совпадают:

$$I_{(1)}(t_1,t_2) = I_{(2)}(t_1,t_2).$$

Доказательство дается следующей последовательностью преобразований:

$$I_{(1)}(t_1,t_2)=\sum_{1\le i\le n} I_i(t_1,t_2)W_i(t_1)=\sum_{1\le i\le n}\frac{r_i(t_2)}{r_i(t_1)}\times\frac{r_i(t_1)Q_i(t_1)}{C(t_1)}=\sum_{1\le i\le n}\frac{r_i(t_2)Q_i(t_1)}{C(t_1)}=$$

$$=\frac{1}{C(t_1)}\sum_{1\le i\le n} r_i(t_2)Q_i(t_1)=\frac{\sum_{1\le ii\le n} r_i(t_2)Q_i(t_1)}{\sum_{1\le ii\le n} r_i(t_1)Q_i(t_1)}=I_{(2)}(t_1,t_2).$$

Обозначим

$$I(t_1,t_2)=I_{(1)}(t_1,t_2)=I_{(2)}(t_1,t_2).$$

Эта величина называется индексом инфляции или индексом роста цен. При этом t_1 называется начальным, или базовым моментом времени, а t_2 - текущим моментом времени.

Теорема 1 допускает следующую полезную переформулировку (интерпретацию).

Теорема 2 (теорема сложения). Индекс инфляции $I(t_1,t_2)$ является средним арифметическим взвешенным индексов $I_i(t_1,t_2)$ для отдельных товаров (услуг) или товарных групп:

$$I(t_1,t_2)=\sum_{1\le i\le n}\beta_i I_i(t_1,t_2),$$

причем весовые коэффициенты β_i положительны и в сумме составляют 1. При этом β_i - это доля стоимости потребительской корзины, приходящаяся на соответствующий товар (услугу) в начальный (базовый) момент времени.

Для доказательства достаточно заметить, что $\beta_i = W_i(t_1), i=1,2,...,n.$

7.2. Практически используемые потребительские корзины и соответствующие индексы инфляции

Выше обсуждался индекс инфляции для отдельного экономического субъекта. Таким образом, каждый человек, семья (домохозяйство), фирма может без больших трудозатрат оценить влияние роста цен на свое экономическое положение.

Однако обычно индекс инфляции рассматривают для более или менее обширной совокупности экономических субъектов - для жителей региона или страны, предприятий определенной отрасли и т.д. В таких случаях $Q_i(t)$ заменяют на общий объем потребления $q_i(t)$, а $r_i(t)$ - на среднюю цену $p_i(t)$.

В этих обозначениях ***индекс инфляции*** имеет вид:

$$I(t_1, t_2) = \frac{\sum\limits_{1 \le i \le n} p_i(t_2) q_i(t_1)}{\sum\limits_{1 \le i \le n} p_i(t_1) q_i(t_1)}.$$

Таким образом, индекс инфляции определяется номенклатурой (т.е. набором, перечнем) товаров и услуг, для которых он вычисляется, объемами потребления и ценами этих товаров и услуг на начальный момент времени и ценами на текущий момент времени.

Из принятого определения вытекает, в частности, что индекс инфляции для продовольственных товаров отличается, вообще говоря, от такового для промышленных товаров, для услуг и от индекса инфляции оптовых цен, индекс инфляции для москвича отличается от такового для жителя Краснодара, индекс инфляции для машиностроительной продукции меняется от индекса инфляции в строительстве, этот индекс меняется в зависимости от индивидуальных структур потребностей в семьях и т.д. Для точного определения индекса инфляции необходимо знать объемы купли-продажи и цены в каждом акте купли-продажи, иначе можно говорить только о той или иной оценке этого индекса.

Конкретизация задачи вычисления индекса инфляции. Прежде всего необходимо сформулировать цель экономического анализа роста цен. Будем ориентироваться на положение основной массы населения. Это означает, в частности, что рост цен на персональные компьютеры (в 2001 г. их имеет 6% российских семей) и автомашины иностранных марок (в 2001 г. их имеет 0,5% российских семей) нас не интересует.

Количество видов товаров и услуг измеряется тысячами (а в классификаторах промышленной продукции указаны миллионы марок различных товаров). Поэтому первый шаг - ограничение номенклатуры товаров и услуг, используемых для вычисления индекса инфляции.

В настоящее время существенная часть доходов населения (зачастую не менее половины) идет на покупку продовольственных товаров (что по классическому закону Энгеля - см., например, учебник нобелевского лауреата по экономике П.Самуэльсона [2] - свидетельствует о сравнительно низком жизненном уровне). Поэтому представляется естественным рассчитать индекс инфляции для продовольственных товаров. Методике такого расчета и посвящена оставшаяся часть настоящего пункта главы 7.

Потребительская корзина Центра экономической конъюнктуры. Центр экономической конъюнктуры при Правительстве Российской Федерации и Государственный комитет Российской Федерации по статистике следят за движением цен по фиксированному набору товаров, которые относительно постоянно бывают в магазинах (по различным причинам время от времени этот набор меняется). Это такие товары (в скобках указаны объемы годового потребления в килограммах; причем, если не оговорено противное, первое число соответствует потребительской корзине Центра экономической конъюнктуры по состоянию на 1993 г., второе - потребительской корзине Института высоких статистических технологий и эконометрики, о которой речь пойдет ниже):

1. Хлеб ржаной (92,0/65,3);

2. Хлеб пшеничный (86,7/59,8);
3. Пшено (18,1/4,9);
4. Вермишель (7,3/4,9);
5. Сахар (24,8/19,0);
6. Масло растительное (10,0/3,8);
7. Масло животное (3,6/2,5);
8. Говядина (42,0/4,4);
9. Колбаса вареная (2,2/0,7);
10. Колбаса полукопченая (1,1/0,7);
11. Молоко (184,3/110,0);
12. Сметана (4,2/1,6);
13. Сыр твердый (2,0/2,3);
14. Яйца (183 шт./152 шт.);
15. Картофель (146,0/124,2);
16. Свежая капуста (29,8/30,4);
17. Репчатый лук (10,2/27,9);
18. Яблоки (11,0/15,1);
19. Сигареты (96 пачек/-).

Потребительская корзина Института высоких статистических технологий и эконометрики на основе данных Института питания РАМН. Однако приведенный выше набор не полностью соответствует перечню продуктов питания, рекомендованному медиками. И дело не только в сигаретах.

Рассмотрим минимальную потребительскую корзину физиологически необходимых продовольственных товаров, разработанную Институтом высоких статистических технологий и эконометрики (ИВСТЭ) на основе исходных данных Института питания Российской академии медицинских наук (РАМН). Эти данные используются также Министерством труда Российской Федерации. Эту минимальную потребительскую корзину обозначим сокращенно "корзина ИВСТЭ". В отличие от приведенной выше корзины Центра экономической конъюнктуры в ней содержание белков, жиров и углеводам соответствует (минимальным) медицинским нормам. В корзине ИВСТЭ продукты питания разделены на 11 групп:

1. Хлеб и хлебопродукты ;
2. Картофель ;
3. Овощи ;
4. Фрукты и ягоды ;
5. Сахар ;
6. Мясопродукты ;
7. Рыба и рыбопродукты ;
8. Молоко и молочные продукты ;
9. Яйца ;
10. Масло растительное и маргарин ;
11. Прочие.

Общая стоимость "прочих" видов продуктов - до 6 % от стоимости первых 10 групп продуктов данной потребительской корзины.

На основе физиологических норм потребления Института питания РАМН в ИИСТЭ составлена минимальная потребительская корзина, т.е. указан годовой объем потребления по основным продовольственным товарам, необходимый для поддержания нормальной жизнедеятельности человеческого организма.

Расчет стоимости минимальной потребительской корзины продовольственных товаров. Чтобы получить индекс инфляции, рассчитаем стоимость минимальной потребительской корзины продовольственных товаров, исходя из объемов потребления, заданных в разработках Института питания, и цен по состоянию на март 1991 г. (т.е. до первого значительного повышения цен в апреле 1991 г. и их "либерализации" в январе 1992 г.) и - в качестве примера - на март 1994 г. (очевидно, расчеты могут быть проведены и на любой иной момент времени) с целью установить динамику цен за полные три года.

Исходные данные для расчета приведены в табл.2. Мы видим, что темпы роста цен на различные продукты питания существенно отличаются друг от друга. Минимальный рост цен - в 633 раза (яблоки сушеные), максимальный - в 5946 раз (минтай).

Табл. 2. Номенклатура, годовые нормы потребления и цены для потребительской корзины ИВСТЭ на основе данных Института питания РАМН

Наименование продукта питания	Годовая норма, кг	Цена на 14.03.91	Цена на 14.03.94	Рост цен, раз
1. Хлеб и хлебопродукты				
1.1 Мука пшеничная	18,5	0-46	646	1404
1.2 Рис	3,5	0-88	620	705
1.3 Другие крупы	4,9	0-62	750	1210
1.4 Хлеб пшеничный	59,8	0-50	720	1440
1.5 Хлеб ржаной	65,3	0-20	390	1950
1.6 Макаронные изделия	4,9	0-70	1200	1714
2. Картофель	124,22.	0-10	490	4900
3. Овощи				
3.1 Капуста	30,4	0-20	500	2500
3.2 Огурцы и помидоры	2,8	0-85	2500	2941
3.3 Столовые корнеплоды	40,6	0-20	450	2250
3.4 Прочие (лук и др.)	27,9	0-50	900	1800
4. Фрукты и ягоды				
4.1 Яблоки свежие	15,1	1-50	960	640
4.2 Яблоки сушены	1,0	3-00	1900	633
5. Сахар и кондитерские изделия				
5.1 Сахар	19,0	0-90	650	722
5.2 Конфеты	0,8	4-50	3500	778
5.3 Печенье и торты	1,2	1-40	14700	3357
6. Мясо и мясопродукты				
6.1 Говядина	4,4	2-00	42700	1350
6.2 Баранина	0,8	1-80	1940	1078
6.3 Свинина	1,4	2-00	2300	1150
6.4 Субпродукты (печень)	0,5	1-40	3500	2500
6.5 Птица	16,1	2-40	2600	1083
6.6 Сало	0,7	2-40	3300	1375
6.7 Копчености	0,7	3-70	15000	4054
7. Рыба и рыбопродукты				
7.1 Свежая (минтай)	10,9	0-37	2200	5946
7.2 Сельди	0,8	1-40	2500	1786

Наименование продукта питания	Годовая норма, кг	Цена на 14.03.91	Цена на 14.03.94	Рост цен, раз
8. Молоко и молочные продукты				
8.1 Молоко, кефир	110,0	0-32	520	1625
8.2 Сметана, сливки	1,6	1-70	2500	1471
8.3 Масло животное	2,5	3-60	4000	1111
8.4 Творог	9,8	1-00	2000	2000
8.5 Сыр и брынза	2,3	3-60	6000	1667
9. Яйца, шт.	152,0	10-09	100	1111
10. Масло растительное, маргарин				
10.1 Масло растительное	3,8	1-80	2000	1111
10.2 Маргарин	6,3	1-20	2000	1667

Примечание. Пункт 1.3 - геркулес (в этой таблице и далее).

Для нахождения расходов на определенные продукты питания (в расчете на год) достаточно умножить цену на объем потребления, как это сделано в табл.3. Там же приведены годовые расходы для каждой из 11 товарных групп.

Табл..3. Годовые расходы на покупку продуктов.

Наименование продукта питания	Годовые расходы по ценам на 14.03.91	Годовые расходы по ценам на 14.03.94
1. Хлеб и хлебопродукты		
1.1 Мука пшеничная	8-51	11951
1.2 Рис	3-08	2170
1.3 Прочие крупы	3-04	3675
1.4 Хлеб пшеничный	29-90	43056
1.5 Хлеб ржаной	13-06	25467
1.6 Макаронные изделия	3-43	5880
Всего:	61-02	92199
2. Картофель	12-42	60858
3. Овощи		
3.1 Капуста	6-08	15200
3.2 Огурцы и помидоры	2-38	7000
3.3 Столовые корнеплоды	8-12	18270
3.4 Прочие (лук и др.)	13-95	25110
Всего:	30-53	65580
4. Фрукты и ягоды		
4.1 Яблоки свежие	22-65	14496
4.2 Яблоки сушеные	3-00	1900
Всего:	25-65	16396
5. Сахар и конд. изделия		
5.1 Сахар	17-10	12350
5.2 Конфеты	3-60	2800
5.3 Печенье и торты	1-68	5640
Всего:	22-38	20790
6. Мясо и мясопродукты		
6.1 Говядина	8-80	11880
6.2 Баранина	1-44	1552
6.3 Свинина	2-80	3220

Наименование продукта питания	Годовые расходы по ценам на 14.03.91	Годовые расходы по ценам на 14.03.94
6.4 Субпродукты (печень) 6.5 Птица 6.6 Сало 6.7 Копчености Всего:	0-70 38-64 1-68 2-59 56-65	1750 41860 2310 10500 73072
7. Рыба и рыбопродукты 7.1 Свежая (минтай) 7.2 Сельди Всего:	 4-03 1-12 5-15	 23980 2000 25980
8. Молоко и молочные продукты 8.1 Молоко, кефир 8.2 Сметана, сливки 8.3 Масло животное 8.4 Творог 8.5 Сыр и брынза Всего:	 35-20 2-72 9-00 9-80 8-28 65-00	 57200 4000 10000 19600 13800 104600
9. Яйца	13-68	15200
10. Масло растительное, маргарин 10.1 Масло растительное 10.2 Маргарин Всего:	 6-84 7-56 14-40	 7600 12600 20200
Всего по 10 группам:	306-89	490675
11. Прочие (6% от групп 1-10)	18-41	29441
Суммарная стоимость минимальной потребительской корзины продуктов питания в расчете на год	325-30	520116
Ее стоимость в расчете на месяц	27-11	43499

7.3. Свойства индексов инфляции

Таблицы типа приведенных выше табл.2 и 3 могут быть составлены любым студентом или иным заинтересованным гражданином, любой профсоюзной организацией, любым менеджером, любой фирмой с целью изучения динамики своего реального экономического положения.

Значения индексов инфляции. Как вытекает из табл.3, индекс инфляции (роста цен) по продуктам питания исходя из минимальной потребительской корзины ИВСТЭ, составленной по физиологическим нормам потребления продуктов питания для города Москвы (согласно разработкам Института питания РАМН и Министерства труда РФ), за три года (14.03.91 - 14.03.94) составил

(520116-00) / (325-30)= 1598,88 или 159788 %

7 Есть два основных способа записи индекса инфляции - в "разах" и в "процентах". В "разах" - это именно тот способ, что использовался до настоящего момента. Однако часто мы слышим или читаем: инфляция за месяц составила 5%. Что это значит? Имеется в виду, что индекс инфляции за месяц составил 1,05. Следовательно, связь между двумя "ликами" индекса инфляции задается соотношениями

$$a\% = 100(I-1)\%, \quad I = 1 + \frac{a\%}{100}.$$

Итак, чтобы перейти к выражению индекса инфляции в процентах, надо значение "в разах" уменьшить на 1 и результат умножить на 100, как это и сделано выше.

Среднемесячная инфляция, как и средний темп роста для любого временного ряда, рассчитывается в предположении, что ежемесячный рост цен не меняется от месяца к месяцу. Она равна

$$(1598{,}88)^{1/36} = 1{,}227424776 \text{ или } 22{,}74\ \%.$$

Другими словами, цены каждый месяц увеличивались в среднем на 22,74% .

Приведем некоторые данные о стоимости потребительской корзины ИВСТЭ и индексах инфляции в различных административных округах Москвы и подмосковных городах (табл.4).

Табл. 4. Стоимости потребительской корзины и индексы инфляции за различные периоды в течение 1991-1994 гг.

<table>
<tr><td>Дата</td><td colspan="3">Стоимости потребительской корзины ИВСТЭ</td><td colspan="3">Индексы инфляции</td></tr>
<tr><td>14.03.91
14.03.94</td><td colspan="3">26,85
40889</td><td colspan="3">1522,9</td></tr>
<tr><td>14.08.93
14.03.94</td><td colspan="3">17691
40889</td><td colspan="3">2,31</td></tr>
<tr><td>15.11.93
14.03.94</td><td colspan="3">28050
40889</td><td colspan="3">1,46</td></tr>
<tr><td></td><td>Север</td><td>Юго-Запад</td><td>Юг</td><td>Север</td><td>Юго-Запад</td><td>Юг</td></tr>
<tr><td>31.03.94</td><td>44933</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>07.04.94</td><td>45196</td><td></td><td></td><td>1,01</td><td></td><td></td></tr>
<tr><td>14.04.94</td><td>48148</td><td>44441</td><td></td><td>1,05</td><td></td><td></td></tr>
<tr><td>21.04.94</td><td>49592</td><td>45306</td><td></td><td>1,02</td><td>1,02</td><td></td></tr>
<tr><td>28.04.94</td><td>51576</td><td>47284</td><td>53941</td><td>1,03</td><td>1,04</td><td></td></tr>
<tr><td>05.05.94</td><td>53123</td><td>47413</td><td>57337</td><td>1,02</td><td>1,01</td><td>1,05</td></tr>
<tr><td>12.05.94</td><td>55295</td><td>47773</td><td>55070</td><td>1,03</td><td>1,01</td><td>0,96</td></tr>
<tr><td>19.05.94</td><td>55615</td><td></td><td>57167</td><td>1,01</td><td></td><td>1,05</td></tr>
<tr><td>26.05.94</td><td>56332</td><td></td><td>61200</td><td>1,01</td><td></td><td>1,06</td></tr>
<tr><td>14.04.94
12.05.94</td><td colspan="3">46662,44
54326,29</td><td colspan="3">1,16</td></tr>
<tr><td></td><td colspan="3">г. Ногинск</td><td colspan="3">г. Ногинск</td></tr>
<tr><td>02.03.94
09.04.94
05.05.94</td><td colspan="3">34785
41141
47091</td><td colspan="3">1,18
1,15</td></tr>
<tr><td></td><td colspan="3">г. Электрогорск</td><td colspan="3">г. Электрогорск</td></tr>
<tr><td>17.04.94
24.04.94
30.04.94
05.05.94</td><td colspan="3">42330
42337
42412
43555</td><td colspan="3">1,01
1,01
1,03</td></tr>
</table>

Замечание. В табл.4 и дальнейших таблицах стоимости потребительской корзины ИВСТЭ приведены по 10 группам (т.е. без умножения на 1.06).

По данным табл.4 можно рассчитать индексы и за иные интервалы времени. Так, например, поскольку на 15.08.93 стоимость потребительской корзины равна 17691 руб., то индекс инфляции за период 15.03.91 - 15.08.93 равен I(15.03.91; 15.08.93)= S(15.08.93)/S(15.03.91)= 17691/26,85 = 659.

Сопоставим инфляцию со средней заработной платой. В марте 1991 г. она равнялась приблизительно 300 руб. в месяц, т.е. минимальная продуктовая корзина ИВСТЭ составляла около 8,9 % от средней заработной платы. В марте 1994 г. среднемесячная зарплата в Москве составила 206076 руб. (данные Московского городского статистического комитета), следовательно, стоимость корзины ИВСТЭ составляла 4349900 / 206076 %=21,11% от нее. Если судить по ценам на продукты, за три года уровень жизни основной массы населения понизился в 21,1/11,4 = 1,85 раз.

При этом надо иметь в виду, что типичное распределение доходов таково, что мода величин доходов весьма меньше медианы, которая в свою очередь существенно меньше среднего арифметического (центральная часть распределения доходов - за исключением больших доходов - хорошо приближается логарифмически нормальным распределением). Дифференциация доходов в России быстро нарастала вплоть до второй половины 1990-х годов и сильно превзошла уровень всех т.н. промышленно развитых стран. Правда, уровень Бразилии и Кении пока не достигнут, но и климат в этих странах существенно иной, так что минимальное жизнеобеспечение требует существенно меньших затрат.

В октябре 1995 г. в Москве средняя заработная плата - 520 тыс. руб., а стоимость потребительской корзины ИВСТЭ - 196,6 тыс. руб., т.е. 37,8% от средней зарплаты, падение уровня жизни - в 4,2 раза.

Нет необходимости связывать индекс инфляции с каким-либо определенным интервалом времени и даже с определенным социально-экономическим строем. Можно формально вычислить индексы инфляции и за весьма длительные промежутки времени. Так, например, рост цен на основные продукты питания с 1913 г. по апрель 1994 г. представлен в табл.5.

Табл. 5. Цены в 1913 г. и в апреле 1994 г.

Наименование продукта	Цена в 1913 г. (руб/ кг)	Цена в апр. 1994 г. (руб/кг)
Хлеб пшеничный	0-05	740
Хлеб ржаной	0-03	400
Молоко	0-14	625
Сыр	0-40	6150
Масло сливочное	0-55	5100
Масло растительное.	0-13	2300
Сметана	0-30	2500
Говядина	0-23	2760
Свинина	0-20	4000
Баранина	0-17	2000

Используя объемы потребления из потребительской корзины ИВСТЭ, получаем, что индекс инфляции за 1913-1994 гг. составил 1129600,6 %.

Результаты расчетов по различным потребительским корзинам дают, естественно, различные значения индексов инфляции, хотя эти различия, как представляется, не слишком значительны (табл.6). Близость различных индексов инфляции за большой промежуток времени объясняется тем, что цены растут в целом достаточно согласованно, "аномалии" выправляются: если темп роста цены определенного продукта отстает от среднего роста цен, то имеются основания полагать, что его цена в ближайшее время сильно возрастет. Однако на малых и средних промежутках времени проявляется различие роста цен на отдельные товары.

Тем более интересно, что официально публикуемые индексы инфляции Госкомстата РФ при отсчете с 1990 г. (или с 14.03.91) дают по крайней мере вдвое меньшие значения, чем расчеты Института высоких статистических технологий и эконометрики (подробнее см. коллективную монографию [3]).

Табл. 6. Сравнение результатов подсчета индексов инфляции и стоимостей потребительских корзин по нормам Госкомстата РФ и ИВСТЭ

	По нормам Госкомстата РФ	По нормам ИВСТЭ
с 14.03.91 по 14.03.94	30,82 / 48990,33 1589,8	26,85 / 40889,1 1598,88
с 15.11.93 по 14.03.94	31255 / 48990,33 1,57	28050 / 40889,1 1,46
с 19.05.94 по 26.05.94	56670,2 / 57667,75 1,02	55615 / 56332 1,01

Примечание. *В табл.6 верхние числа - стоимости потребительских корзин соответственно на первую указанную дату и через дробь - на вторую, нижнее число - индекс инфляции за данный период.*

Соотношение индексов инфляции для трех моментов времени. Рассмотрим три момента времени t_1, t_2, t_3 и соответствующие индексы инфляции $I(t_1,t_2)$, $I(t_2,t_3)$ и $I(t_1,t_3)$. Из определения индекса инфляции как отношения стоимостей потребительской корзины в соответствующие моменты времени вытекает следующее утверждение.

Теорема 3 (теорема умножения). Для любых трех моментов времени t_1, t_2, t_3 справедливо равенство

$$I(t_1,t_3) = I(t_1,t_2)I(t_2,t_3)$$

Теорема умножения позволяет переходить от индексов инфляции за отдельные недели к индексам инфляции за месяц (четыре недели), от помесячных индексов инфляции - к квартальным и годовым, от годовых - к индексам инфляции за несколько лет. Например, индекс инфляции за второй квартал - с 01.04.94 по 01.07.94 - т.е. I(01.04.94, 01.07.94), выражается через индексы инфляции за апрель I(01.04.94, 01.05.94), май I(01.05.94, 01.06.94) и июнь I(01.06.94, 01.07.94) соответственно как произведение этих индексов, т.е. находится по формуле

I(01.04.94, 01.07.94) = I(01.04.94,01.05.94) I(01.05.94,01.06.94) I(01.06.94,01.07.94).

Аналогично индекс инфляции за год равен произведению двенадцати индексов инфляции: за январь, февраль, март и остальные девять месяцев.

В приведенных выше рассуждениях индекс инфляции рассматривался как положительное число, как говорят, выражался "в разах". Распространено его выражение "в процентах". Напомним, что индексы инфляции "в разах" $I(t_1,t_2)$ и "в процентах" $i(t_1,t_2)$ связаны соотношениями

$i(t_1,t_2) = (I(t_1,t_2) - 1)\ 100\%$,

$I(t_1,t_2) = 1 + i(t_1,t_2)\ / 100$.

Таким образом, выражения "индекс инфляции за месяц составил 1,16" и "индекс инфляции за месяц составил 16%" означают одно и то же.

Поскольку для любых чисел x и y справедливо тождество

$$(1 + x)(1 + y) = 1 + x + y + xy,$$

то, как легко проверить, для индексов инфляции "в процентах" справедливо тождество

$$i(t_1, t_3) = i(t_1, t_2) + i(t_2, t_3) + \frac{i(t_1, t_2) i(t_2, t_3)}{100}.$$

Если индексы инфляции "в процентах" $i(t_1, t_2)$ и $i(t_2, t_3)$ малы, т.е. индексы инфляции "в разах" $I(t_1, t_2)$ и $I(t_2, t_3)$ мало отличаются от единицы, то справедлива приближенная формула

$$i(t_1,t_3) = i(t_1,t_2) + i(t_2,t_3).$$

Погрешность этой формулы, измеряемая в процентах, равна

$$\frac{i(t_1, t_2) i(t_2, t_3)}{100}.$$

Эта величина становится заметной, если сомножители - порядка десятков (процентов).Если формула применяется несколько раз, то погрешность накапливается.

Рассмотрим пример. В известном учебнике экономической теории [4] рассмотрена связь между ежегодным увеличением цен и числом лет, необходимых для увеличения цен вдвое. Приведено правило, которое вначале выглядит совершенно непонятным:

(приблизительное количество лет, необходимое для удвоения удвоение цен) = 70 / (темп ежегодного увеличения уровня цен в %).

Действительно, пусть n - количество лет, необходимое для удвоения цен, а x- темп ежегодного увеличения уровня цен (в % - $100x$ %). При "подходе профана" рост за n лет составит $100nx$, а потому срок удвоения цен должен находиться из условия

$$100nx = 100, \quad n = 100 / (100x),$$

т.е. в числителе дроби должно стоять число 100, а не 70. В чем дело?

А дело в том, что рост описывается не линейной функцией, а экспоненциальной, надо не складывать, а возводить в степень. За n лет рост цен составит $(1+x)^n$. Период удвоения находится из уравнения

$$(1+x)^n = 2$$

Тогда

$$n = \frac{\ln 2}{\ln(1+x)}.$$

Воспользуемся приближенной формулой математического анализа

$$\ln (1 + x) = x + O (x^2),$$

тогда с точностью до бесконечно малых более высокого порядка

$$n = 100 \ln 2 \; / \; 100\, x.$$

Остается заметить, что

$$100 \ln 2 = 100 * 0{,}69314718,$$

т.е. с достаточной для подобных расчетов точностью $100 \ln 2 = 70$.

Некоторые ошибки при расчетах с индексами инфляции. Информация об индексах инфляции и рассуждения, связанные с ними, постоянно появляются на страницах печати и обсуждаются в иных средствах массовой информации. К сожалению, достаточно распространены ошибки.

Так, в одной из экономических (!) газет была помещена публикация, в которой основной исходный материал для обсуждения - следующие индексы инфляции (по отношению к предыдущему месяцу):
январь - 1, февраль - 1,23, март - 1,19, апрель - 1,25,
май - 1,29, июнь - 1,3, июль - 1,23, август - 1,22,
сентябрь - 1,22, октябрь - 1,19, ноябрь - 1,23, декабрь - 1,25.

Автору публикации были нужны индексы инфляции за несколько месяцев. Рассчитывая их, он без каких-либо сомнений пользовался приближенной формулой предыдущего подпункта (сложение приращений в процентах) вместо точной (перемножение индексов инфляции, выраженных "в разах"). В результате он получил для периода январь-декабрь значение индекса инфляции 3,60 (23% + 19% + 25% + 29% + 30% + 23% + 22% + 22% + 19% + 23% + 25% = 260%), в то время как на самом деле индекс инфляции, рассчитанный в результате перемножения индексов по месяцам, равен 10,23. Допущенная ошибка в 10,23/3,60 = 2,84 раза существенно исказила дальнейшие расчеты (фонда оплаты труда, средней зарплаты) в рассматриваемой публикации, названной в специализированной экономической газете не как-нибудь, а "консультацией" !

В еженедельнике "Аргументы и факты" в апреле 1994 г. в рубрике "Прогноз" помещена беседа журналистки Татьяны Коростиковой с первым заместителем министра экономики России Яковом Уринсоном [5], в которой Я. Уринсон прогнозирует:

"...Мы предполагаем рост цен за 1994 г. в 5 раз... В месяц- 7-8%..."

Сказанное противоречиво. Если индекс инфляции за год равен 5,0, то за месяц, очевидно, рост цен равен в среднем

$$12\sqrt{5{,}0} = 1{,}1435,$$

т.е. 14,35% в месяц, а не 7-8%. Если же рост цен составляет 7-8% в месяц, то индекс инфляции за год лежит между

$$(1{,}07)^{12} = 2{,}25 \quad \text{и} \quad (1{,}08)^{12} = 2{,}52,$$

т.е. по крайней мере в два раза меньше, чем названный в беседе достаточно реальный прогноз - рост в 5 раз. Остается неясным, кто дезориентировал читателя многотиражного издания - чиновник или журналист. Наш запрос об этом в редакцию "Аргументов и фактов" остался без ответа.

Приведенных примеров достаточно для констатации того, что к сообщениям в средствах массовой информации, посвященным росту цен, следует относиться с известной осторожностью.

Потребительские корзины, включающие в себя промтовары и услуги, и соответствующие индексы инфляции. В настоящее время, чтобы не только быть в курсе проблем, касающихся инфляции в нашей стране, но и хорошо ориентироваться в создавшейся ситуации, недостаточно отслеживать только изменение цен на продовольственные товары. Необходимо также фиксировать инфляцию и в сфере коммунальных, транспортных, медицинских, образовательных и других услуг, а также анализировать цены на промышленные товары широкого потребления. Рост цен в этих областях достаточно заметен (если раньше проезд в метро обходился в 5 коп., то в ноябре 1995 г. он стоил 1000 руб., а в феврале 1999 г. , после деноминации в 1000 раз - 4 рубля, в 2001 г. - разовая поездка обходилась в 5 руб.). Следует также отметить, что темпы роста цен на те или иные промышленные товары и услуги не всегда совпадают с темпами ростом цен на продовольственные товары. Например, наблюдалось подорожание хлебобулочных товаров примерно в 2000 раз за три года (1991-1994), а цены на компьютерные товары выросла за это время в среднем только в 80 раз.

При обсуждении проблем инфляции студенты часто обращают внимание на то, что в настоящее время заметная часть доходов каждой семьи идет на оплату коммунальных услуг и покрытие расходов на транспорт и связь. Необходимо учитывать расходы на услуги прачечной, парикмахерской, на ремонт обуви и т.д. Увеличиваются расходы на удовлетворение культурных потребностей из-за роста цен на книги, журналы, газеты, билеты в театры и кино, спортивный инвентарь и т.д. С течением времени подобные расходы могут и сокращаться из-за прекращения покупок книг, журналов, газет, прекращения походов в театры и т.д.

Дорогими сегодня являются и промышленные товары. Но при подсчете индекса инфляции по этим товарам возникает ряд трудностей. Например, наблюдается разброс цен по магазинам или имеет место временное отсутствие в магазинах некоторых товаров. Кроме того, меняется мода, многие виды одежды выходят из употребления, вместо них появляются новые. То же самое, в связи с развитием техники, происходит и с товарами длительного пользования (когда-то раньше не было телевизоров, холодильников, стиральных машин, железных дорог и самолетов). Кроме того, пока еще мы можем пользоваться бесплатными услугами в области медицины и образования, но скоро, очевидно, и это будет платным, по крайней мере, частично.

Для того, чтобы подсчитать индекс инфляции по достаточно обширной потребительской корзине, включающей не только продовольственные товары, но и одежду, товары длительного пользования, услуги и т.п., необходимо иметь соответствующие нормы потребления. Определить их весьма трудно. (При нормативном подходе к экономическим явлениям - откуда взять нормы? При позитивном - как в нестабильной ситуации замерить потребительские бюджеты?) Поэтому в настоящей главе мы ограничились индексами инфляции, рассчитанными для продовольственной потребительской корзины. Индекс инфляции можно считать не только для Москвы в целом, но и для отдельных ее районов и даже для покупателей отдельных магазинов - достаточно измерить соответствующие цены; не только для населения в целом, но и для отдельных слоев и даже отдельных семей - достаточно знать соответствующие потребительские корзины.

Индекс инфляции целесообразно сопоставлять со средним (средним арифметическим) доходом. Данные Московского городского статистического комитета приведены в табл.7. О реальной динамике благосостояния можно судить по временному ряду, составленному из частных от деления среднего дохода на стоимость потребительской корзины.

Табл. 7

Средняя (средняя арифметическая) зарплата в Москве

1993,	январь	- 15997,6	1993,	сентябрь	- 88621,0
	февраль	- 18966,7		октябрь	- 99621,0
	март	- 23939,7		ноябрь	- 111807,1
	июнь	- 48424,2	1994,	январь	- 177061,0
	июль	- 56568,9		февраль	- 179626,0
	август	- 62986,5		март	- 206076,0

При проведении мониторинга за ценами на продукты питания обычно у нас есть некоторая компьютерная база данных, в которую заносятся сведения вида:

1. НАЗВАНИЕ ПРОДУКТА ПИТАНИЯ
2. ЦЕНА НА ПРОДУКТ
3. ОБЪЕМ ПОТРЕБЛЕНИЯ ПРОДУКТА
4. ДАТА СНЯТИЯ ЦЕНЫ
5. НАЗВАНИЕ ТОРГОВОЙ ТОЧКИ.

Кроме того, есть вполне понятная последовательность действий по вычислению индекса инфляции с момента времени t_1 по момент времени t_2 :

1. Вычислить сумму произведений ЦЕНА*ОБЪЕМ ПОТРЕБЛЕНИЯ для момента времени t_1 ;
2. Вычислить сумму произведений ЦЕНА*ОБЪЕМ ПОТРЕБЛЕНИЯ для момента времени t_2 ;
3. Найти их отношение.

Для нахождения индекса инфляции по группам эти действия выполняются для продуктов искомой группы.

Для вычисления индекса инфляции по продуктам питания разработаны различные программные средства для IBM PC и для персональных компьютеров фирмы "Apple".

Табл.8. Индекс инфляции и стоимость потребительской корзины Института высоких статистических технологий и эконометрики

№ п/п	Дата снятия цен	Стоимость потребитель-ской корзины *S(t)* (руб.)	Индекс инфляции *I(31.3.91;t)*
1	31.3.91	26.60	1.00
2	14.8.93	17,691.00	665.08
3	15.11.93	28,050.00	1054.51
4	14.3.94	40,883.00	1536.95
5	14.4.94	44,441.00	1670.71
6	28.4.94	47,778.00	1796.17
7	26.5.94	52,600.00	1977.44
8	8.9.94	58,614.00	2203.53
9	6.10.94	55,358.00	2081.13
10	10.11.94	72,867.00	2739.36
11	1.12.94	78,955.00	2968.23
12	29.12.94	97,897.00	3680.34
13	2.2.95	129,165.00	4855.83
14	2.3.95	151,375.00	5690.79
15	30.3.95	160,817.00	6045.75
16	27.4.95	159,780.00	6006.77
17	1.6.95	167,590.00	6300.38
18	29.6.95	170,721.00	6418.08
19	27.7.95	175,499.00	6597.71

20	31.8.95	173,676.00	6529.17
22	28.9.95	217,542.00	8178.27
23	26.10.95	243,479.00	9153.35
24	30.11.95	222,417.00	8361.54
25	28.12.95	265,716.00	9989.32
26	1.2.96	287,472.55	10,807.24
27	5.3.96	297,958.00	11,201.43
28	5.4.96	304,033.44	11,429.83
29	8.5.96	305,809.55	11,496.60
30	5.6.96	302,381.69	11,367.73
31	3.7.96	306,065.21	11,506.21
32	3.8.96	308,963.42	11,615.17
33	7.9.96	288,835.07	10,858.46
34	1.10.96	278,235.35	10,459.98
35	5.11.96	287,094.77	10,793.04
36	4.12.96	298,024.76	11,203.94
37	3.1.97	314,287.16	11,815.31
38	4.2.97	334,738.24	12,584.14
39	4.1.98	345.72	12.997
40	3.1.99	622.30	23.395
41	5.1.00	851.32	32.004
42	3.1.01	949.21	35.684
43	2.7.01	1157.23	43.505

Примечание 1. В таблице целая часть отделяется от дробной десятичной точкой, а запятая используется для деления числа по разрядам (на западный манер). Учитывается проведенная деноминация рубля. Если ее не учитывать, то за 10 лет (1991-2001) цены (в Москве) выросли примерно в 40 тысяч раз. Поскольку экономические связи между регионами ослабли, то темпы роста цен в регионах различаются, но, видимо, не более чем в 2 раза.

Примечание 2. Стоимости потребительской корзины ИВСТЭ на один и тот же момент времени, как мог заметить внимательный читатель, несколько отличаются. В этом нет ничего странного, так как исходные цены на продукты несколько отличались. Строго говоря, цены, стоимости потребительских корзин, индексы инфляции и многие другие экономические величины следовало бы считать нечисловыми данными (см. главу 8), например, интервальными или нечеткими.

7.4. Возможности использования индекса инфляции в экономических расчетах

Хорошо известно, что стоимость денежных единиц со временем меняется. Например, на один доллар США полвека назад можно было купить примерно в восемь раз больше материальных ценностей (например, продовольствия), чем сейчас (см. таблицу пересчета в учебнике [4]), а если сравнивать с временами Тома Сойера - в 100 раз больше. Причем стоимость денежных единиц с течением времени, как правило, падает. Этому есть две основные причины - банковский процент и инфляция. В экономике есть инструменты для учета изменения стоимости денежных единиц с течением времени. Один из наиболее известных - расчет NPV (Net Present Value) - чистой текущей стоимости. Однако бухгалтерский учет и построенный на данных баланса предприятия экономический анализ финансово-хозяйственной деятельности предприятия пока что, как правило, игнорируют сам факт наличия инфляции. Обсудим некоторые возможности использования индекса инфляции в экономических расчетах.

Переход к сопоставимым ценам. Индекс инфляции даст возможность перехода к сопоставимым ценам, расходам, доходам и другим экономическим величинам. Например, по нашим данным индекс инфляции за 4 года - с 14.03.91 г. по 16.03.95 г. - составил 5936. Это означает, что покупательной способности 1 рубля марта 1991 г. соответствует примерно 6000 (а точнее 5936) рублей марта 1995 г.

Рассмотрим приведение доходов к неизменным ценам. Пусть Иван Иванович Иванов получал в 1990 г. 300 руб. в месяц, а в мае 1995 г. - 1 миллион руб. в месяц. Увеличились его доходы или уменьшились?

Номинальная заработная плата выросла в 1000000/300 = 3333 раза. Однако индекс инфляции на 18 мая 1995 г. составлял 7080, т.е. 1 руб. 1990 г. соответствовал по покупательной способности 7080 руб. в ценах на 18.05.95 г. Следовательно, в ценах 1990 г. доход И.И. Иванова составлял 1000000/7080 = 142 руб. 24 коп., т.е. 47,4% от дохода в 1990 г.

Можно поступить наоборот, привести доход 1990 г. к ценам на 18 мая 1995 г. Для этого достаточно умножить его на индекс инфляции: доход 1990 г. соответствует 300 х 7080 = 2 миллиона 124 тыс. руб. в ценах мая 1995 г.

Средняя зарплата. По данным Госкомстата РФ средняя заработная плата составляла в 1990 г. 297 руб., в октябре 1993 г. - 93 тыс. руб., в январе 1995 г. - 303 тыс. руб. Поскольку зарплата тратится в основном в следующем месяце после получки, то рассмотрим индексы инфляции на 15.11.93 г. и 2.02.95 г., равные 1045 и 4811 соответственно. В ценах 1990 г. средняя зарплата составила 89 руб. и 62 руб.98 коп. соответственно, т.е. 30% и 21,2% от зарплаты 1990 г.

Средняя зарплата рассчитывается путем деления фонда оплаты труда на число работников. При этом объединяются доходы и низкооплачиваемых и сравнительно высокооплачиваемых. Известно, что распределение доходов резко асимметрично, большому числу низкооплачиваемых работников соответствует малое число лиц с высокими доходами. За 1991-1995-е годы дифференциация доходов резко увеличилась. Это означает, что доходы основной массы трудящихся сдвинулись влево относительно средней зарплаты. По нашей оценке 50% получают не более 70% от средней зарплаты, т.е. не более 212100 руб. по состоянию на январь 1995 г., а наиболее массовой является оплата в 50% от средней, т.е. около 150 тыс. руб. в месяц.

Доходы отдельных слоев трудящихся снизились еще существеннее. Зарплата профессора Московского государственного института электроники и математики (технического университета) составляла в марте 1994 г. - 42 руб.92 коп. (в ценах 1990 г.), в июле 1995 г. - 43 руб. 01 коп., т.е. с 1990 г. (400 руб.) снизилась в 9,3 раза, дошла до уровня прежней студенческой стипендии. А студенческие стипендии снизились примерно в той же пропорции и составляли 4-5 руб. в ценах 1990 г.

Кроме того, необходимо учесть, что Госкомстат учитывает начисленную зарплату, а не выплаченную. В отдельные периоды отечественной истории выплата заработной платы откладывалась надолго.

Минимальная зарплата и прожиточный минимум. Минимальная зарплата в сентябре 1994 г. (22500 руб.) и в мае 1995 г. (43700 руб.) составляла 38% и 23,4% соответственно от стоимости минимальной физиологически необходимой продовольственной корзины. После подъема до 55 тыс. руб. она в сентябре 1995 г. составляла около 26,34% от стоимости корзины, т.е. реально уменьшилась в 1,44 раза по сравнению с сентябрем 1994 г. В дальнейшем уменьшение стало еще более заметным.

Минимальная зарплата вместе с единой тарифной сеткой во многом определяет зарплату работников бюджетной сферы. Учитывая снижение коэффициентов тарифной сетки, проведенное весной 1995 г., снижение в 1,5 раза покупательной способности минимальной зарплаты, необходимо заключить, что в сентябре 1995 г. доход бюджетников в 2 раза меньше, чем год назад.

Оценим прожиточный минимум. Бюджетные обследования 1990 года показали, что для лиц с низкими доходами расходы на продовольствие составляют около 50% всех расходов, т.е. на промтовары и услуги идет около 50% доходов. Это соотношение подтвердило и проведенное нашим коллективом бюджетное обследование конца 1995 г. Исходя из него, среднедушевой прожиточный минимум можно оценить, умножая на 2 стоимость минимальной продовольственной корзины ИВСТЭ. Например, на 1 сентября 1995 г. - 418220 руб.. Т.е. прожиточный уровень для семьи из трех человек - муж, жена и ребенок - должен был на 1 сентября 1995 г. составлять 1,25 миллиона руб. (в месяц). Например, муж должен получать 800 тыс. руб., жена - 450 тыс. руб. в месяц. Очевидно, доходы большинства трудящихся меньше прожиточного уровня.

Численные значения стоимостей потребительских корзин и индексов инфляции в настоящей главы рассчитаны в основном по ценам на продукты в Москве и Подмосковье. Однако для других регионов численные значения отличаются мало. Для Москвы индекс инфляции на 1.09.95 г. - 7759, а для Иванова на 1.08.95 г. - 7542. Поскольку потребительская корзина на 14.03.91 г. в Иванове была на 95 коп. дешевле, то и на 1.08.95 г. она несколько дешевле - 195337 руб., а прожиточный минимум равен 390673 руб.. Приведенные выше численные значения для Москвы в качестве первого приближения можно использовать для различных регионов России.

Индексы инфляции с помощью описанной выше методики можно рассчитать для любого региона, профессиональной или социальной группы, отдельного предприятия или даже конкретной семьи. Эти значения могут быть эффективно использованы на трехсторонних переговорах между профсоюзами, работодателями и представителями государства.

Проценты по вкладам в банк, плата за кредит и инфляция. Рассмотрим банк, честно выполняющий свои обязательства. Пусть он дает 10% в месяц по депозитным вкладам. Тогда 1 руб., положенный в банк, через месяц превращается в 1,1 руб., а через 2 - по формуле сложных процентов - в $1{,}1^2 = 1{,}21$ руб., ..., через год - в $1{,}1^{12} = 3{,}14$ руб. Однако за год росли не только вклады, но и цены. Например, с 19.05.94 г. по 18.05.95 г. индекс инфляции составил 3,73. Значит, в ценах на момент оформления вкладов итог годового хранения равен 3,14 / 3,73 = 0,84 руб. Хранение оказалось невыгодным - реальная стоимость вклада уменьшилась на 16%, несмотря на, казалось бы, очень выгодные условия банка.

Пусть фирма получила кредит под 200% годовых. Значит, вместо 1 рубля, полученного в настоящий момент в кредит, через год ей надо отдать 3 рубля. Пусть она взяла кредит 19.05.94 г., а отдает 18.05.95 г. Тогда в ценах на момент взятия кредита она отдает 3/3,73 = 0,80 руб. за 1 руб. кредита, т.е. кредит превратился в подарок - возвращать надо на 20% меньше, чем получил, реальная ставка кредита отрицательна, она равна (- 20)% ! Такова была типичная ситуация в России в течении ряда лет начиная с 1992 г., особенно в 1992-1994 гг. Но бесплатных подарков в бизнесе не бывает - за них надо платить по другим каналам, как правило, криминальным.

Сколько стоит доллар? В июле 1995 г. индекс инфляции около 7000, а курс доллара США - около 4500 руб. за доллар. Следовательно, доллар США стоит 4500 / 7000 = 0,64 руб. в ценах 1990 г., т.е. примерно соответствует официальному обменному курсу в 1980-х годах. В сентябре 1994 г. курс доллара был около 2000, а индекс инфляции - около 2200, т.е. доллар стоил около 0,9 руб. в ценах 1990 г. Реальная покупательная способность доллара упала за 10 месяцев в 1,42 раза.

Ошибочно думать, что на Московской межбанковской валютной бирже курс доллара определяется по законам свободного рынка. На самом деле участвующие в торгах коммерческие банки административно зависят от Центрального Банка РФ. Другой инструмент влияния Центрального Банка - долларовые или рублевые интервенции. Реально курс доллара определяется руководством страны, действующим через Центральный Банк РФ. Одно из следствий реального понижения доллара - легальное присвоение средств тех граждан, которые пытаются сохранить свои сбережения (например на летний отдых), купив доллары США. Другой пример - спекулятивная инфляция, являющаяся следствием искусственного подъема курса доллара после "дефолта" августа 1998 г. Цель этой спекуляции очевидна - выжать рубли из населения с целью увеличения доходов государства (путем увеличения сбора налогов) и поддержки коммерческих банков, существенная часть активов которых "заморожена" в ГКО.

Инфляция, показатели работы предприятия и ВВП. Индексы инфляции используются для пересчета номинальных цен в неизменные, для приведения доходов и расходов к ценам определенного момента времени. Потребительские корзины для промышленных предприятий, конечно, должны отличаться от используемых при изучении жизненного уровня.

Сколько стоит предприятие? Важно оценить основные фонды. Для этого нужно взять их стоимость в определенный момент времени и умножить на индекс инфляции. Интересно отметить, что официальный индекс инфляции Госкомстата более чем в 2 раза меньше , чем наш, если отсчитывать от 1990 г. Причем разрыв все более увеличивается! Есть много способов исказить экономические показатели, и "специалисты" ими умело пользуются. Занижение индекса инфляции выгодно тем, кто хочет обзавестись государственной (т.е. общенародной) собственностью за бесценок, в ущерб истинным владельцам - трудящимся России.

Валовой внутренний продукт, валовой национальный продукт и другие характеристики экономического положения страны рассчитываются в текущих ценах. Для перехода к неизменным ценам надо поделить на индекс инфляции (т.е. умножить на дефлятор). В 2 раза занизишь индекс инфляции - в 2 раза завысишь валовой национальный продукт, валовой внутренний продукт, национальный доход и иные макроэкономические характеристики.

По данным Правительства РФ к концу 1994 г. валовой внутренний продукт составил 53% от уровня 1990 г. Падение больше, чем в Германии в результате разгрома фашизма. И это по официальным данным! Используя же коэффициент занижения со стороны Госкомстата , равный 2 (т.е. заниженный), получаем более реальную цифру - 25% от уровня 1990 г.. Падение в 4 раза! Эта оценка близка к выводам ряда специалистов, независимых от правительства.

Проблема учета инфляции при экономическом анализе финансово-хозяйственной деятельности предприятия. Как известно, разработана и широко применяется развернутая система коэффициентов, используемых при экономическом анализе финансово-хозяйственной деятельности предприятия на основе данных бухгалтерского баланса. Она опирается на два столбца баланса - данные на "начало периода" и данные на "конец периода". Записывают в эти столбцы номинальные значения. В настоящее время инфляцию полностью игнорируют. Это приводит к приукрашиванию реального положения предприятия. По официальной отчетности оно может считаться получившим хорошую прибыль, а по существу - не иметь средств для продолжения деятельности.

Ясно, что учитывать инфляцию надо. Вопрос в другом - как именно. Потребительская корзина должна, видимо, состоять из тех товаров и услуг, которые предприятие покупает. Стоимость основных фондов может не убывать в соответствии с амортизацией, а возрастать согласно отраслевой инфляции, и т.д.. Здесь мы только ставим проблему, не пытаясь ее решить.

7.5. Динамика цен на продовольственные товары с Москве и Московской области

Уже более десяти лет в России осуществляется т.н. радикальная экономическая реформа. Одним из сопутствующих эффектов этой реформы является изменение сложившейся к 1991 г. системы цен на все товары, услуги, труд (рабочую силу). Эти изменения цен приобрели ярко выраженный инфляционный характер. В течении десяти лет "радикальной реформы" произошли изменения не только абсолютных величин цен, но и их пропорций.

Масштабы инфляции были определены не только дисбалансом между скопившейся к 1992 г. у населения значительной массой наличных денег и наличием товаров, но и массовым преобразованием безналичных средств предприятий в наличные деньги в период расцвета совместных предприятий и кооперативов в 1989-91 гг.. В дальнейшем в результате применения жестких мер (например, невыплаты заработной платы), ограничивающих поступление на рынок товаров и услуг наличных денег, а также ограничивающих количество покупателей среднего класса и тем самым обеспечивающих искусственное снижение спроса, темпы инфляции заметно снизились, но инфляция не прекратилась. Болезнь не исчезла. Удалось сбить температуру больного, т.е. отключить сигнальную систему, но не вылечить болезнь. Стоит только лишь начать платить людям наемного труда заработанные ими деньги при условии корректной оценки труда как рыночного товара, как инфляционная болезнь возобновится. В августе 1998 г. инфляция была подстегнута руководством страны путем искусственного подъема курса доллара.

Институт высоких статистических технологий и эконометрики изучает динамику экономического положения граждан России на основе независимо собранной информации. Приведение к сопоставимым ценам (с помощью индексов инфляции и дефляторов) - составная часть любого экономического расчета, связанного более чем с одним моментом времени. Как показали наши наблюдения над ценами, использование публикуемых Госкомстатом РФ значений индексов инфляции приводит к систематическим ошибкам. Так, например, по нашим данным цены за 6 с небольшим лет (с марта 1991 г. по май 1997 г.) выросли в среднем примерно в 10000 раз, а по данным Госкомстата РФ - примерно в 6000 раз. Сказанное определяет актуальность использования независимой информации о ценах и индексах инфляции при анализе экономического положения России, а также при разработке прикладных моделей и методов управления в рыночной экономике России.

Предметом описанного здесь исследования является оценка изменения в ходе реформ фактического среднего и минимального физиологически необходимого уровней жизни граждан РФ через сравнение индексов инфляции, вычисленных на основании потребительских корзин, и индекса изменения величины средней заработной платы.

В 1994-97 гг. еженедельно собрались данные о ценах 35 продуктов в 12 точках Москвы, Подмосковья и Крыма. А именно, информация о ценах собиралась:

в 9 точках г. Москвы;

в 2-х точках Московской области (г. Раменское и г. Ногинск)

и в Крыму (г. Симферополь).

Регулярное измерение цен производилось с интервалом в одну неделю начиная с декабря 1995 г. по 35 различным товарам. Измерение цен в 1994 - 1995 гг. производилось менее регулярно и имеющиеся в нашем распоряжении временные ряды в этом интервале времени имеют пропуски (как правило, в летнее время).

Расчеты по собранным ценам продовольственных товаров проводились для следующих 5 потребительских корзин:

ИВСТЭ - продовольственная потребительская (продуктовая) корзина Института высоких статистических технологий и эконометрики Составлена с учетом разработок Института питания РАМН. Является сбалансированной по белкам, жирам и углеводам. Обеспечивает минимальные физиологически необходимые потребности человека. Приведена выше;

ГКС-1 - продуктовая корзина из 19 продуктов питания (включая сигареты) Государственного комитета РФ по статистике, применявшаяся в 1993-1996 гг. Приведена выше ;

ГКС-2 - продуктовая корзина Государственного комитета РФ по статистике, используемая с 1 января 1997 г. Нормы потребления предложены Министерством труда;

Бюдж-1 - продуктовая корзина, разработанная на основе бюджетного обследования "бедных семей" студентов Московского государственного института электроники и математики (технического университета) (среднедушевое потребление не превосходит 90 % от медианы обследованной совокупности семей). Эта корзина определяет усредненные фактические объемы потребления (кг./год/чел. и, соответственно, кг./мес./чел.) 35 продуктовых товаров, по которым производились измерения цен при использовании корзины **ИВСТЭ**. Общий объем затрат "бедных семей" на продуктовые и иные товары предлагается находить умножением стоимости корзины **Бюдж-1** на соответствующие коэффициенты;

Бюдж-2 - продовольственная потребительских корзина, разработанная на основе бюджетного обследования семей студентов Московского государственного института электроники и математики (технического университета) (октябрь-ноябрь 1995 г.). Совокупность обследованных семей в целом характеризуется средним уровнем потребления. Эта корзина определяет усредненные фактические объемы потребления (кг./год/чел. и, соответственно, кг./мес./чел.) 35 продуктовых товаров, по которым производились измерения цен в корзине **ИВСТЭ**. Общий объем затрат "семей со средним достатком" на продуктовые и иные товары предлагается находить умножением стоимости корзины **Бюдж-2** на соответствующие коэффициенты.

Количество элементарных измерений (значений собранных цен продовольственных товаров) приблизительно равно 30000. Собранные данные о ценах обрабатывались на компьютерах Macintosh как известными, так и оригинальными методами. Точность вычислений равна обычной компьютерной точности на компьютере *Macintosh* при работе с числами с плавающей точкой. В дальнейшем все средние величины цен приведены с точностью до одного рубля, а величины процентов - с точностью до 1/10 процента.

Приведем некоторые результаты анализа данных. Начнем с временных рядов стоимостей потребительских корзин в Москве .Оказалось, что стоимость потребительской корзины **ГКС-2** примерно в 1,5 раза меньше стоимости потребительской корзины **ГКС-1**. Потребительская корзина **ИВСТЭ** располагалась по стоимости примерно посередине между **ГКС-1** и **ГКС-2**. Несмотря на различие стоимостей, индексы инфляции для всех трех корзин **ГКС-1, ГКС-2, ИВСТЭ** близки и составляют 8233-8896 на конец декабря 1995 г. и 10396-10890 на конец февраля 1997 г. Любопытно отметить, что **ГКС-1** имеет наименьшие значения индекса из трех корзин, а **ГКС-2** - наибольшие, если сравнивать с мартом 1991 г. (Госкомстат РФ такие сравнения не проводит), а то время как рост цен за исследуемый промежуток времени (с конца декабря 1995 г. по конец февраля 1997 г.) наибольший рост дает корзина **ИВСТЭ** (28,05%), а наименьший - **ГКС-2** (22,42%).

Совсем иная картина со стоимостями потребительских корзин **Бюдж-1** и **Бюдж-2**. Они относятся к реальному потреблению сравнительно обеспеченных москвичей, включают в себя стоимости не только продуктов, но и других товаров и услуг, в то время как корзины **ИВСТЭ**, **ГКС-1** и **ГКС-2** дают представление о стоимости минимального набора товаров и услуг, обеспечивающего физиологические потребности человека. В конце декабря 1995 г. стоимость корзины **Бюдж-1** (для "бедных") составляла 659852 руб., а корзины **Бюдж-2** (для "средних" семей) - 726364 руб., а к февралю 1997 г. они "подросли" до 832498 руб. (на 26,16%) и 950989 руб. (на 30,92%) соответственно. Эти величины больше прожиточного минимума согласно данным Московской федерации профсоюзов (750 тыс. руб. в мае 1997 г.), хотя разницу нельзя назвать заметной.

Интереснее другое - общий рост цен (на февраль 1997 г.) составил 8060-8446, т.е. примерно на 20% меньше, чем рост стоимостей корзин **ИВСТЭ**, **ГКС-1**, **ГКС-2**. Значит, "реформы" тяжелее всего ударили по наиболее дешевым товарам, предназначенным для наиболее бедной части населения. Это связано, видимо, с сокращением и прекращением дотаций для таких товаров. Правда, затем темпы роста выровнялись - при сравнении февраля 1997 г. с декабрем 1995 г. они составляют 28,05% для корзины **ИВСТЭ**, 26,27% - для **ГКС-1**, 26,16% - для **Бюдж-1** и 30,92% для **Бюдж-2**. Особняком стоит **ГКС-2** - 22,42%, заметно меньше, чем для других корзин. В то же время наибольший рост для корзины

Бюдж-2 может указывать на тенденцию более быстрого роста цен на товары, предназначенные для более состоятельных людей.

Анализ временных рядов стоимостей потребительских корзин и индексов инфляции по Подмосковью в целом подтверждает приведенные выше выводы, сделанные по московским данным. Снова наблюдаем близость роста цен с 1991 г. для корзин **ИВСТЭ**, **ГКС-1**, **ГКС-2**, снова **ГКС-2** в полтора раза дороже **ГКС-1**, снова темп роста с декабря 1995 г. меньше всего из этих трех корзин у **ГКС-2.** Снова индексы с 1991 г. для корзин **Бюдж-1** и **Бюдж-2** на 20-25% меньше, чем для первых трех корзин. Однако с декабря 1995 г. наибольший рост стоимости корзины - не у двух последних, а у **ГКС-1** (на втором месте - корзина **ИВСТЭ**). Возможно, это отражает меньшую долю состоятельных людей в Подмосковье и соответственно меньшую ориентацию торговцев на их "покупательные" возможности.

Обращает на себя внимание меньшая величина индексов с марта 1991 г. в Подмосковье по сравнению с Москвой. Возможно, дело в том, что стоимости потребительских корзин по состоянию на 31 марта 1991 г. брались те же, что и в Москве, поскольку сведения о ценах на тот момент в Московской области у нас отсутствуют. Это приводит к занижению истинных значений индексов инфляции, поскольку и до 31 марта 1991 г. цены в Подмосковье были несколько ниже, чем в Москве. Это относится, в частности, к ценам на овощи и фрукты, молочные продукты и др.

Вполне естественно, что с марта 1991 г. цены на различные товары выросли по-разному. Так, цены на рыбу (треска, минтай) выросли примерно в 25000 раз, а цена на сахар - менее чем в 4000 раз. Цены на творог выросли в 2,5 раза больше, чем на сыр, и т.д. В Москве и Московской области рост цен достаточно хорошо согласован. Можно было бы предположить, что в рыночных условиях были исправлены диспропорции прежней дотационной плановой системы. Тогда рост цен после декабря 1995 г. должен был бы быть примерно равномерным, отражающим динамику общеэкономических процессов. Однако конкретные эмпирические данные о динамике цен отвергают это предположение.

В Москве при общем среднем росте цен на 20-30 % больше всего выросли цены на огурцы (74,8%), баранину (75,9%), птицу (74,5%), упали цены на капусту (- 4,6%), сахар (- 5,5%). В Московской области при таком же среднем росте цен больше всего выросли цены на мясо - на говядину (82,6%), свинину (88,6%), баранину (107,6%), при этом упали цены на картофель (-10%), капусту (-10%), сахар (-13,1%), конфеты (-21,1%), минтай (-6,5%), растительное масло (-20,6%) и маргарин (- 13 %).

Приходится констатировать, что цены растут непропорционально, стабилизация цен не наступила, более того, динамика цен на отдельные товары не только не согласована, но и отнюдь не близка. **Нет никаких признаков приближения к равновесным ценам**, чего можно было бы ожидать после пяти лет "либерализации". В качестве дополнительного следствия из сказанного вытекает, что, подбирая нужным образом номенклатуру товаров для потребительской корзины, можно получить индекс инфляции желательной величины - от значительного роста (+ 80 %) до падения цен (- 20 %).

Временные ряды наименьшей, средней и наибольшей из зарегистрированных по Москве цен 35 продовольственных товаров показывают, что **такое понятие, как "цена товара", строго говоря, не корректно**. Оно применимо к единственному акту купли-продажи определенного товара в фиксированном месте, в крайнем случае - к актам купли-продажи в определенном магазине, но не к огромному городу в целом. Действительно, зафиксированные нашими сотрудниками цены на один и тот же товар в один и тот же день могут различаться в несколько раз. Так, 26 июня 1996 г. максимальная зафиксированная цена на рис превышает минимальную в 3,04 раза, а на картофель - в 3,13 раза. Аналогичное превышение для баранины 27 декабря 1996 г. равно 2,79. Типовое же превышение максимальной цены над максимальной - в 1,5 раза. Ничего странного в сказанном нет - всем московским потребителям известно, что наибольшие цены - в центральных престижных магазинах, средние - в рядовых магазинах, наименьшие - на "оптовых" рынках.

С целью обеспечение сопоставимости данных наши сотрудники собирали данные в одних и тех же местах (магазинах, киосках, рынках). Это позволяло отслеживать рост цен и получать корректные значения индексов инфляции. Однако это делало несколько условной стоимость потребительской корзины - потребитель, потратив время и обойдя достаточное количество мест продажи,, мог обеспечить себя теми же продуктами по менее высоким ценам. Дополнительную сложность вносит большая номенклатура видов одного и того же товара. На какой тип батона белого хлеба ориентироваться? Что понимать под говядиной - отечественную или импортную, вырезку или кости для супа? Объективно существующая свобода при решении организаторами исследования жизненного уровня подобных вопросов дает возможность для сдвига результатов в заранее заданном направлении. Объективно цены не являются стабильными в пространстве и во времени.

На практике указанные сложности в основном преодолимы. Оказалось, в частности, что стоимость потребительских корзин в различных районах Москвы отличается хотя и отличается, но не более чем на 5-10%. Отклонения в стоимости отдельных продуктов частично компенсируют друг друга.

Нами изучены вклады отдельных продовольственных товаров в стоимости потребительских корзин. Обращает на себя внимание различие между нормативными (т.е. заданными априори) корзинами **ИВСТЭ**, **ГКС-1**, **ГКС-2** и полученными в результате анализа реального потребления корзинами **Бюдж-1** и **Бюдж-2**. В реальном потреблении гораздо меньше муки, пшена, геркулеса, ржаного хлеба, картофеля, трески, минтая, молока, маргарина, но гораздо больше лука, яблок, конфет, колбасы, сельди, сливочного масла, сыра. Объяснение достаточно очевидное: корзины **ИВСТЭ**, **ГКС-1**, **ГКС-2** - это "корзины выживания", действительно минимальные по стоимости корзины, в то время как корзины **Бюдж.1** и **Бюдж.2** - это корзины реального потребления в семьях студентов Московского государственного института электроники и математики (технического университета) различного достатка.

Продовольственные товары, на наш взгляд, можно разделить на две группы. Цены на одни растут монотонно, без всякой связи со временем года, т.е. ведут себя примерно также, как промышленные товары. Можно предположить, что индексы инфляции, построенные по подмножеству таких товаров, представляют собой общие индексы, "очищенные от сезонности", а потому лучше описывающие реальное состояние экономики, чем исходные индексы. Однако при их применении теряется связь со стоимостью корзины выживания, обеспечивающей существование без физиологического вырождения.

Второе подмножество - это товары с ярко выраженной сезонностью, прежде всего овощи, цены на которые падают во второй половине лета и осенью, а затем начинают возрастать. Наличие этой составляющей приводит к тому, что рост стоимостей корзин практически останавливается летом, а наиболее быстрым является зимой.

Можно ли управлять процессом роста цен? Мы наблюдали результаты явно административного воздействия: в ноябре 1995 г., перед выборами в Государственную Думу, цены в Москве внезапно упали на 9 %, хотя в ноябре цены обычно растут быстрее, чем в иное время года. Тем не менее необходимо констатировать, что обычно изменение цен происходит на микроэкономическом уровне, хотя и провоцируется макроэкономическими процессами, в частности, монопольными изменениями цен на энергоносители.

Ложная, на наш взгляд, идея монетаристов состоит в том, что они считают необходимым бороться с инфляцией, сокращая денежную массу в стране, например, не выплачивая вовремя зарплату и пенсии. Однако, как пишет академик-секретарь Отделения экономики РАН Д.С. Львов: "Макроэкономические расчеты показывают, что за каждый процент сокращения инфляции приходится расплачиваться тремя-пятью процентами спада производства" [6, с.11]. Основной удар монетаристской политики приходится не по инфляции, а по производству.

Процесс инфляции частично управляем административными методами. Осенью 1996 г. спрогнозированного ИВСТЭ роста цен не произошло, что объясняется изменением условий - правительство перешло к борьбе с инфляцией путем гигантского роста задолженностей по зарплате, пенсиям и другим платежам (например, детским пособиям, стипендиям студентов).

Если у населения нет денег - торговцы не поднимают цены. Так, в Москве за 2 года - с лета 1995 г. по лето 1997 г. цены выросли примерно на 50 %, в то время как в г. Иваново - лишь на 15 %, а импортные товары на ивановских рынках стоят на 1/3 дешевле, чем на московских (хотя эти импортные товары закупаются в Москве). Объяснить это можно тем, что экономическое положение в Иваново гораздо хуже, чем в Москве, ниже уровень доходов, больше безработных, что вынуждены учитывать торговцы.

Расчет индекса инфляции - вспомогательная задача. решение которой необходимо для приведения экономических характеристик к сопоставимому виду. Важнейшей задачей является расчет реальной заработной платы, равной частному от деления номинальной заработной платы на индекс инфляции. Известно, что цены на промышленные товары и на услуги, как правило, растут быстрее, чем на продовольствие. Поэтому рассчитываемые по продовольственным потребительским корзинам значения индексов инфляции дают оценку снизу для роста потребительских цен и стоимости жизни в целом.

Минимальный прожиточный минимум оцениваем по методу американской исследовательницы польского происхождения М. Оршански с коэффициентом Энгеля 0,5. Этот метод основан на расчете стоимости минимальной продовольственной корзины и учете стоимостей остальных минимально необходимых затрат с помощью коэффициентов. Так, для "бедных семей" студентов Московского государственного института электроники и математики (технического университета) во время пробного бюджетного обследования в октябре-ноябре 1995 г. затраты на продовольствие составили 52% от всех расходов. Поэтому стоимость прожиточного минимума для них получим, приняв за 52% стоимость минимальной продовольственной корзины **ИВСТЭ**, т.е. умножив ее стоимость на 1/0,52 = 1,92.

Метод М. Оршански предполагает, что структура затрат практически не меняется. Однако, как уже отмечалось, цены на промышленные товары и на услуги растут быстрее, чем на продовольствие. Поэтому замена 1,92 на 2,00 представляется обоснованной. Полученные значения (на май 1997 г. - 700 тыс. руб. в месяц на человека) хорошо согласуется с уже цитированными данными Московской федерации профсоюзов (750 тыс. руб.). Отметим, что для всей совокупности семей, чьи бюджеты были обследованы в 1996 г., затраты на продовольствие составили 42 %, т.е. для них коэффициент Оршански равен 1/0,42 = 2,38.

Средняя (начисленная) заработная плата в Москве составляла в декабре 1996 г. 1,12 миллиона руб. (в России - 0.84 млн.). В сопоставлении со сказанным выше (с учетом логнормального характера функции распределения доходов и наличия детей) это означает, что даже в Москве по крайней мере половина семей живет ниже прожиточного уровня. В 1990 г. средняя зарплата превышала прожиточный минимум в 5.5 раз, а в 1997 г. - лишь в 1.2 раза (по России), т.е. уровень жизни упал в среднем в 4,6 раза. Он весной 1998 г. соответствовал концу 50-х - началу 60-х годов. За август-сентябрь 1998 г. корзина **ИВСТЭ** подорожала в 1,5 раза, следовательно, уровень жизни упал уже в 7 раз, и по покупательной способности зарплаты рядовые граждане "приблизились" к возможностям начала 50-х годов.

Переход к сопоставимым ценам необходимо использовать также при расчете таких макроэкономических характеристик, как валовой внутренний продукт, объем бюджетных ассигнований и т.д. С учетом сказанного выше можно утверждать, что экономика России с 1990 г. по 1998 г. была "сокращена" в 4-6 раз, что соответствует сдвигу назад по времени на 35-45 лет.

Материалы настоящего пункта обсуждаются в работе [7].

Эконометрика описывает инфляцию. Причины инфляции - это предмет иных экономических наук. Однако несколько слов сказать об этом полезно.

Всегда говорят об *инфляции спроса.* Это ситуация, когда у населения много денег*, которые оно хочет истратить*. А товаров мало. Тогда цены растут. Либо непосредственно, либо через механизм " черного рынка".

Другой вид инфляции - *инфляция издержек*. Производитель вынужден повышать цену на свою продукцию, потому что его поставщики повышают цены на собственную продукцию. Этот порочный круг очень трудно разорвать.

Третий вид инфляции - *административная инфляция.* Цены повышает государство. Естественно, на то, что оно контролирует. Например, с августа по декабрь 1998 г. курс доллара США был поднят в 4 раза. Последствия были понятные: адекватный подъем цен на импортные товары, рост цен на продукцию, для изготовления которой использовались импортные комплектующие, а затем и рост цен на чисто отечественную продукцию, если такая вообще существует. В результате инфляция за год составила 80%.

Выше уже приводились примеры административного регулирования цен. Политика государственных органов в области энергетики, транспорта, экспорта и импорта и других сфер государственного регулирования экономики оказывает непосредственное влияние на инфляцию.

Ряд вопросов анализа и прогнозирования инфляционных процессов рассмотрен в главах 5, 6, 12 и др. Судя по опыту последних десяти лет, инфляционные процессы стали постоянной составляющей отечественной экономической жизни, и экономистам, менеджерам, инженерам различных специальностей придется учитывать их свойства в своей работе.

Цитированная литература

1. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979. - 296 с.

2. Самуэльсон П. Экономика. Тт.1,2. - М.: МГП "Алгон" - ВНИИСИ, 1992. - 333 с. + 415 с.

3. Математическое моделирование процессов налогообложения (подходы к проблеме) / Под ред. В.Г. Кольцова, В.Н. Жихарева, Н.Ю. Ивановой, А.И. Орлова. - М.: Изд-во ЦЭО Министерства общего и профессионального образования РФ, 1997. - 232 с. (Авторский коллектив: Балашов В. В., Букина Е. П., Жихарев В. Н., Иванова И. Г., Иванова Н. Ю., Иванова Р. К., Кастосов М. А., Кольцов В. Г., Кулага Е.В., Нечаева Е. Г., Орлов А. И., Орлова Л. А., Рафальская А. Э., Светлов С.В., Семенова О. В., Стешов И. В., Цупин В. А.)

4. Макконнелл Кэмпбелл Р., Брю Стэнли Л. Экономикс: Принципы, проблемы и политика. В 2-х т.: Пер. с англ. 11-го изд. Т.1. - М.: Республика, 1995. - 400 с.

5. Коростикова Т. Цены вырастут в 5 раз // Аргументы и факты, 1994, No.16, с.5.

6. Львов Д.С. Реформы с позиции современной науки. - Научные труды Международного союза экономистов и Вольного экономического общества России. Том второй. - М.- СПб: 1995, с.7-16.

7. Орлов А.И., Жихарев В.Н., Цупин В.А., Балашов В.В. Как оценивать уровень жизни? (На примере московского региона.) - Журнал «Обозреватель-Observer». 1999. No.5 (112). С. 80-83.

Глава 8. Статистика нечисловых данных

Статистика нечисловых данных - это направление в эконометрике, в котором в качестве исходных статистических данных (результатов наблюдений) рассматриваются объекты нечисловой природы. Так принято называть объекты, которые нецелесообразно описывать числами, в частности элементы нелинейных пространств. Примерами являются бинарные отношения (ранжировки, разбиения, толерантности и др.), результаты парных и множественных сравнений, множества, нечеткие множества, измерение в шкалах, отличных от абсолютных. Этот перечень примеров не претендует на законченность. Он складывался постепенно в соответствии с развитием теоретических исследований в области статистики нечисловых данных и расширением опыта применений этого направления эконометрики.

Объекты нечисловой природы широко используются в теоретических и прикладных исследованиях по экономике, менеджменту и другим проблемам управления, в частности управления качеством продукции, в технических науках, социологии, психологии, медицине и т.д., а также практически во всех отраслях народного хозяйства.

8.1. Объекты нечисловой природы

Начнем с первоначального знакомства с основными видами объектов нечисловой природы.

Результаты измерений в шкалах, отличных от абсолютной. Рассмотрим конкретное исследование в области маркетинга образовательных услуг, послужившее поводом к развитию отечественных исследований по теории измерений (см. главу 3). При изучении привлекательности различных профессий для выпускников новосибирских школ был составлен список из 30 профессий. Опрашиваемых просили оценить каждую из этих профессий одним из баллов 1,2,...,10 по правилу: чем больше нравится, тем выше балл. Для получения социологических выводов необходимо было дать единую оценку привлекательности определенной профессии для совокупности выпускников школ. В качестве такой оценки в работе [1] использовалось среднее арифметическое баллов, выставленных профессии опрошенными школьниками. В частности, физика получила средний балл 7,69, а математика - 7,50. Поскольку 7,69 больше, чем 7,50, был сделан вывод, что физика более предпочтительна для школьников, чем математика.

Однако этот вывод противоречит данным работы [2], согласно которым ленинградские школьники средних классов больше любят математику, чем физику. Обсудим одно из возможных объяснений этого противоречия, которое сводится к указанию на неадекватность (с точки зрения теории измерений) методики обработки эконометрических данных, примененной в работе [1].

Дело в том, что баллы 1,2,...,10 введены конкретными исследователями, т.е. субъективно. Если одна профессия оценена в 10 баллов, а вторая - в 2, то из этого нельзя заключить, что первая ровно в 5 раз привлекательней второй. Другой коллектив социологов мог бы принять иную систему баллов, например

1,4,9,16,...,100. Естественно предположить , что упорядочивание профессий по привлекательности, присущее школьникам, не зависит от того, какой системой баллов им предложит пользоваться маркетолог. Раз так, то распределение профессий по градациям десятибалльной системы не изменится, если перейти к другой системе баллов с помощью любого допустимого преобразования в порядковой шкале (см. главу 3), т.е. с помощью строго возрастающей функции $g: R^1 \to R^1$. Если , $Y_1, Y_2,...,Y_n$ -ответы n выпускников школ, касающихся математики, а $Z_1, Z_2,...,Z_n$ -физики, то после перехода к новой системе баллов ответы относительно математики будут иметь вид $g(Y_1), g(Y_2),...,g(Y_n)$, а относительно физики - $g(Z_1), g(Z_2),...,g(Z_n)$.

Пусть единая оценка привлекательности профессии вычисляется с помощью функции $f(X_1, X_2,...,X_n)$. Какие требования естественно наложить на функцию $f: R^n \to R^1$, чтобы полученные с ее помощью выводы не зависели от того, какой именно системой баллов пользовался специалист по маркетингу образовательных услуг?

Замечание. Обсуждение можно вести в терминах экспертных оценок. Тогда вместо сравнения математики и физики n экспертов (а не выпускников школ) оценивают по конкурентоспособности на мировом рынке, например, две марки стали. Однако в настоящее время маркетинговые и социологические исследования более привычны, чем экспертные.

Единая оценка вычислялась для того, чтобы сравнивать профессии по привлекательности. Пусть $f(X_1, X_2,...,X_n)$ - среднее по Коши. Пусть среднее по первой совокупности меньше среднего по второй совокупности:

$$f(Y_1, Y_2,...,Y_n) < f(Z_1, Z_2,...,Z_n).$$

Тогда согласно теории измерений (см. главу 3) необходимо потребовать, чтобы для любого допустимого преобразования g из группы допустимых преобразований в порядковой шкале было справедливо также неравенство

$$f(g(Y_1), g(Y_2),...,g(Y_n)) < f(g(Z_1), g(Z_2),...,g(Z_n)).$$

т.е. среднее преобразованных значений из первой совокупности также было меньше среднего преобразованных значений для второй совокупности. Причем сформулированное условие должно быть верно для любых двух совокупностей $Y_1, Y_2,...,Y_n$ и $Z_1, Z_2,...,Z_n$ и, напомним, любого допустимого преобразования. Средние величины, удовлетворяющие сформулированному условию, называют допустимыми (в порядковой шкале). Согласно теории измерений только такими средними можно пользоваться при анализе мнений выпускников школ, экспертов и иных данных, измеренных в порядковой шкале.

Какие единые оценки привлекательности профессий $f(X_1, X_2,...,X_n)$ устойчивы относительно сравнения? Ответ на этот вопрос дан в главе 3. В частности, оказалось, что средним арифметическим, как в работе [1] новосибирских специалистов по маркетингу образовательных услуг, пользоваться нельзя, а порядковыми статистиками, т.е. членами вариационного ряда (и только ими) - можно.

Методы анализа конкретных экономических данных, измеренных в шкалах, отличных от абсолютной, являются предметом изучения в статистике нечисловых данных как части эконометрики. Как известно, основные шкалы измерения делятся на качественные (шкалы наименований и порядка) и количественные (шкалы интервалов, отношений, разностей, абсолютная). Методы анализа статистических данных в количественных шкалах сравнительно мало отличаются от таковых в абсолютной шкале. Добавляется только требование инвариантности относительно преобразований сдвига и/или масштаба. Методы анализа качественных данных - принципиально иные. О них пойдет речь в настоящей главе.

Напомним, что исходным понятием теории измерений является совокупность $\Phi = \{\varphi\}$ допустимых преобразований шкалы (обычно Ф- группа), $\varphi : R^1 \to R^1$. Алгоритм обработки данных *W*, т.е. функция $W : R^n \to A$ (здесь *A*- множество возможных результатов работы алгоритма) называется адекватным в шкале с совокупностью допустимых преобразований Ф, если

$$W(x_1, x_2,..., x_n) = W(\varphi(x_1), \varphi(x_2),..., \phi(x_n))$$

для всех $x_i \in R^1, i = 1,2,..., n,$ и всех $\varphi \in \Phi$. Таким образом, теорию измерений рассматриваем как теорию инвариантов относительно различных совокупностей допустимых преобразований Ф. Интерес вызывают две задачи:

а) дана группа допустимых преобразований Ф (т.е. задана шкала); какие алгоритмы анализа данных *W* из определенного класса являются адекватными?

б) дан алгоритм анализа данных *W*; для каких шкал (т.е. групп допустимых преобразований Ф) он является адекватным?

В главе 3 первая задача рассматривалась для алгоритмов расчета средних величин. Информацию о других результатах решения задач указанных типов можно найти в работах [3-5].

Бинарные отношения. Пусть $W : R^n \to A$ - адекватный алгоритм в шкале наименований. Можно показать, что этот алгоритм задается некоторой функцией от матрицы $B = \| b_{ij} \| = B(x_1, x_2,..., x_n),$ где

$$b_{ij} = \begin{cases} 1, & x_i = x_j, i, j = 1,2,..., n, \\ 0, & x_i \neq x_j, i, j = 1,2,..., n. \end{cases}$$

Если $W : R^n \to A$ - адекватный алгоритм в шкале порядка, то этот алгоритм задается некоторой функцией от матрицы $C = \| c_{ij} \| = C(x_1, x_2,..., x_n)$ порядка $n \times n$, где

$$c_{ij} = \begin{cases} 1, & x_i \leq x_j, i, j = 1,2,..., n, \\ 0, & x_i > x_j, i, j = 1,2,..., n. \end{cases}$$

Матрицы *B* и *C* можно проинтерпретировать в терминах бинарных отношений. Пусть некоторая характеристика измеряется у *n* объектов $q_1, q_2, ..., q_n$, причем x_i - результат ее измерения у объекта q_i Тогда матрицы *B* и *C* задают бинарные отношения на множестве объектов $Q = \{q_1, q_2, ..., q_n\}$. Поскольку бинарное отношение можно рассматривать как подмножество декартова квадрата $Q \times Q$, то любой матрице $D = ||d_{ij}||$ порядка $n \times n$ из 0 и 1 соответствует бинарное отношение *R(D),* определяемое следующим образом: $(q_i, q_j) \in R(D)$ тогда и только тогда, когда $d_{ij} = 1$.

Бинарное отношение *R(B)* - отношение эквивалентности, т.е. рефлексивное симметричное транзитивное отношение. Оно задает разбиение *Q* на классы эквивалентности. Два объекта q_i и q_j входят в один класс эквивалентности тогда и только тогда, когда $x_i = x_j, b_{ij} = 1$.

Выше показано, как разбиения возникают в результате измерений в шкале наименований. Разбиения могут появляться и непосредственно. Так, при оценке качества промышленной продукции эксперты дают разбиение показателей качества на группы. Для изучения психологического состояния людей их просят разбить предъявленные рисунки на группы сходных между собой. Аналогичная методика применяется и в иных экспериментальных психологических исследованиях, необходимых для оптимизации управления персоналом.

Во многих эконометрических задачах разбиения получаются "на выходе" (например, в кластер - анализе) или же используются на промежуточных этапах анализа данных (например, сначала проводят классификацию с целью выделения однородных групп, а затем в каждой группе строят регрессионную зависимость).

Бинарное отношение *R(C)* задает разбиение *Q* на классы эквивалентности, между которыми введено отношение строгого порядка. Два объекта q_i и q_j входят в один класс тогда и только тогда, когда c_{ij}= 1 и c_{ji}= 1, т.е. $x_i = x_j$. Класс эквивалентности Q_1 предшествует классу эквивалентности Q_2 тогда и только тогда, когда для любых $q_i \in Q_1, q_j \in Q_2$ имеем c_{ij} = 1, c_{ji}= 0, т.е. $x_i < x_j$. Такое бинарное отношение в статистике часто называют ранжировкой со связями; связанными считаются объекты, входящие в один класс эквивалентности. В литературе встречаются и другие названия: линейный квазипорядок, упорядочение, квазисерия, ранжирование. Если каждый из классов эквивалентности состоит только из одного элемента, то имеем обычную ранжировку (другими словами, линейный порядок).

Как известно, ранжировки возникают в результате измерений в порядковой шкале. Так, при описанном выше опросе ответ выпускника школы - это ранжировка (со связями) профессий по привлекательности. Ранжировки часто возникают и непосредственно, без промежуточного этапа - приписывания объектам квазичисловых оценок - баллов. Многочисленные примеры тому даны М. Кендэлом [6]. При оценке качества промышленной продукции нормативные методические документы предусматривают использование ранжировок.

Для прикладных областей, кроме ранжировок и разбиений, представляют интерес толерантности, т.е. рефлексивные симметричные отношения. Толерантность - математическая модель для выражения представлений о сходстве (похожести, близости). Разбиения - частный вид толерантностей. Толерантность, обладающая свойством транзитивности - это разбиение. Однако в общем случае толерантность не обязана быть транзитивной. Толерантности появляются во многих постановках теории экспертных оценок, например, как результат парных сравнений (см. ниже).

Напомним, что любое бинарное отношение на конечном множестве может быть описано матрицей из 0 и 1.

Дихотомические (бинарные) данные. Это данные, которые могут принимать одно из двух значений (0 или 1), т.е. результаты измерений значений альтернативного признака. Как уже было показано, измерения в шкале наименований и порядковой шкале приводят к бинарным отношениям, а те могут быть выражены как результаты измерений по нескольким альтернативным признакам, соответствующим элементам матриц, описывающих отношения. Дихотомические данные возникают в прикладных исследованиях и многими иными путями.

В настоящее время в большинстве стандартов на конкретную продукцию предусмотрен контроль по альтернативному признаку. Это означает, что единица продукции относится к одной из двух категорий - "годных" или "дефектных", т.е. соответствующих или не соответствующих требованиям стандарта. Отечественными специалистами проведены обширные теоретические исследования проблем статистического приемочного контроля по альтернативному признаку. Основополагающими в этой области являются работы академика А.Н.Колмогорова. Подход советской вероятностно-статистической школы к проблемам контроля качества продукции отражен в монографиях [7,8] (см. также главу 13).

Дихотомические данные - давний объект математической статистики. Особенно большое применение они имеют в экономических и социологических исследованиях, в которых большинство переменных, интересующих специалистов, измеряется по качественным шкалам. При этом дихотомические данные зачастую являются более адекватными, чем результаты измерений по методикам, использующим большее число градаций. В частности, психологические тесты типа MMPI используют только дихотомические данные. На них опираются и популярные в технико-экономическом анализе методы парных сравнений [9].

Элементарным актом в методе парных сравнений является предъявление эксперту для сравнения двух объектов (сравнение может проводиться также прибором). В одних постановках эксперт должен выбрать из двух объектов лучший по качеству, в других - ответить, похожи объекты или нет. В обоих случаях ответ эксперта можно выразить одной из двух цифр (меток)- 0 или 1. В первой постановке: 0, если лучшим объявлен первый объект; 1 - если второй. Во второй постановке: 0, если объекты похожи, схожи, близки; 1 - в противном случае.

Подводя итоги изложенному, можно сказать, что рассмотренные выше данные представимы в виде векторов из 0 и 1 (при этом матрицы, очевидно, могут быть записаны в виде векторов). Поскольку все результаты наблюдений имеют лишь несколько значащих цифр, то, используя двоичную систему счисления, любые виды анализируемых эконометрическими методами данных можно записать в виде векторов из 0 и 1. Представляется, что эта возможность имеет лишь академический интерес, но во всяком случае можно констатировать, что анализ дихотомических данных необходим во многих прикладных постановках.

Множества. Совокупность X^n векторов $X = (x_1, x_2,...,x_n)$ из 0 и 1 размерности n находится во взаимно-однозначном соответствии с совокупностью 2^n всех подмножеств множества $N = \{1, 2, ..., n\}$. При этом вектору $X = (x_1, x_2,...,x_n)$ соответствует подмножество $N(X) \subseteq N$, состоящее из тех и только из тех i, для которых $x_i = 1$. Это объясняет, почему изложение вероятностных и статистических результатов, относящихся к анализу данных, являющихся объектами нечисловой природы перечисленных выше видов, можно вести на языке конечных случайных множеств, как это было сделано в монографии .[3].

Множества как исходные данные появляются и в иных постановках. Из геологических реалий исходил Ж. Матерон, из электротехнических - Н.Н. Ляшенко и др. Случайные множества применялись для описания процесса случайного распространения, например распространения информации, слухов, эпидемии или пожара, а также в математической экономике. В монографии [3] рассмотрены приложения случайных множеств в теории экспертных оценок и в теории управления запасами и ресурсами (логистике).

Отметим, что реальные объекты можно моделировать случайными множествами как из конечного числа элементов, так и из бесконечного, однако при расчетах на ЭВМ неизбежна дискретизация, т.е. переход к первой из названных возможностей.

Нечеткие множества. Пусть A - некоторое множество. Подмножество B множества A характеризуется своей характеристической функцией

$$\mu_B(x) = \begin{cases} 1, & x \in B, \\ 0, & x \notin B. \end{cases} \qquad (1)$$

Что такое нечеткое множество? Обычно говорят, что нечеткое подмножество *C* множества *A* характеризуется своей функцией принадлежности $\mu_C : A \to [0,1]$. Если функция принадлежности $\mu_C(x)$ имеет вид (1) при некотором *B*, то *C* есть обычное (четкое) подмножество *A*.

Обычное подмножество можно было бы отождествить с его характеристической функцией. Этого математики не делают, поскольку для задания функции (в ныне принятом подходе) необходимо сначала задать множество. Нечеткое же подмножество с формальной точки зрения можно отождествить с его функцией принадлежности. Однако термин "нечеткое подмножество" предпочтительнее при построении математических моделей реальных явлений.

Начало современной теории нечеткости положено работой 1965 г. американского ученого азербайджанского происхождения Л.А.Заде. К настоящему времени по этой теории опубликованы тысячи книг и статей, издается несколько международных журналов, выполнено достаточно много как теоретических, так и прикладных работ.

Л.А. Заде рассматривал теорию нечетких множеств как аппарат анализа и моделирования гуманистических систем, т.е. систем, в которых участвует человек. Его подход опирается на предпосылку о том, что элементами мышления человека являются не числа, а элементы некоторых нечетких множеств или классов объектов, для которых переход от "принадлежности" к "непринадлежности" не скачкообразен, а непрерывен. В настоящее время методы теории нечеткости используются почти во всех прикладных областях, в том числе при управлении предприятием, качеством продукции и технологическими процессами.

Л.А. Заде использовал термин "fuzzy set" (нечеткое множество). На русский язык термин "fuzzy" переводили как нечеткий, размытый, расплывчатый, и даже как пушистый и туманный.

Аппарат теории нечеткости громоздок. В качестве примера дадим определения теоретико-множественных операций над нечеткими множествами. Пусть *C* и *D*- два нечетких подмножества *A* с функциями принадлежности $\mu_C(x)$ и $\mu_D(x)$ соответственно. Пересечением $C \cap D$, произведением *CD*, объединением $C \cup D$, отрицанием $\overline{C}$, суммой *C*+*D* называются нечеткие подмножества *A* с функциями принадлежности

$$\mu_{C \cap D}(x) = \min(\mu_C(x), \mu_D(x)), \quad \mu_{CD}(x) = \mu_C(x)\mu_D(x), \quad \mu_{\overline{C}}(x) = 1 - \mu_C(x),$$

$$\mu_{C \cup D}(x) = \max(\mu_C(x), \mu_D(x)), \quad \mu_{C+D}(x) = \mu_C(x) + \mu_D(x) - \mu_C(x)\mu_D(x), \quad x \in A,$$

соответственно.

Теория нечетких множеств в определенном смысле сводится к теории вероятностей, а именно, к теории случайных множеств. Соответствующий цикл теорем приведен в книгах [3,10]. Однако при решении прикладных задач вероятностно-статистические методы и методы теории нечеткости обычно рассматриваются как различные.

Объекты нечисловой природы как статистические данные. В эконометрике и прикладной математической статистике наиболее распространенный объект изучения - выборка $x_1, x_2, \ldots, x_n$, т.е. совокупность результатов *n* наблюдений. В различных областях статистики результат наблюдения - это или число, или конечномерный вектор, или функция... Соответственно проводится деление прикладной математической статистики: одномерная статистика, многомерный статистический анализ, статистика временных рядов и случайных процессов... В статистике нечисловых данных в качестве результатов наблюдений рассматриваются объекты нечисловой природы, в частности, перечисленных выше видов - измерения в шкалах, отличных от абсолютной, бинарные отношения, вектора из 0 и 1, множества, нечеткие множества. Выборка может состоять из *n* ранжировок или *n* толерантностей, или *n* множеств, или *n* нечетких множеств и т.д.

Отметим необходимость развития методов статистической обработка "разнотипных данных", обусловленную большой ролью в прикладных исследованиях "признаков смешанной природы". Речь идет о том, что результат наблюдения состояния объекта зачастую представляет собой вектор, у которого часть координат измерена по шкале наименований, часть - по порядковой шкале, часть - по шкале интервалов и т.д. Статистические методы ориентированы обычно либо на абсолютную шкалу, либо на шкалу наименований (анализ таблиц сопряженности), а потому зачастую непригодны для обработки разнотипных данных. Есть и более сложные модели разнотипных данных, например, когда некоторые координаты вектора наблюдений описываются нечеткими множествами.

Для обозначения подобных неклассических результатов наблюдений в 1979 г. в монографии [3] предложен собирательный термин - объекты нечисловой природы. Термин "нечисловой" означает, что структура пространства, в котором лежат результаты наблюдений, не является структурой действительных чисел, векторов или функций, она вообще не является структурой линейного (векторного) пространства. При расчетах объекты числовой природы, разумеется, изображаются с помощью чисел, но эти числа нельзя складывать и умножать.

С целью "стандартизации математических орудий" целесообразно разрабатывать методы статистического анализа данных, пригодные одновременно для всех перечисленных выше видов результатов наблюдений. Кроме того, в процессе развития прикладных исследований выявляется необходимость использования новых видов объектов нечисловой природы, отличных от рассмотренных выше, например, в связи с развитием статистических методов обработки текстовой информации. Поэтому целесообразно ввести еще один вид объектов нечисловой природы - объекты произвольной природы, т.е. элементы множества, на которые не наложено никаких условий (кроме "условий регулярности", необходимых для справедливости доказываемых теорем). Другими словами, в этом случае предполагается, что результаты наблюдений (элементы выборки) лежат в произвольном пространстве *X*. Для получения теорем необходимо потребовать, чтобы *X* удовлетворяло некоторым условиям, например, было так называемым топологическим пространством. Как известно, ряд результатов классической математической статистики получен именно в такой постановке. Так, при изучении оценок максимального правдоподобия элементы выборки могут лежать в пространстве произвольной природы. Это не влияет на рассуждения, поскольку в них рассматривается лишь зависимость плотности вероятности от параметра. Методы классификации, использующие лишь расстояние между классифицируемыми объектами, могут применяться к совокупностям объектов произвольной природы, лишь бы в пространстве, где они лежат, была задана метрика. Цель статистики нечисловых данных (в некоторых литературных источниках используется термин "статистика объектов нечисловой природы") состоит в том, чтобы систематически рассматривать методы статистической обработки данных как произвольной природы, так и относящихся к указанным выше конкретным видам объектов нечисловой природы, т.е. методы описания данных, оценивания и проверки гипотез. Взгляд с общей точки зрения позволяет получить новые результаты и в других областях эконометрики.

Использование объектов нечисловой природы при формировании математической модели реального явления. Использование объектов нечисловой природы часто порождено желанием обрабатывать более объективную, более освобожденную от погрешностей информацию. Как показали многочисленные опыты, человек более правильно (и с меньшими затруднениями) отвечает на вопросы качественного например, сравнительного, характера, чем количественного. Так, ему легче сказать, какая из двух гирь тяжелее, чем указать их примерный вес в граммах. Другими словами, использование объектов нечисловой природы - средство повысить устойчивость эконометрических и экономико-математических моделей реальных явлений. Сначала конкретные области статистики объектов нечисловой природы (а именно, прикладная теория измерений, нечеткие и случайные множества) были рассмотрены в монографии [3] как частные постановки проблемы устойчивости математических моделей социально-экономических явлений и процессов к допустимым отклонениям исходных данных и предпосылок модели, а затем была понята необходимость проведения работ по развитию статистики объектов нечисловой природы как самостоятельного научного направления.

Начнем со шкал измерения. Науку о единстве мер и точности измерений называют метрологией. Таким образом, репрезентативная теория измерений - часть метрологии. Методы обработки данных должны быть адекватны относительно допустимых преобразований шкал измерения в смысле репрезентативной теории измерений. Однако установление типа шкалы, т.е. задание группы преобразований Ф - дело специалиста соответствующей прикладной области. Так, оценки привлекательности профессий мы считали измеренными в порядковой шкале. Однако отдельные социологи не соглашались с этим, считая, что выпускники школ пользуются шкалой с более узкой группой допустимых преобразований, например, интервальной шкалой. Очевидно, эта проблема относится не к математике, а к наукам о человеке. Для ее решения может быть поставлен достаточно трудоемкий эксперимент. Пока же он не поставлен, целесообразно принимать порядковую шкалу, так как это гарантирует от возможных ошибок.

Порядковые шкалы широко распространены не только в социально-экономических исследованиях. Они применяются в медицине - шкала стадий гипертонической болезни по Мясникову, шкала степеней сердечной недостаточности по Стражеско-Василенко-Лангу, шкала степени выраженности коронарной недостаточности по Фогельсону; в минералогии - шкала Мооса (тальк - 1, гипс - 2, кальций - 3, флюорит - 4, апатит - 5, ортоклаз - 6, кварц - 7, топаз - 8, корунд - 9, алмаз - 10), по которому минералы классифицируются согласно критерию твердости; в географии - бофортова шкала ветров ("штиль", "слабый ветер", "умеренный ветер" и др.) и т.д. Напомним, что по шкале интервалов измеряют величину потенциальной энергии или координату точки на прямой, на которой не отмечены ни начало, ни единица измерения; по шкале отношений - большинство физических единиц: массу тела, длину, заряд, а также цены в экономике. Время измеряется по шкале разностей, если год принимаем естественной единицей измерения, и по шкале интервалов в общем случае. В процессе развития соответствующей области знания тип шкалы может меняться.

Так, сначала температура измерялась по порядковой шкале (холоднее - теплее), затем - по интервальной (шкалы Цельсия, Фаренгейта, Реомюра) и, наконец, после открытия абсолютного нуля температур - по шкале отношений (шкала Кельвина). Следует отметить, что среди специалистов иногда имеются разногласия по поводу того, по каким шкалам следует считать измеренными те или иные реальные величины.

Отметим, что термин "репрезентативная" использовался, чтобы отличить рассматриваемый подход к теории измерений от классической метрологии, а также от работ А.Н.Колмогорова и А. Лебега, связанных с измерением геометрических величин, от "алгоритмической теории измерения" и др.

Необходимость использования в математических моделях реальных явлений таких объектов нечисловой природы, как бинарные отношения, множества, нечеткие множества, кратко была показана выше. Здесь же обратим внимание, что используемые в классической статистике результаты наблюдений также "не совсем числа". А именно, любая величина X измеряется всегда с некоторой погрешностью ΔX и результатом наблюдения является

$$Y = X + \Delta X.$$

Как уже отмечалось, погрешностями измерений занимается метрология. Отметим справедливость следующих фактов:

а) для большинства реальных измерений невозможно полностью исключить систематическую ошибку, т.е. $M(\Delta X) \neq 0$;

б) распределение ΔX в подавляющем большинстве случаев не является нормальным (см. главу 4);

в) измеряемую величину X и погрешность ее измерения ΔX обычно нельзя считать независимыми случайными величинами;

г) распределение погрешностей оценивается по результатам специальных наблюдений, следовательно, полностью известным считать его нельзя; зачастую исследователь располагает лишь границами для систематической погрешности и оценками таких характеристик для случайной погрешности, как дисперсия или размах.

Приведенные факты показывают ограниченность области применимости распространенной модели погрешностей, в которой X и ΔX рассматриваются как независимые случайные величины, причем ΔX имеет нормальное распределение с нулевым математическим ожиданием.

Строго говоря, результаты наблюдения всегда имеют дискретное распределение, поскольку описываются числами с небольшими (1 - 5) числом значащих цифр. Возникает дилемма: либо признать, что непрерывные распределения - фикция, и прекратить ими пользоваться, либо считать, что непрерывные распределения имеют "реальные" величины X, которые мы наблюдаем с принципиально неустранимой погрешностью ΔX. Первый выход в настоящее время нецелесообразен, так как потребует отказаться от большей части разработанного математического аппарата. Из второго следует необходимость изучения влияния неустранимых погрешностей на статистические выводы.

Погрешности ΔX можно учитывать либо с помощью вероятностной модели (ΔX - случайная величина, имеющая функцию распределения, вообще говоря, зависящую от X), либо с помощью нечетких множеств. Во втором случае приходим к теории нечетких чисел и к ее частному случаю - статистике интервальных данных (см. главу 9).

Другой источник появления погрешности ΔX связан с принятой в конструкторской и технологической документации системой допусков на контролируемые параметры изделий и деталей, с использованием шаблонов при проверке контроля качества продукции. В этих случаях характеристики ΔX определяются не свойствами средств измерения, а применяемой технологией проектирования и производства. В терминах математической статистики сказанному соответствует группировка данных, при которой мы знаем, какому из заданных интервалов принадлежит наблюдение, но не знаем точного значения результата наблюдения. Применение группировки может дать экономический эффект, поскольку зачастую легче (в среднем) установить, к какому интервалу относится результат наблюдения, чем точно измерить его.

Объекты нечисловой природы как результат статистической обработки данных. Объекты нечисловой природы появляются не только на "входе" статистической процедуры, но и в процессе обработки данных, и на "выходе" в качестве итога статистического анализа.

Рассмотрим простейшую прикладную постановку задачи регрессии (см. также главу 5). Исходные данные имеют вид $(x_i, y_i) \in R^2, i = 1,2,...,n.$. Цель состоит в том, чтобы с достаточной точностью описать *y* как полином от *x*, т.е. модель имеет вид

$$y_i = \sum_{k=0}^{m} a_j x_i^k + \varepsilon_i, \qquad (2)$$

где m - неизвестная степень полинома; $a_0, a_1, a_2, ..., a_m$ - неизвестные коэффициенты многочлена; $\varepsilon_i, i = 1,2,...,n$, - погрешности, которые для простоты примем независимыми и имеющими одно и то же нормальное распределение. (Здесь наглядно проявляется одна из причин живучести модель на основе нормального распределения. Такие модели, хотя и неадекватны реальной ситуации, с математической точки зрения позволяет проникнуть глубже в суть изучаемого явления. Поэтому они пригодны для первоначального анализа ситуации, как и в рассматриваемом случае. Дальнейшие научные исследования должны быть направлены на снятие нереалистического предположения нормальности и перехода к непараметрическим моделям погрешности.) Распространенная процедура такова: сначала пытаются применить модель (2) для линейной функции (m = 1), при неудаче (неадекватности модели) переходят к многочлену второго порядка (m = 2), если снова неудача, то берут модель (2) с m= 3 и т.д. (адекватность модели проверяют по F-критерию Фишера).

Обсудим свойства этой процедуры в терминах математической статистики. Если степень полинома задана ($m = m_0$), то его коэффициенты оценивают методом наименьших квадратов, свойства этих оценок хорошо известны (см., например, главу 5 или монографию [10, гл.26]). Однако в описанной выше реальной постановке m тоже является неизвестным параметром и подлежит оценке. Таким образом, требуется оценить объект (m, a_0, a_1, a_2, …, a_m), множество значений которого можно описать как $R^1 \cup R^2 \cup R^3 \cup ...$ Это - объект нечисловой природы, обычные методы оценивания для него неприменимы, так как m - дискретный параметр. В рассматриваемой постановке разработанные к настоящему времени методы оценивания степени полинома носят в основном эвристический характер (см., например, гл. 12 монографии [11]). Свойства описанной выше распространенной процедуры рассмотрены в главе 5; где показано, что m при этом оценивается несостоятельно, и найдено предельное распределение оценки этого параметра, оказавшееся геометрическим.

В более общем случае линейной регрессии данные имеют вид $(y_i, X_i), i = 1,2,...,n,$ где $X_i = (x_{i1}, x_{i2}, ..., x_{iN}) \in R^N$ - вектор предикторов (факторов, объясняющих переменных), а модель такова:

$$y_i = \sum_{j \in K} a_j x_{ij} + \varepsilon_i, i = 1,2,...,n \qquad (3)$$

(здесь K - некоторое подмножество множества {1,2,…,n}; ε_i - те же, что и в модели (2); a_j - неизвестные коэффициенты при предикторах с номерами из K). Модель (2) сводится к модели (3), если

$$x_{i1} = 1, x_{i1} = x_i, x_{i2} = x_i^2, x_{i3} = x_i^3, ..., x_{ij} = x_i^{j-1}, ...$$

В модели (2) есть естественный порядок ввода предикторов в рассмотрение - в соответствии с возрастанием степени, а в модели (3) естественного порядка нет, поэтому здесь стоит произвольное подмножество множества предикторов. Есть только частичный порядок - чем мощность подмножества меньше, тем лучше. Модель (3) особенно актуальна в задачах управления качеством продукции и других технико-экономических исследованиях, в экономике, маркетинге и социологии, когда из большого числа факторов, предположительно влияющих на изучаемую переменную, надо отобрать по возможности наименьшее число значимых факторов и с их помощью сконструировать прогнозирующую формулу (3).

Задача оценивания модели (3) разбивается на две последовательные задачи: оценивание множества K - подмножества множества всех предикторов, а затем - неизвестных параметров a_j. Методы решения второй задачи хорошо известны и подробно изучены. Гораздо хуже обстоит дело с оцениванием объекта нечисловой природы K. Как уже отмечалось, существующие методы - в основном эвристические, они зачастую не являются даже состоятельными. Даже само понятие состоятельности в данном случае требует специального определения. Пусть K_0 - истинное подмножество предикторов, т.е. подмножество, для которого справедлива модель (3), а подмножество предикторов K_n - его оценка. Оценка K_n называется состоятельной, если

$$\lim_{n \to \infty} Card(K_n \Delta K_0) = 0,$$

где Δ - символ симметрической разности множеств; *Card(K)* означает число элементов в множестве K, а предел понимается в смысле сходимости по вероятности.

Задача оценивания в моделях регрессии, таким образом, разбивается на две - оценивание структуры модели и оценивание параметров при заданной структуре. В модели (2) структура описывается неотрицательным целым числом m, в модели (3) - множеством K. Структура - объект нечисловой природы. Задача ее оценивания сложна, в то время как задача оценивания численных параметров при заданной структуре хорошо изучена, разработаны эффективные (в смысле математической статистики) методы.

Такова же ситуация и в других методах многомерного статистического анализа - в факторном анализе (включая метод главных компонент) и в многомерном шкалировании. Ряд иных примеров можно найти в списке оптимизационных постановок основных проблем прикладного многомерного статистического анализа, приведенном в монографии [12].

Перейдем к объектам нечисловой природы на "выходе" статистической процедуры. Примеры многочисленны. Разбиения - итог работы многих алгоритмов классификации, в частности, алгоритмов кластер-анализа. Ранжировки - результат упорядочения профессий по привлекательности или автоматизированной обработки мнений экспертов - членов комиссии по подведению итогов конкурса научных работ. (В последнем случае используются ранжировки со связями; так, в одну группу, наиболее многочисленную, попадают работы, не получившие наград.) Из всех объектов нечисловой природы, видимо, наиболее часты на "выходе" дихотомические данные - принять или не принять гипотезу, в частности, принять или забраковать партию продукции. Результатом статистической обработка данных может быть множество, например зона наибольшего поражения при аварии, или последовательность множеств, например, "среднемерное" описание распространения пожара (см. главу 4 в монографии [3]). Нечетким множеством Э. Борель [13] еще в начале XX в. предлагал описывать представление людей о числе зерен, образующем "кучу". С помощью нечетких множеств формализуются значения лингвистических переменных, выступающих как итоговая оценка качества систем автоматизированного проектирования, сельскохозяйственных машин, бытовых газовых плит, надежности программного обеспечения или систем управления. Можно констатировать, что все виды объектов нечисловой природы могут появляться " на выходе" статистического исследования.

8.2. Вероятностные модели конкретных видов объектов нечисловой природы

В настоящем пункте рассмотрены основные вероятностные модели объектов нечисловой природы: дихотомических данных, результатов парных сравнений, бинарных отношений, рангов, объектов общей природы. Обсуждаются различные варианты вероятностных моделей, приведены краткие сведения об их практическом использовании (см. также обзор [14]).

Дихотомические данные. Рассмотрим базовую вероятностную модель дихотомических данных - бернуллиевский вектор (в терминологии энциклопедии [15] - люсиан), т.е. конечную последовательность $X = (X_1, X_2, ..., X_k)$ независимых испытаний Бернулли X_i, для которых $P(X_i = 1) = p_i$ и $P(X_i = 0) = p_i$, $i = 1,2,...,k$, причем вероятности p_i могут быть различны.

Бернуллиевские вектора часто применяются при практическом использовании эконометрических методов. Так, они использованы в монографии [3] для описания равномерно распределенных случайных толерантностей. Как известно, толерантность на множестве из m элементов можно задать симметричной матрицей $\|\delta_{ij}\|$ из 0 и 1, на главной диагонали которой стоят 1. Тогда случайная толерантность описывается распределением $m(m-1)/2$ дихотомических случайных величин δ_{ij}, $1 \le i \prec j \le m$, а для равномерно распределенной (на множестве всех толерантностей) толерантности эти случайные величины, как можно доказать, оказываются независимыми и принимают значения 0 и 1 с равными вероятностями 1/2. Записав элементы δ_{ij} задающей такую толерантность матрицы в строку, получим бернуллиевский вектор с $k=m(m-1)/2$ и $p_i = 1/2$, $i = 1,2,...,k$.

В связи с оцениванием по статистическим данным функции принадлежности нечеткой толерантности в 1970-е годы была построена теория случайных толерантностей с такими независимыми δ_{ij}, что вероятности $P(\delta_{ij} = 1) = p_{ij}$ произвольны (см. об этом монографию [3]).

Случайные множества с независимыми элементами использовались как общий язык для описания парных сравнений и случайных толерантностей. В статьях [16] и [17] термин "люсиан" применялся как сокращение для выражения "случайные множества с независимыми элементами". В работе [18], являющейся продолжением [17] и содержащей описание расчетных методов, вытекающих из результатов [17], этот термин не употреблялся вообще, хотя указанный объект (т.е. бернуллиевский вектор) был основным предметом изучения. Это объясняется тем, что изложение в работе [18] шло на языке обработки результатов парных сравнений, которые для прикладника никак не связаны с множествами.

В дальнейшем был выявлен ещё ряд областей, в которых может оказаться полезным разработанный математический аппарат решения различных эконометрических задач, связанных с бернуллиевскими векторами. Перечислим эти области, включая ранее названные: анализ случайных толерантностей; случайные множества с независимыми элементами; обработка результатов независимых парных сравнений; статистические методы анализа точности и стабильности технологического процесса, а также анализ и синтез планов статистического приемочного контроля (по альтернативным, т.е. дихотомическим, признакам); обработка маркетинговых и социологических анкет (с закрытыми вопросами типа "да"-"нет"); обработка социально-психологических и медицинских данных, в частности, ответов на психологические тесты типа MMPI (используемых в задачах управления персоналом), топографических карт (применяемых для анализа и прогноза зон поражения при технологических авариях, распространении коррозии, распространении экологически вредных загрязнений в других ситуациях) и т.д.

Теорию бернуллиевских векторов можно выразить в терминах любой из этих теоретических и прикладных областей. Однако терминология одной из этих областей "режет слух" и приводит к недоразумениям в другой из них. Поэтому мы считаем целесообразным использовать термины "бернуллиевский вектор" в указанном выше значении, не связанном ни с какой из перечисленных областей приложения этой теории (в ряде публикаций в том же значении использовался термин "люсиан").

Распределение бернуллиевского вектора *X* полностью описывается вектором $P = (p_1, p_2, ..., p_k)$,т.е. нечетким подмножеством множества *{1,2,...,k}*. Действительно, для любого детерминированного вектора $x = (x_1, x_2, ..., x_k)$ из 0 и 1 имеем

$$P(X = x) = \prod_{1 \le j \le k} h(x_j, p_j),$$

где *h(x,p)=p* при *x*=1 и *h(x,p)=1-p* при *x*=0.

Теперь можно уточнить способы использования люсианов при эконометрическом моделировании. Бернуллиевскими векторами можно моделировать: результаты статистического контроля (0-годное изделие, 1-дефектное); результаты маркетинговых и социологических опросов (0-опрашиваемый выбрал первую из двух подсказок, 1-вторую); распределение посторонних включений в материале (0 - нет включения в определенном объеме материала, 1 - есть); результаты испытаний и анализов (0 - нет нарушений требований нормативно-технической документации, 1 - есть такие нарушения); процессы распространения, например, пожаров (0 - нет загорания, 1 - есть; подробнее см. [3, с.215-223]); технологические процессы (0 - процесс находится в границах допуска,1 - вышел из них); ответы экспертов (опрашиваемых) о сходстве объектов (проектов, образцов) и т.д.

Парные сравнения. Общую модель парных сравнений опишем согласно монографии Г. Дэвида [9, с.9]. Предположим, что *t* объектов $A_1, A_2, ..., A_t$ сравниваются попарно каждым из *n* экспертов. Всего возможных пар для сравнения имеется $s = t(t-1)/2$. Эксперт с номером γ делает r_γ повторных сравнений для каждой из *s* возможностей. Пусть $X(i, j, \gamma, \delta)$, *i,j*=1,2,...,*t,* $i \neq j$, γ=1,2,...,*n*; δ=1,2,..., r_γ-случайная величина, принимающая значение 1 или 0 в зависимости от того, предпочитает ли эксперт γ объект A_i или объект A_j в δ-м сравнении двух объектов. Предполагается, что все сравнения проводятся независимо друг от друга, так что случайные величины $X(i,j,\gamma,\delta)$ независимы в совокупности, если не считать того, что $X(i,j,\gamma,\delta)+X(j,i,\gamma,\delta)=1$. Положим

$$P(X(i,j,\gamma,\delta)=1)=\pi(i,j,\gamma,\delta).$$

Ясно, что описанная эконометрическая модель парных сравнений представляет собой частный случай бернуллиевского вектора. В этой модели число наблюдений равно числу неизвестных параметров, поэтому для получения статистических выводов необходимо положить априорные условия на $\pi(i,j,\gamma,\delta)$, например [9, с.9]:

$\pi(i,j,\gamma,\delta)=\pi(i,j,\gamma)$ (нет эффекта от повторений);

$\pi(i,j,\gamma,\delta)=\pi(i,j)$ (нет эффекта от повторений и от экспертов).

Теорию независимых парных сравнений целесообразно разделить на две части - непараметрическую, в которой статистические задачи ставятся непосредственно в терминах $\pi(i,j,\gamma,\delta)$, и параметрическую, в которой вероятности $\pi(i,j,\gamma,\delta)$ выражаются через меньшее число иных параметров. Ряд результатов непараметрической теории парных сравнений непосредственно вытекает из теории бернуллиевских векторов.

В параметрической теории парных сравнений наиболее популярна так называемая линейная модель [9, с.11], в которой предполагается , что каждому объекту A_i можно сопоставить некоторую "ценность" V_i так, что вероятность предпочтения $\pi(i,j)$ (т.е. предполагается дополнительно, что эффект от повторений и от экспертов отсутствует) выражается следующим образом:

$$\pi(i,j)=H(V_i-V_j), \qquad (1)$$

где $H(x)$ - функция распределения, симметричная относительно 0, т.е.

$$H(-x)=1-H(x) \qquad (2)$$

при всех x.

Широко применяются модели Терстоуна - Мостеллера и Брэдли - Терри , в которых $H(x)$ - соответственно функции нормального и логистического распределений. Поскольку функция $\Phi(x)$ стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1 и функция

$$\Psi(x)=e^x(1+e^x)^{-1}$$

стандартного логистического распределения удовлетворяют (см., например, [19]) соотношению

$$\sup_{x\in R^1}|\Phi(x)-\Psi(1{,}7x)|<0{,}01,$$

то для обоснованного выбора по статистическим данным между моделями Терстоуна-Мостеллера и Брэдли-Терри необходимо не менее тысячи наблюдений (ср. п.4.2 выше).

Соотношение (1) вытекает из следующей модели поведения эксперта: он измерят "ценность" V_i и V_j объектов A_i и A_j, но с ошибками ε_i и ε_j соответственно, а затем сравнивает свои оценки ценности объектов $y_i=V_i+\varepsilon_i$ и $y_j=V_j+\varepsilon_j$. Если $y_i>y_j$, то он предпочитает A_i, в противном случае - A_j. Тогда

$$\pi(i,j)=P(\varepsilon_i-\varepsilon_j<V_i-V_j)=H(V_i-V_j). \qquad (3)$$

Обычно предполагают, что субъективные ошибки эксперта ε_i и ε_j независимы и имеют одно и то же непрерывное распределение. Тогда функция распределения $H(x)$ из соотношения (3) непрерывна и удовлетворяет функциональному уравнению (2).

Существует много разновидностей моделей парных сравнений, постоянно предполагаются новые. В качестве примера опишем модель парных сравнений, основанную не на процедуре упорядочения, а на определении сходства объектов. Пусть каждому объекту A_i соответствует точка a_i в r-мерном евклидовом пространстве R^r. Эксперт "измеряет" a_i и a_j с ошибками ε_i и ε_j соответственно и в случае, если евклидово расстояние между $a_i+\varepsilon_i$ и $a_j+\varepsilon_j$ меньше 1, заявляет о сходстве объектов A_i и A_j, в противном случае - об их различии. Предполагается, что ошибки ε_i и ε_j независимы и имеют одно и то же распределение, например, круговое нормальное распределение с нулевым математическим ожиданием и дисперсией координат σ^2. Целью статистической обработки является определение по результатам парных сравнений оценок параметров $a_1, a_2,\dots,a_r$, и σ^2, а также проверка согласия опытных данных с моделью.

Рассмотренные модели парных сравнений могут быть обобщены в различных направлениях. Так, можно ввести понятие "ничья "- ситуации, когда эксперт оценивает объекты одинаково. Модели с учетом "ничьих" предполагают, что эксперт может отказаться от выбора одного из объектов и заявить об их эквивалентности, т. е. число возможных ответов увеличивается с 2 до 3. В моделях множественных сравнений эксперту представляется не два объекта , а три или большее число

Модели, учитывающие "ничьи", строятся обычно с помощью используемых в психофизике "порогов чувствительности": если $|y_i-y_j|\le r$ (где r- порог чувствительности), то объекты A_i и A_j эксперт объявляет неразличимыми. Приведем пример модели с "ничьими", основанной на другом принципе. Пусть каждому объекту A_i соответствует точка a_i в r-мерном линейном пространстве. Как и прежде , эксперт "измеряет " объектные точки " a_i и a_j с ошибками ε_i и ε_j соответственно, т.е. принимает решение на основе $y_i = a_i+\varepsilon_i$ и $y_j = a_j+\varepsilon_j$. Если все координаты y_i больше соответствующих координат y_j , то A_i предпочитается A_j. Соответственно, если каждая координата y_i меньше координаты y_j с тем же номером , то эксперт считает наилучшим объект A_j. Во всех остальных случаях эксперт объявляет о ничейной ситуации. Эта модель при r=1 переходит в описанную выше линейную модель. Она связана с принципом Парето в теории группового выбора и предусматривает выбор оптимального по Парето объекта, если он существует (роль согласуемых критериев играют процедуры сравнения значений отдельных координат), и отказ от выбора, если такого объекта нет.

Можно строить модели, учитывающие порядок предъявления объектов при сравнении, зависимость результата сравнения от результатов предшествующих сравнений. Опишем одну из подобных моделей.

Пусть эксперт сравнивает три объекта - *A, B, C,* причем сначала сравниваются *A* и *B*, потом - *B* и *C* и, наконец, *A* и *C*. Для определенности пусть *A*>*B* будет означать, что *A* более предпочтителен, чем *B*. Пусть при предъявлении двух объектов

$$P(A>B)=\pi_{AB}, P(B>C)=\pi_{BC}, P(A>C)=\pi_{AC}.$$

Теперь пусть пара *B, C* предъявляется после пары *A, B*. Естественно предположить, что высокая оценка *B* в первом сравнении повышает вероятность предпочтения *B* и во втором, и, наоборот, отрицательное мнение о *B* в первом сравнении сохраняется и при проведении второго сравнения. Это предположение проще всего учесть в модели следующим образом:

$$P(B > C \mid B > A) = \pi_{BC} + \delta, \quad P(B > C \mid A > B) = \pi_{BC} - \delta,$$

где δ - некоторое положительное число, показывающее степень влияния первого сравнения на второе. По аналогичным причинам вероятности исхода третьего сравнения в зависимости от результатов первых двух можно описать так:

$$P(A > C \mid A > B, B > C) = \pi_{AC} + 2\delta, \quad P(A > C \mid A > B, B < C) = \pi_{AC},$$

$$P(A > C \mid A < B, B > C) = \pi_{AC}, \quad P(A > C \mid A < B, B < C) = \pi_{AC} - 2\delta.$$

Статистическая задача состоит в определении параметров π_{AB}, π_{BC}, π_{AC} и δ по результатам сравнений, проведенных n экспертами, и в проверке адекватности модели.

Ясно, что можно рассматривать и другие модели, в частности, учитывающие тягу экспертов к транзитивности ответов. Очевидно, что проблемы построения моделей парных сравнений относятся не к эконометрической теории, а к тем прикладным областям, для решения задач которых развиваются методы парных сравнений, например, к экономике предприятия, стратегическому менеджменту, производственной психологии, изучению поведения потребителей, экспертным оценкам и т. д.

Метод парных сравнений был введен в 1860 г. Г. Т. Фехнером для решения задач психофизики. Расскажем об этом несколько подробнее. Как известно, основателем психофизики по праву считается Густав Теодор Фехнер (1801 - 1887), а год выхода в свет его фундаментальной работы "Элементы психофизики"(1860) - датой рождения новой науки; в этой работе широко применялся предложенный Г.Т. Фехнером метод парных сравнений (обсуждение событий тех лет с современных позиций дано в монографии [9, с.14-16]).

С точки зрения математической статистики приведенные выше модели не представляют большого теоретического интереса: оценки параметров находятся обычно методом максимального правдоподобия, а проверка согласия проводится по критерию отношения правдоподобия или асимптотически эквивалентными ему критериями типа хи-квадрат [9]. Вычислительные процедуры обычно сложны и плохо исследованы; их можно упростить и одновременно повысить обоснованность, перейдя от оценок максимального правдоподобия к одношаговым оценкам [20].

Отметим некоторые сложности при обосновании возможности использовании линейных моделей типа (1) - (3). Эконометрическая теория достаточно проста, когда предполагается , что каждому отдельному сравнению двух объектов соответствуют свои собственные ошибки экспертов, причем все ошибки независимы в совокупности. Однако это предположение отнюдь не очевидно с содержательной точки зрения. В качестве примера рассмотрим три объекта *A, B* и *C*, которые сравнивают попарно: *A* и *B*, *B* и *C*, *A* и *C*. В соответствии со сказанным, в рассмотрение вводят 6 ошибок одного и того же эксперта: ε_A и ε_B в первом сравнении, ε_B' и ε_C'-во втором, ε_A' и ε_C'- в третьем, причем все эти 6 случайных величин независимы в совокупности. Между тем естественно думать, что мнения эксперта об одном и том же объекте связаны между собой, т. е. ε_A и ε_A' зависимы, равно как ε_B и ε_B', а также ε_C и ε_C'. Более того, если принять, что точка зрения эксперта полностью определена для него самого, то следует положить $\varepsilon_A = \varepsilon_A'$ и соответственно $\varepsilon_B = \varepsilon_B'$ и $\varepsilon_C = \varepsilon_C'$. При этом, напомним, случайные величины ε_A, ε_B и др. интерпретируется как отклонения мнений отдельных экспертов от истины. Видимо, ошибку эксперта целесообразно считать состоящей из двух слагаемых, а именно: отклонения от истины, вызванного внутренними особенностями эксперта (систематическая погрешность) и колебания мнения эксперта в связи с очередным парным сравнением (случайная погрешность). Игнорирование систематической погрешности облегчает развитие математико-статистической теории, а ее учет приводит к необходимости изучения зависимых парных сравнений.

При обработке результатов парных сравнений первый этап - проверка согласованности. Понятие согласованности уточняется различными способами, но все они имеют один и тот же смысл проверки однородности обрабатываемого материала, т.е. того, что целесообразно агрегировать мнения отдельных экспертов, объединить данные и совместно их обрабатывать. При отсутствии однородности данные разбиваются на группы (классы, кластеры, таксоны) с целью обеспечения однородности внутри отдельных групп. Естественно, согласованность целесообразно проверять, вводя возможно меньше гипотез о структуре данных. Следовательно, целесообразно пользоваться для этого непараметрической теорией парных сравнений, основанной на теории бернуллиевских векторов.

Хорошо известно, что модели парных сравнений можно с успехом применять в экспертных и экспериментальных процедурах упорядочивания и выбора, в частности, для анализа голосований, турниров, выбора наилучшего объекта (проекта, образца, кандидатуры); в планировании и анализе сравнительных экспериментов и испытаний; в органолептической экспертизе (в частности, дегустации); при изучении поведения потребителей; визуальной колоритмии, определении индивидуальных рейтингов и вообще изучении предпочтений при выборе и т. д. (подробнее см. [3,9]).

Бинарные отношения. Теорию ранговой корреляции [6, 21] можно рассматривать как теорию статистического анализа случайных ранжировок, равномерно распределенных на множестве всех ранжировок. Так, при обработке данных классического психофизического эксперимента по упорядочению кубиков соответственно их весу, подробно описанного в работе [22], оказалась адекватной следующая т.н. *T*-модель ранжирования.

Пусть имеется t объектов $A_1, A_2, ..., A_t$, причем каждому объекту A_i соответствует число a_i, описывающее его положение на шкале изучаемого признака. Испытуемый упорядочивает объекты так, как если бы оценивал соответствующие им значения с ошибками, т.е. находил $y_i = a_i + \varepsilon_i$, i=1,2,…n, где ε_i - ошибка при рассмотрении i-го объекта, а затем располагал бы объекты в том порядке, в каком располагаются $y_1, y_2, ..., y_t$. В этом случае вероятность появления упорядочения $A_{i1}, A_{i2}, ...A_{it}$ есть $P(y_{i1} < (y_{i2} < ... < y_{it})$, а ранги $R_1, R_2, ...R_t$ объектов являются рангами случайных величин $y_1, y_2, ..., y_t.$, полученных при их упорядочении в порядке возрастания. Кроме того, для простоты расчетов в модели предполагается, что ошибки испытуемого $\varepsilon_1, \varepsilon_2, ..., \varepsilon_t$ независимы и имеют нормальное распределение с математическим ожиданием 0 и дисперсией σ^2. Как уже отмечалось, бинарное отношение на множестве из t элементов полностью описывается матрицей из 0 и 1 порядка $t \times t$. Поэтому задать распределение случайного бинарного отношения - это то же самое, что задать распределение вероятностей на множестве всех матриц описанного вида, состоящем из $2^{(t^2)}$ элементов. Пространства ранжировок, разбиений, толерантностей зачастую удобно считать подпространствами пространства всех бинарных отношений, тогда распределения вероятностей на них - частные случаи описанного выше распределения, выделенные тем, что вероятности принадлежности соответствующим подпространствам равны 1. Распределение произвольного бинарного отношения описывается $2^{(t^2)}$-1 параметрами, распределение случайной ранжировки (без связей) - *(t!-1)* параметрами, а описанная выше *Т*-модель ранжирования - *(t+1)* параметром. При *t*=4 эти числа равны соответственно 65535, 23 и 5. Первое из этих чисел показывает практическую невозможность использования в эконометрических моделях произвольных бинарных отношений, поскольку по имеющимся данным невозможно оценить столь большое число параметров. Приходится ограничиваться теми или иными семействами бинарных отношений - ранжировками, разбиениями, толерантностями и др. Модель произвольной случайной ранжировки при *t*=5 описывается 119 параметрами, при *t*=6 - уже 719 параметрами, при *t*=7 число параметром достигает 5049, что уже явно за возможностями оценивания. В то же время *Т*-модель ранжирования при *t*=7 описывается всего 8-ю параметрами, а потому она практически пригодна.

Что естественно предположить относительно распределения случайного элемента со значениями в том или ином пространстве бинарных отношений? Зачастую целесообразно считать, что распределение имеет некий центр, попадание в который наиболее вероятно, а по мере удаления от центра вероятности убывают. Это соответствует естественной модели измерения с ошибкой; в классическом одномерном случае результат подобного измерения описывается унимодальной симметричной плотностью, монотонно возрастающей слева от модального значения, в котором плотность максимальна, и монотонно убывающей справа от него. Чтобы ввести понятие монотонного распределения в пространстве бинарных отношений, будем исходить из метрики в этом пространстве. Воспользовавшись тем, что бинарные отношения *C* и *D* однозначно описываются матрицами $\| c_{ij} \|$ и $\| d_{ij} \|$ порядка $t \times t$ соответственно, рассмотрим расстояние (в несколько другой терминологии - метрику) в пространстве бинарных отношений

$$d(C,D) = \sum_{1 \le i,j \le t} | c_{ij} - d_{ij} |. \qquad (4)$$

Метрика (4) в различных пространствах бинарных отношений - ранжировок, разбиений, толерантностей - может быть введена с помощью соответствующих систем аксиом. В работах [3, 23] дан обзор аксиоматическим подходам к получению метрики (4) в различных пространствах объектов нечисловой природы. В настоящее время метрику (4) обычно называют расстоянием Кемени в честь американского исследователя Джона Кемени, впервые получившего эту метрику исходя из предложенной им системы аксиом для расстояния между упорядочениями (ранжировками). Этой тематике посвящена первая глава учебника [24], на английском языке выпущенном под названием "Математические методы в социальных науках".

В статистике нечисловых данных используются и иные метрики, отличающиеся от расстояния Кемени. Более того, для использования понятия монотонного распределения, о котором сейчас идет речь, нет необходимости требовать выполнения неравенства треугольника, а достаточно, чтобы *d(C,D)* можно было рассматривать как показатель различия. Под показателем различия понимаем такую функцию *d(C,D)* двух бинарных отношений *C* и *D*, что *d(C,D)*=0 при *C=D* и увеличение *d(C,D)* интерпретируется как возрастание различия между *C* и *D*.

Определение 1. Распределение бинарного отношения *X* называется монотонным относительно расстояния (показателя различия) *d* с центром в C_0, если из $d(C,C_0)<d(D,C_0)$ следует, что $P(X=C)>P(X=D)$

Это определение впервые введено в монографии [3, с.196]. Оно может использоваться в любых пространствах бинарных отношений и, более того, в любых пространствах из конечного числа элементов, лишь бы в них была введена функция $d(C,D)$ - показатель различия элементов C и D этого пространства. Монотонное распределение унимодально, мода находится в C_0.

Определение 2. Распределение бинарного отношения X называется симметричным относительно расстояния d с центром в C_0, если существует такая функция $f: R^1_+ \to [0,1]$, что

$$P(X = C) = f(d(C, C_0)). \qquad (5)$$

Если распределение X монотонно и таково, что из $d(C,C_0) = d(D,C_0)$ следует $P(X=C) = P(X=D)$, то оно симметрично. Если функция f в формуле (5) монотонно строго убывает, то соответствующее распределение монотонно в смысле определения 1.

Поскольку толерантность на множестве из t элементов задается $0{,}5t(t-1)$ элементами матрицы из 0 и 1 порядка $t \times t$, лежащими выше главной диагонали, то распределение на множестве толерантностей задается в общем случае $2^{0,5t(t-1)}$ параметрами. Естественно выделить семейство распределений, соответствующее независимым элементам матрицы. Оно задается бернуллиевским вектором (люсианом) с $0{,}5t(t-1)$ параметрами (выше бернуллиевские вектора рассмотрены подробнее). Математическая техника, необходимая для изучения толерантностей с независимыми элементами, существенно проще, чем в случае ранжировок и разбиений. Здесь легко отказаться от условия равномерности распределения. Этому условию соответствует $p_{ij} = 1/2$, в то время как статистические методы анализа люсианов, развитые в статистике нечисловых данных (см., например, работы [3,17, 18]) не налагают никаких существенных ограничений на p_{ij} .

Как уже отмечалось, при обработке мнений экспертов сначала проверяют согласованность. В частности, если мнения экспертов описываются монотонными распределениями, то для согласованности необходимо совпадение центров этих распределений. К сожалению, рассмотренные выше классические методы проверки согласованности для ранжировок, основанные на коэффициентах ранговой корреляции и конкордации, позволяют лишь отвергнуть гипотезу о равнораспределенности, но не установить, можно ли считать, что центры соответствующих экспертам распределений совпадают или же, например, существует две группы экспертов, каждая со своим центром. Теория случайных толерантностей лишена этого недостатка. Отсюда вытекают следующие практические рекомендации.

Пусть цель обработки экспертных данных состоит в получении ранжировки, отражающей групповое мнение. Однако согласно рекомендуемой процедуре экспертного опроса пусть эксперты не упорядочивают объекты, а проводят парные сравнения, сравнивая каждый из рассматриваемых объектов со всеми остальными, причем ровно один раз. Когда ответ эксперта - толерантность, но, вообще говоря, не ранжировка, поскольку в ответах эксперта может нарушаться транзитивность.

Возможны два пути обработки данных. Первый - превратить ответ эксперта в ранжировку (тем или иным способом "спроектировав" на пространство ранжировок), а затем проверять согласованность ранжировок с помощью известных критериев. При этом от толерантности перейти к ранжировке можно, например, так. Будем выбирать ближайшую (в смысле применяемого расстояния) матрицу к матрице ответов эксперта из всех, соответствующих ранжировкам без связей.

Второй путь - проверить согласованность случайных толерантностей, а групповое мнение искать с помощью медианы Кемени (см. ниже) непосредственно по исходным данным, т.е. по толерантностям. Групповое мнение при этом может быть найдено в пространстве ранжировок. Второй путь мы считаем более предпочтительным, поскольку при этом обеспечивается более адекватная проверка согласованности и исключается процедура укладывания мнения эксперта в "прокрустово ложе "ранжировки" (эта процедура может приводить как к потере информации, так и к принципиально неверным выводам).

Области применения статистики бинарных отношений многообразны: ранговая корреляция - оценка величины связи между переменными, измеренными в порядковой шкале; анализ экспертных или экспериментальных упорядочений; анализ разбиений технико-экономических показателей на группы сходных между собой; обработка данных о сходстве (взаимозаменяемости); статистический анализ классификаций; математические вопросы теории менеджмента и др.

Случайные множества. Будем рассматривать случайные подмножества некоторого множества Q. Если Q состоит из конечного числа элементов, то считаем, что случайное подмножество S - это случайный элемент со значениями в 2^Q - множестве всех подмножеств множества Q, состоящем из $2^{card(Q)}$ элементов. Чтобы удовлетворить математиков, считаем, что все подмножества Q измеримы. Тогда распределение случайного подмножества $S = S(\omega)$ множества Q - это

$$P_S(A) = P(S = A) = \mathrm{P}(\{\omega : S(\omega) = A\}), A \subseteq Q. \qquad (6)$$

В формуле (6) предполагается, что $S : \Omega \to 2^Q$, где (Ω, F, P) - вероятностное пространство (здесь Ω - пространство элементарных событий, F- σ-алгебра случайных событий, P-вероятностная мера на F), на котором определен случайный элемент $S(\omega)$.Через распределение $P_S(A)$ выражаются вероятности различных событий, связанных с S. Так ,чтобы найти вероятность накрытия фиксированного элемента q случайным множеством S, достаточно вычислить

$$P(q \in S) = \mathrm{P}(\{\omega : q \in S(\omega)\}) = \sum_{A: q\in A, A\subseteq 2^Q} P(S = A),$$

где суммирование идет по всем подмножествам A множества Q, содержащим q. Пусть $Q=\{q_1, q_2,...,q_k\}$. Рассмотрим случайные величины, определяемые по случайному множеству S следующим образом

$$\chi_i(\omega) = \begin{cases} 1, q_i \in S(\omega), \\ 0, q_i \notin S(\omega). \end{cases}$$

Определение 3. Случайное множество S называется случайным множеством с независимыми элементами, если случайные величины $\chi_i(\omega), i = 1,2,...,k,$ независимы (в совокупности).

Последовательность случайных величин $\chi_1, \chi_2,..., \chi_k$ --бернуллиевский вектор с $X_i = \chi_i$ и $p_i = P(q_i \in S(\omega)), i = 1,2,...,k.$ Из последней формулы подпункта "Дихотомические данные" следует, что распределение случайного множества с независимыми элементами задается формулой

$$P(S = A) = \prod_{q_i\in A} p_i \prod_{q_i \in Q\setminus A}(1 - p_i),$$

т.е. такие распределения образуют k = $card(Q)$ - мерное параметрическое семейство, входящее в $(2^{card(Q)} - 1)$ - одномерное семейство всех распределений случайных подмножеств множества Q.

При исследовании случайных подмножеств произвольного множества Q будем рассматривать их как случайные величины со значениями в некотором пространстве подмножеств множества Q, например, в пространстве замкнутых подмножеств 2^Q множества Q. Представляющими интерес лишь для математиков способами введения измеримой структуры в 2^Q интересоваться не будем. Отсутствие специального интереса к проблеме измеримости связано с тем, что при эконометрическом моделировании и обработке на ЭВМ все случайные подмножества рассматриваются как конечные (т.е. подмножества конечного множества).

Случайные множества находят разнообразные применения в многообразных проблемах эконометрики и математической экономики, в том числе в задачах управлении запасами и ресурсами (см. об этом главу 5 в монографии [3]), в задачах менеджмента и маркетинга, в экспертных оценках, в частности, при анализе мнений голосующих или опрашиваемых, каждый из которых отмечает несколько пунктов из списка и т.д. Кроме того, случайные множества применяются в гранулометрии, при изучении пористых сред и объектов сложной природы в таких областях, как металлография, петрография, биология, в частности, математическая морфология, в изучении структуры веществ и материалов, в исследовании процессов распространения, в частности, просачивания, распространения пожаров, экологических загрязнений, при районировании, в том числе в изучении областей поражения, в частности, поражения металла коррозией и сердечной мышцы при инфаркте миокарда, и т.д., и т.п. Можно вспомнить о компьютерной томографии, о наглядном представлении сложной информации на экране компьютера, об изучении распространения рекламной информации, о картах Кохонена (популярный метод представления информации при применении нейросетей) и т.д.

Ранговые методы. Ранее установлено, что любой адекватный алгоритм в порядковой шкале является функцией от некоторой матрицы C. Пусть никакие два из результатов наблюдений $x_1, x_2,\dots,x_n$ не совпадают, а $r_1, r_2,\dots,r_n$ - их ранги. Тогда элементы матрицы C и ранги результатов наблюдений связаны взаимно однозначным соответствием:

$$r_i = 1 + \sum_{1\le j\le n}(1 - c_{ij}),$$

а c_{ij} через ранги выражаются так: $c_{ij}=1$, если $r_i<r_j$, и $c_{ij}=0$ в противном случае.

Сказанное означает, что при обработке данных, измеренных в порядковой шкале, могут применяться только ранговые статистические методы. Отметим, что часто используемое в непараметрической статистике преобразование $Y=F(x)$ (здесь $F(x)$ - непрерывная функция распределения случайной величины X, причем F предполагается произвольной) фактически означает переход к порядковой шкале, поскольку статистические выводы при этом инвариантны относительно допустимых преобразований в порядковой шкале.

Разумеется, ранговые статистические методы могут применяться не только при обработке данных, измеренных в порядковой шкале. Так, для проверки независимости двух количественных признаков в случае, когда нет уверенности в нормальности соответствующего двумерного распределения, целесообразно пользоваться коэффициентами ранговой корреляции Кендалла или Спирмена.

Как было подробно обосновано в главах 4 и 5, в настоящее время с помощью непараметрических и прежде всего ранговых методов можно решать все те задачи эконометрики и прикладной статистики, что и с помощью параметрических методов, в частности, основанных на предположении нормальности. Однако параметрические методы вошли в массовое сознание исследователей и инженеров и мешают широкому внедрению более обоснованной и прогрессивной ранговой статистики. Так, при проверке однородности двух выборок вместо критерия Стьюдента целесообразно использовать ранговые методы, но пока это делается редко.

Объекты общей природы. Вероятностная модель объекта нечисловой природы в общем случае- случайный элемент со значениями в пространстве произвольного вида, а модель выборки таких объектов - совокупность независимых одинаково распределенных случайных элементов. Именно такая модель была использована для обработки наблюдений, каждое из которых - нечеткое множество [10].

Из-за имеющего разнобоя в терминологии приведем математические определения из справочника по теории вероятностей академика РАН Ю.В Прохорова и проф. Ю.А. Розанова [25].

Пусть (X,B)-некоторое измеримое пространство; (F,B)-измеримая функция $\xi=\xi(\omega)$ на пространстве элементарных событий (Ω,F,P) (где P - вероятностная мера на σ-алгебре F - измеримых подмножеств Ω, называемых событиями) со значениями в (X,B) называется случайной величиной (чаще этот математический объект называют случайным элементом, оставляя термин "случайная величина" за частным случаем, когда X - числовая прямая) в фазовом пространстве (X,B). Распределением вероятностей этой случайной величины ξ называется функция $P_\xi=P_\xi(B)$ на σ-алгебре B фазового пространства, определенная как

$$P_\xi=P\{\xi\in B\}\quad (B\in \mathrm{B})\qquad (7)$$

(распределение вероятностей P_ξ представляет собой вероятностную меру в фазовом пространстве (X,B)) [25, с. 132].

Пусть $\xi_1,\xi_2,...,\xi_n$ - случайные величины на пространстве случайных событий (Ω,F,P) в соответствующих фазовых пространствах $(\mathrm{X}_k,\mathrm{B}_k)$. Совместным распределением вероятностей этих величин называется функция $P_{\xi_1,\xi_2,...,\xi_n}=P_{\xi_1,\xi_2,...,\xi_n}(B_1,B_2,...,B_n)$, определенная на множествах $B_1\in \mathrm{B}_1$, $B_2\in \mathrm{B}_2$, $\dots,B_n\in \mathrm{B}_n$ как

$$P_{\xi_1,\xi_2,...,\xi_n}(B_1,B_2,...,B_n)=P_{\xi_1,\xi_2,...,\xi_n}(\xi_1\in B_1,\xi_2\in B_2,...,\xi_n\in B_n).\qquad (8)$$

Распределение вероятностей $P_{\xi_1,\xi_2,...,\xi_n}$ как функция на полукольце множеств вида $B_1\times B_2\times...\times B_n$, $B_1\in \mathrm{B}_1, B_2\in \mathrm{B}_2,...,B_n\in \mathrm{B}_n$, в произведении пространств $\mathrm{X}_1,\mathrm{X}_2,...,\mathrm{X}_n$ представляет собой функцию распределения. Случайные величины $\xi_1,\xi_2,...,\xi_n$ называются независимыми, если при любых B_1, $B_2,…,B_n$ (см. [25, с.133])

$$P_{\xi_1,\xi_2,...,\xi_n}(B_1,B_2,...,B_n)=P_{\xi_1}(B_1)P_{\xi_2}(B_2)...P_{\xi_n}(B_n)\,.\qquad (9)$$

Предположим, что совместное распределение вероятностей $P_{\xi,\eta}(A,B)$ случайных величин ξ и η абсолютно непрерывно относительно некоторой меры Q на произведении пространств $\mathrm{X}\times\mathrm{Y}$, являющейся произведением мер Q_{X} и Q_{Y}, т.е.:

$$P_{\xi,\eta}(A,B)=\int\limits_{A\times B}p(x,y)Q(dx,dy)\qquad (10)$$

для любых $A\in \mathrm{A}$ и $B\in \mathrm{B}$, где $p(x,y)$ - соответствующая плотность распределения вероятностей [25, с.145].

В формуле (10) предполагается, что $\xi=\xi(\omega)$ и $\eta=\eta(\omega)$ - случайные величины на одном и том же пространстве элементарных событий Ω со значениями в фазовых пространствах (X, A) и (Y, B). Существование плотности *p(x,y)* вытекает из абсолютной непрерывности $P_{\xi,\eta}(A, B)$ относительно Q в соответствии с теоремой Радона - Никодима.

Условное распределение вероятностей $P_\xi(A \mid \eta), A \in \mathrm{A},$ может быть выбрано одинаковым для всех $\omega \in \Omega,$ при которых случайная величина $\eta = \eta(\omega)$ сохраняет одно и то же значение: $\eta(\omega) = y.$ При почти каждом $y \in \mathrm{Y}$ (относительно распределения P_η в фазовом пространстве (Y, B)) условное распределение вероятностей $P_\xi(A \mid y) = P_{\omega,\xi}(A),$ где $\omega \in \{\eta = y\}$ и $A \in \mathrm{A},$ будет абсолютно непрерывно относительно меры Q_X:

$$Q_\mathrm{X}(A) = \int_{A \times \mathrm{X}} Q(dx, dy).$$

Причем соответствующая плотность условного распределения вероятностей будет иметь вид (см. [25, с.145-146]):

$$p_\xi(x \mid y) = \frac{P_\xi(dx \mid y)}{Q_\mathrm{X}(dx)} = \frac{p(x, y)}{\int_\mathrm{X} p(x, y) Q_\mathrm{X}(dx)}. \qquad (11)$$

При построении вероятностных моделей реальных явлений важны вероятностные пространства из конечного числа элементарных событий. Для них перечисленные выше общие понятия становятся более прозрачными, в частности, снимаются вопросы измеримости (все подмножества конечного множества обычно считаются измеримыми). Вместо плотностей и условных плотностей рассматриваются вероятности и условные вероятности. Отметим, что вероятности можно рассматривать как плотности относительно меры, приписывающей каждому элементу пространства элементарных событий вес 1, т.е. считающей меры

$$Q(A) = Card(A)$$

(мера каждого множества равна числу его элементов). В целом ясно, что определения основных понятий теории вероятностей в общем случае практически не отличаются от таковых в элементарных курсах, во всяком случае с идейной точки зрения.

За последние двадцать лет в эконометрике и прикладной математической статистике сформировалась новая область - статистика нечисловых данных, она же - статистика объектов нечисловой природы. К настоящему времени она развита не менее, чем ранее выделенные статистика случайных величин, многомерный статистический анализ, статистика временных рядов и случайных процессов. Краткая сводка основных постановок и результатов математической статистики в пространствах нечисловой природы даны ниже в настоящей главе. Теория, построенная для результатов наблюдений, лежащих в пространствах общей природы, является центральным стержнем в статистике нечисловой природы. В ее рамках удалось разработать и изучить методы оценивания параметров и характеристик, проверки гипотез (в частности, с помощью статистик интегрального типа), параметрической и непараметрической регрессии (восстановления зависимостей), непараметрического оценивания плотности, дискриминантного и кластерного анализов и т.д.

Вероятностно-статистические методы, развитые для результатов наблюдений из пространств произвольного вида, позволяют единообразно проводить анализ данных из любого конкретного пространства. Так, в монографии [3] они применены к конечным случайным множествам, в работе [10] - к нечетким множествам. С их помощью установлено поведение обобщенного мнения экспертной комиссии (медианы Кемени) при увеличении числа экспертов, когда ответы экспертов лежат в том или ином пространстве бинарных отношений (см. см. пункт 4 настоящей главы и главу 12 ниже). В пункте 5 настоящей главы методы распознавания образов, основанные на непараметрических оценках плотности распределения вероятностей в пространстве общей природы, применены для разработки алгоритма диагностики в пространстве разнотипных данных (часть координат вектора измерена по количественным шкалам, часть - по качественным - см. главу 3).

8.3. Структура статистики объектов нечисловой природы

Как уже отмечалось, термин "статистика объектов нечисловой природы" впервые появился в 1979 г. в монографии [3]. В том же году в статье [16] была сформулирована программа развития этого нового направления прикладной математической статистики, которая к 1985 г. в основном была реализована.

Статистика объектов нечисловой природы как самостоятельное научное направление была выделена в нашей стране. Со второй половины 80-х годов существенно возрос интерес к этой тематике и у зарубежных исследователей. Это нашло отражение, в частности, на Первом Всемирном Конгрессе Общества математической статистики и теории вероятностей им. Бернулли, состоявшемся в сентябре 1986 г. в Ташкенте. Статистика объектов нечисловой природы используется в нормативно-технической и методической документации, ее применение позволяет получить существенный технико-экономический эффект (см. например, сводку [26]).

Однако тематика статистики объектов нечисловой природы обсуждалась до сих пор в основном в кругу развивающих ее специалистов, в результате она недостаточно отражена в монографической литературе. Цель настоящего пункта - дать введение в статистику объектов нечисловой природы, выделить ее структуру, указать основные идеи и результаты.

Напомним, что объектами нечисловой природы (см. также предыдущие пункты настоящей главы) называют элементы пространств, не являющихся линейными. Примерами являются бинарные отношения (ранжировки, разбиения, толерантности), множества, последовательности символов (тексты). Объекты нечисловой природы нельзя складывать и умножать на числа, не теряя при этом содержательного смысла. Этим они отличаются от издавна используемых в прикладной статистике (в качестве элементов выборок) чисел, векторов и функций.

Прикладную статистику по виду статистических данных принято делить на следующие направления:

статистика случайных величин (одномерная статистика);
многомерный статистический анализ;
статистика временных рядов и случайных процессов;
статистика объектов нечисловой природы.

При создании теории вероятностей и математической статистики исторически первыми были рассмотрены объекты нечисловой природы - белые и черные шары в урне. На основе соответствующих вероятностных моделей были введены биномиальное, гипергеометрическое и другие распределения, получены теоремы Муавра-Лапласа, Пуассона и др. Современное развитие этой тематики привело, в частности, к созданию теории статистического контроля качества продукции по альтернативному признаку (годен - не годен) в работах

А.Н.Колмогорова, Б.В. Гнеденко, Ю.К. Беляева, Я.П. Лумельского и многих других (см., например, классические монографии [7,8]).

В семидесятых годах в связи с запросами практики весьма усилился интерес к статистическому анализу нечисловых данных. Московская группа, организованная Ю.Н. Тюриным и другими специалистами вокруг созданного в 1973 г. научного семинара "Экспертные оценки и нечисловая статистика", развивала в основном вероятностную статистику нечисловых данных. Были установлены разнообразные связи между различными видами объектов нечисловой природы и изучены свойства этих объектов. Московской группой выпущены десятки сборников и обзоров, перечень которых приведен в итоговой работе [4]. Хотя в названиях многих из этих изданий стоят слова "экспертные оценки", анализ содержания сборников показывает, что подавляющая часть статей посвящена математико-статистическим вопросам, а не проблемам проведения экспертиз. Частое употребление указанных слов отражает лишь один из импульсов, стимулирующих развитие статистики объектов нечисловой природы и идущих от запросов практики. При этом необходимо подчеркнуть, что полученные результаты могут и должны активно использоваться в теории и практике экспертных оценок.

Новосибирская группа (Г.С. Лбов, Б.Г. Миркин и др.), как правило, не использовала вероятностные модели, т.е. вела исследования в рамках анализа данных. В московской группе в рамках анализа данных также велись работы, в частности, Б.Г.Литваком. Исследования по статистике объектов нечисловой природы выполнялись также в Ленинграде, Ереване, Киеве, Таллине, Тарту, Красноярске, Минске, Днепропетровске, Владивостоке, Калинине и других научных центрах.

Внутреннее деление статистики объектов нечисловой природы. Внутри рассматриваемого направления эконометрики и прикладной статистики выделим следующие области.

1. Статистика конкретных видов объектов нечисловой природы.
2. Статистика в пространствах общей (произвольной) природы.
3. Применение идей, подходов и результатов статистики объектов нечисловой природы в классических областях прикладной статистики.

Единство рассматриваемому направлению придает прежде всего вторая составляющая, позволяющая с единой точки зрения подходить к статистическим задачам описания данных, оценивания, проверки гипотез при рассмотрении выборки, элементы которой имеют ту или иную конкретную природу. Внутри первой составляющей рассмотрим:

1.1) теорию измерений;
1.2) статистику бинарных отношений;
1.3) теорию люсианов (бернуллиевских векторов);
1.4) статистику случайных множеств;
1.5) статистику нечетких множеств;
1.6) многомерное шкалирование;
1.7) аксиоматическое введение метрик.

Перечисленные разделы тесно связаны друг с другом, как продемонстрировано, в частности, в работах [3,15] и первых двух пунктах настоящей главы. Вне данного перечня остались работы по хорошо развитым классическим областям - статистическому контролю, таблицам сопряженности, а также по анализу текстов и некоторые другие (см.[4]). Таким образом, рассмотрим постановки 1970-2000 гг. вероятностной статистики объектов нечисловой природы.

Статистика в пространствах общей природы. Пусть $x_1,x_2,...,x_n$ - элементы пространства X, не являющегося линейным. Как определить среднее значение для $x_1,x_2,...,x_n$? Поскольку нельзя складывать элементы X, сравнивать их по величине, то необходимы подходы, принципиально новые по сравнению с классическими. В статистике объектов нечисловой природы предложено использовать показатель различия $d: X^2 \to [0,+\infty)$ (содержательный смысл показателя различия: чем больше $d(x,y)$, тем больше различаются x и y) и определять среднее как решение экстремальной задачи

$$E_n(d) = Arg\min\{\sum_{1\le i\le n} d(x_i, x), x \in X\}. \qquad (1)$$

Таким образом, среднее $E_n(d)$- это совокупность всех тех $x \in X$, для которых функция

$$f_n x) = \frac{1}{n}\sum_{1\le i\le n} d(x_i, x) \qquad (2)$$

достигает минимума на X.

Для классического случая $X = R^1$ при $d(x,y) = (x-y)^2$ имеем $E_n(d) = \bar{x}$, а при $d(x,y)=|x-y|$ среднее $E_n(d)$ совпадает с выборочной медианой (при нечетном объеме выборки; а при четном - $E_n(d)$ является отрезком с концами в двух средних элементах вариационного ряда).

Для ряда конкретных объектов среднее как решение экстремальной задачи вводилось рядом авторов. В 1929 г. итальянские статистики Джини и Гальвани применили такой подход для усреднения точек на плоскости и в пространстве Американский исследователь Джон Кемени решение задачи (1) называл медианой или средним для выборки, состоящей из ранжировок (см. монографию [24]). При моделировании лесных пожаров согласно выражению (1) было введено "среднеуклоняемое множество" для описания средней выгоревшей площади (см. об этом в монографии [3]). Общее определение среднего вида (1) было впервые введено в работе [16].

Основной результат, связанный со средними вида (1) - аналог закона больших чисел. Пусть $x_1,x_2,...,x_n$ - независимые одинаково распределенные случайные элементы со значениями в пространстве общей природы X. Теоретическим средним, или математическим ожиданием, в статистике объектов нечисловой природы называют

$$E_n(x_1, d) = Arg\min\{Ed(x_1, x), x \in X\}. \qquad (3)$$

Закон больших чисел состоит в сходимости $E_n(d)$ к $E_n(x_1,d)$ при $n \to \infty$. Поскольку и эмпирическое, и теоретическое средние - множества, то понятие сходимости требует уточнения.

Одно из возможных уточнений, впервые введенное в работе [16], таково. Для функции

$$f(x) = Ed(x_1(\omega), x), f: X \to R^1 \qquad (4)$$

введем понятие "ε-пятки" (ε>0)

$$K_\varepsilon(f) = \{x \in X: f(x) < \inf\{f(y), y \in X\} + \varepsilon\}. \qquad (5)$$

Очевидно, ε - пятка f - это окрестность *Argmin(f)* (если он достигается), заданная в терминах минимизируемой функции. Тем самым снимается вопрос о выборе метрики в пространстве X. Тогда при некоторых условиях регулярности для любого ε>0 вероятность события

$$\{\omega: E_n(d) \subseteq K_\varepsilon(f)\} \qquad (6)$$

стремится к 1 при. $n \to \infty$, т.е. справедлив закон больших чисел. Подробное доказательство приводится в следующем пункте настоящей главы.

Естественное обобщение рассматриваемой задачи позволяет построить общую теорию оптимизационного подхода в статистике. Как известно, большинство задач прикладной статистики может быть представлено в качестве оптимизационных [12]. Как себя ведут решения экстремальных задач? Частные случаи этой постановки: как ведут себя при росте объема выборки оценки максимального правдоподобия, минимального контраста (в том числе робастные в смысле Тьюки-Хьюбера - см. главу 10), оценки нагрузок в факторном анализе и методе главных компонент при отсутствии нормальности, оценки метода наименьших модулей в регрессии и т.д.

Обычно легко устанавливается, что для некоторых пространств X и последовательности случайных функций $f_n(x)$ при $n \to \infty$ найдется функция $f(x)$ такая, что

$$f_n(x) \to f(x) \qquad (7)$$

для любого $x \in X$ (сходимость по вероятности). Требуется вывести отсюда, что

$$Arg\min f_n(x) \to Arg\min f(x), \qquad (8)$$

т.е. решения экстремальных задач также сходятся. Понятие сходимости в соотношении (8) уточняется с помощью ε-пяток, как это сделано выше для закона больших чисел. Условия регулярности, при которых справедливо предельное соотношение (8), приведены в исследовании [27]. В подавляющем большинстве реальных задач эти условия выполняются.

Как оценить распределение случайного элемента в пространстве общей природы? Поскольку понятие функции распределения неприменимо, естественно использовать непараметрические оценки плотности. Что такое плотность распределения вероятностей в пространстве произвольной природы? Это функция $g: X \to [0,+\infty)$ такая, что для любого измеримого множества (т.е. случайного события) $A \subseteq X$ справедливо соотношение

$$P(x_1(\omega) \in A) = \int_A g(x)\mu(dx)\,, \qquad (9)$$

где μ - некоторая мера в X. Ряд непараметрических оценок плотности был предложен в работе [16]. Например, аналогом ядерных оценок плотности является оценка

$$g_n(x) = \frac{1}{\nu(h_n, x)} \sum_{1 \le i \le n} H\left(\frac{d(x_i, x)}{h_n}\right), \qquad (10)$$

где d - показатель различия; H - ядерная функция; h_n - последовательность положительных чисел; $\nu(h_n, x)$ - нормирующий множитель. Удалось установить, что, что статистики типа (10) обладают такими же свойствами, по крайней мере при фиксированном x, что и их классические аналоги при $X = R^1$. В частности, такой же скоростью сходимости. Некоторые изменения необходимы при рассмотрении дискретных X, каковыми являются многие пространства конкретных объектов нечисловой природы. С помощью непараметрических оценок плотности можно развивать регрессионный анализ, дискриминантный анализ и другие направления в пространствах общей природы (см. пункт 5 ниже).

Для проверки гипотез согласия, однородности, независимости в пространствах общей природы могут быть использованы статистики интегрального типа

$$\int f_n(x,\omega) dF_n(x,\omega), \qquad (11)$$

где $f_n(x,\omega)$-последовательность случайных функций на X; $F_n(x,\omega)$ - последовательность случайных распределений (или зарядов). Обычно $f_n(x,\omega)$

при $n \to \infty$ сходится по распределению к некоторой случайной функции $f(x,\omega)$, а $F_n(x,\omega)$ - к распределению *F(x).* Тогда распределение статистики интегрального типа (11) сходится к распределению случайного элемента

$$\int f(x,\omega)dF(x). \qquad (12)$$

Условия, при которых это справедливо, даны в работе [28]. Пример применения - вывод предельного распределения статистики типа омега-квадрат для проверки симметрии распределения (см. главу 4).

Перейдем к статистике конкретных видов объектов нечисловой природы.

Теория измерений. Цель теории измерений - борьба с субъективизмом исследователя при приписывании численных значений реальным объектам. Так, расстояния можно измерять в метрах, микронах, милях, парсеках и других единицах измерения. Выбор единиц измерения зависит от исследователя, т.е. субъективен. Статистические выводы могут быть адекватны реальности только тогда, когда они не зависят от того, какую именно единицу измерения предпочтет исследователь, т.е. когда они инвариантны относительно допустимого преобразования шкалы.

Теория измерений известна в нашей стране уже около 30 лет. С начала семидесятых годов активно работают отечественные исследователи. В настоящее время изложение основ теории измерений включают в справочные издания, помещают в научно-популярные журналы и книги для детей. Однако она еще не стала общеизвестной среди специалистов, в частности, среди метрологов. Поэтому опишем одну из задач теории измерений (ср. главу 3).

Как известно, шкала задается группой допустимых преобразований (прямой в себя). Номинальная шкала (шкала наименований) задается группой всех взаимно-однозначных преобразований, шкала порядка - группой всех строго возрастающих преобразований. Это - шкалы качественных признаков. Группа линейных возрастающих преобразований $\varphi(x) = ax + b, a > 0,$ задает шкалу интервалов. Группа $\varphi(x) = ax, a > 0,$ определяет шкалу отношений. Наконец, группа, состоящая из одного тождественного преобразования, описывает абсолютную шкалу. Это - шкалы количественных признаков. Используют и некоторые другие шкалы.

Практическую пользу теории измерений обычно демонстрируют на примере задачи сравнения средних значений для двух совокупностей одинакового объема $x_1, x_2,\dots,x_n$ и $y_1, y_2,\dots,y_n$. Пусть среднее вычисляется с помощью функции $f : R^n \to R^1$. Если

$$f(x_1, x_2,\dots,x_n)<f(y_1, y_2,\dots,y_n),. \qquad (13)$$

то необходимо, чтобы

$$f(\varphi(x_1),\varphi(x_2),\dots,\varphi(x_n)) < f(\varphi(y_1),\varphi(y_2),\dots,\varphi(y_n)) \qquad (14)$$

для любого допустимого преобразования φ из задающей шкалу группы Φ. (В противном случае результат сравнения будет зависеть от того, какое из эквивалентных представлений шкалы выбрал исследователь.)

Требование равносильности неравенств (13) и (14) вместе с некоторыми условиями регулярности приводят к тому, что в порядковой шкале в качестве средних можно использовать только члены вариационного ряда, в частности, медиану, но нельзя использовать среднее геометрическое, среднее арифметическое, и т.д. В количественных шкалах это требование выделяет из всех обобщенных средних по А.Н. Колмогорову в шкале интервалов - только среднее арифметическое, а в шкале отношений - только степенные средние.

Кроме средних, аналогичные задачи рассмотрены для расстояний, мер связи случайных признаков и других процедур анализа данных.

Приведенные результаты о средних величинах применялись, например, при проектировании системы датчиков в АСУ ТП доменных печей. Велико прикладное значение теории измерений в задачах стандартизации и управления качеством, в частности, в квалиметрии. Так, В.В. Подиновский показал, что любое изменение коэффициентов весомости единичных показателей качества продукции приводит к изменению упорядочения изделий по средневзвешенному показателю, а Н.В. Хованов развил одну из возможных теорий шкал измерения качества. Теория измерений полезна и в других прикладных областях.

Статистика бинарных отношений. Оценивание центра распределения случайного бинарного отношения проводят обычно с помощью медианы Кемени. Состоятельность вытекает из закона больших чисел [3]. Вычислительные процедуры нахождения медианы Кемени здесь не обсуждаем.

Методы проверки гипотез развиты отдельно для каждой разновидности бинарных отношений. В области статистики ранжировок, или ранговой корреляции, классической является книга Кендалла [6]. Современные достижения отражены в работах Ю.Н.Тюрина и Д.С.Шмерлинга. Статистика случайных разбиений развита А.В.Маамяги. Статистика случайных толерантностей (рефлексивных симметричных отношений) изложена в работе [3]. Многие ее задачи являются частными случаями задач теории люсианов.

Теория люсианов (бернуллиевских векторов). Люсиан (бернуллиевский вектор) - это последовательность испытаний Бернулли с, вообще говоря, различными вероятностями успеха. Реализация люсиана (бернуллиевского вектора) - это последовательность из 0 и 1. Люсианы (бернуллиевские вектора) рассматривались как случайные множества с независимыми элементами, а также - как результаты независимых парных сравнений. Последовательность результатов контроля качества последовательности единиц продукции по альтернативному признаку - также реализация люсиана (бернуллиевского вектора). Случайная толерантность может быть записана в виде люсиана. Поскольку один и тот же эконометрический объект применяется в различных областях, естественно для его наименования применять специально введенный термин "бернуллиевский вектор". Используется также термин "люсиан".

В рассматриваемой теории изучают методы проверки согласованности (одинаковой распределенности), однородности двух выборок, независимости люсианов. Методы проверки указанных гипотез нацелены на ситуацию, когда число бернуллиевских векторов фиксировано, а их длина растет. При этом число неизвестных параметров возрастает пропорционально объему данных, т.е. теория построена в асимптотике растущего числа параметров. Ранее подобная асимптотика под названием асимптотики А.Н.Колмогорова использовалась в дискриминантном анализе, но там применялись совсем другие методы.

Непараметрическая теория парных сравнений (в предположении независимости результатов отдельных сравнений) - часть теории бернуллиевских векторов. Параметрическая теория связана в основном с попытками выразить вероятности того или иного исхода через значения гипотетических или реальных параметров сравниваемых объектов. Известны модели Терстоуна, Бредли-Терри-Льюса и др.. В СССР построен ряд новых моделей парных сравнений (см. выше - второй пункт настоящей главы). Имеются модели парных сравнений с тремя исходами (больше, меньше, неразличимо), модели зависимых сравнений, сравнений нескольких объектов (сближающие рассматриваемую область с теорией случайных ранжировок) и т.д.

Статистика случайных и нечетких множеств. Давнюю историю имеет статистика случайных геометрических объектов (отрезков, треугольников, кругов и т.д.). Современная теория случайных множеств сложилась при изучении пористых сред и объектов сложной природы в таких областях, как металлография, петрография, биология. Различные направления внутри этой теории рассмотрены в работе [3, гл.4]. Остановимся на двух.

Случайные множества, лежащие в евклидовом пространстве, можно складывать: сумма множеств A и B- - это объединение всех векторов $x+y$, где $x \in A, y \in B$. Н.Н. Ляшенко получил аналоги законов больших чисел, центральной предельной теоремы, ряда методов прикладной статистики, систематически используя подобные суммы.

Для статистики объектов нечисловой природы интереснее подмножества пространств, не являющихся линейными. В работе [3] рассмотрены некоторые задачи теории конечных случайных множеств. Ряд интересных результатов получил С.А.Ковязин, в частности, он доказал нашу гипотезу о справедливости закона больших чисел при использовании расстояния между множествами

$$d(a,b) = \mu(A \Delta B), \qquad (15)$$

где μ - некоторая мера;. Δ - знак симметрической разности. Расстояние (15) выведено из некоторой системы аксиом в монографии [3]. Прикладники также делают попытки развивать методы статистики случайных множеств.

С теорией случайных множеств тесно связана теория нечетких множеств, начало которой положено статьей Л.А.Заде 1965 г. Это направление прикладной математики получило бурное развитие - к настоящему времени число публикаций измеряется десятками тысяч, имеются международные журналы, постоянно проводятся конференции, практические приложения дали ощутимый технико-экономический эффект. При изложении теории нечетких множеств обычно не подчеркивается связь с вероятностными моделями. Между тем еще в первой половине 1970-х годов было установлено [3], что теория нечеткости в определенном смысле сводится к теории случайных множеств, хотя эта связь и имеет лишь теоретическое значение.

С точки зрения статистики объектов нечисловой природы нечеткие множества - лишь один из видов объектов нечисловой природы. Поэтому к ним применима общая теория в пространствах произвольной природы. Имеются работы, в которых совместно используются соображения вероятности и нечеткости.

Многомерное шкалирование и аксиоматическое введение метрик. Многомерное шкалирование имеет целью представление объектов точками в пространстве небольшой размерности (1-3) с максимально возможным сохранением расстояний между точками.

Из сказанного выше ясно, какое большое место занимают в статистике объектов нечисловой природы метрики (расстояния). Как их выбрать? Предлагают выводить вид метрик из некоторых систем аксиом. Аксиоматически получена метрика в пространстве ранжировок, которая оказалась линейно связанной с коэффициентом ранговой корреляции Кендалла. Метрика (15) в пространстве множеств получена в работе [3] также исходя из некоторой системы аксиом. Г.В.Раушенбахом [23] дана сводка по аксиоматическому подходу к введению метрик в пространствах нечисловой природы. К настоящему времени практически для каждой используемой в прикладных работах метрики удалось подобрать систему аксиом, из которой чисто математическими средствами можно вывести именно эту метрику.

Применения статистики объектов нечисловой природы. Идеи, подходы, результаты статистики объектов нечисловой природы оказались полезными и в классических областях прикладной статистики. Статистика в пространствах общей природы позволила с единых позиций рассмотреть всю прикладную статистику, в частности, показать, что регрессионный, дисперсионный и дискриминантный анализы являются частными случаями общей схемы регрессионного анализа в пространстве произвольной природы. Поскольку структура модели - объект нечисловой природы, то ее оценивание, в частности, оценивание степени полинома в регрессии, также относится к статистике объектов нечисловой природы. Если учесть, что результаты измерения всегда имеют погрешность, т.е. являются не числами, а интервалами или нечеткими множествами, то приходим к необходимости пересмотреть некоторые выводы теоретической статистики. Например, отсутствует состоятельность оценок, нецелесообразно увеличивать объем выборок сверх некоторого предела (см. главу 9).

Технико-экономическая эффективность от применения методов статистики объектов нечисловой природы достаточно высока [114]. К сожалению, из-за изменения экономической ситуации, в частности, из-за инфляции трудно сопоставить конкретные экономические результаты в разные моменты времени. Кроме того, методы статистики объектов нечисловой природы составляют часть эконометрических методов, а те, в свою очередь - часть методов, входящих в систему информационной поддержки принятия решений на предприятии. Какую часть приращения прибыли предприятия надо отнести на эту систему? Мы знаем, как работает система управления фирмой в настоящем виде, но можем только гадать (а точнее, оценивать, скорее всего, с помощью экспертных оценок), каковы были бы результаты финансово-хозяйственной деятельности предприятия, если бы система управления фирмой была бы иной, например, не содержала методов статистики объектов нечисловой природы.

8.4. Законы больших чисел и состоятельность статистических оценок в пространствах произвольной природы

Законы больших чисел состоят в том, что эмпирические средние сходятся к теоретическим. В классическом варианте: выборочное среднее арифметическое при определенных условиях сходится по вероятности при росте числа слагаемых к математическому ожиданию. На основе законов больших чисел обычно доказывают состоятельность различных статистических оценок. В целом эта тематика занимает заметное место в теории вероятностей и математической статистике.

Однако математический аппарат при этом основан на свойствах сумм случайных величин (векторов, элементов линейных пространств). Следовательно, он не пригоден для изучения вероятностных и статистических проблем, связанных со случайными объектами нечисловой природы. Это такие объекты, как бинарные отношения, нечеткие множества, вообще элементы пространств без векторной структуры. Объекты нечисловой природы все чаще встречаются в прикладных исследованиях. Много конкретных примеров приведено выше в настоящей главе. Поэтому представляется полезным получение законов больших чисел в пространствах нечисловой природы. Необходимо решить следующие задачи.

А) Определить понятие эмпирического среднего.

Б) Определить понятие теоретического среднего.

В) Ввести понятие сходимости эмпирических средних к теоретическому.

Г) Доказать при тех или иных комплексах условий сходимость эмпирических средних к теоретическому.

Д) Обобщив это доказательство, получить метод обоснования состоятельности различных статистических оценок.

Е) Дать применения полученных результатов при решении конкретных задач.

Ввиду принципиальной важности рассматриваемых результатов приводим доказательство закона больших чисел, а также результаты компьютерного анализа множества эмпирических средних.

Определения средних величин. Пусть X - пространство произвольной природы, $x_1, x_2, x_3,...,x_n$ - его элементы. Чтобы ввести эмпирическое среднее для $x_1, x_2, x_3,...,x_n$ будем использовать действительнозначную (т.е. с числовыми значениями) функцию $f(x,y)$ двух переменных со значениями в X. В стандартных математических обозначениях, $f: X^2 \to R^1$. Величина $f(x,y)$ интерпретируется как показатель различия между x и y: чем $f(x,y)$ больше, тем x и y сильнее различаются. В качестве f можно использовать расстояние в X, квадрат расстояния и т.п.

Определение 1. Средней величиной для совокупности $x_1, x_2, x_3,...,x_n$ (относительно меры различия f), обозначаемой любым из трех способов:

$$x_{cp} = E_n(f) = E_n(x_1, x_2, x_3,...,x_n; f),$$

называем решение оптимизационной задачи

$$\sum_{i=1}^{n} f(x_i, y) \to \min, \quad y \in X.$$

Это определение согласуется с классическим: если $X = R^1$, $f(x,y) = (x - y)^2$, то x_{cp} - выборочное среднее арифметическое. Если же $X = R^1$, $f(x,y) = |x - y|$, то при $n = 2k+1$ имеем $x_{cp} = x(k+1)$, при $n = 2k$ эмпирическое среднее является отрезком $[x(k), x(k+1)]$. Здесь через $x(i)$ обозначен i-ый член вариационного ряда, построенного по $x_1, x_2, x_3,...,x_n$, т.е. i-я порядковая статистика. Таким образом, при $X = R^1$, $f(x,y) = |x - y|$ решение задачи (1) дает естественное определение выборочной медианы, правда, несколько отличающееся от предлагаемого в курсах "Общей теории статистики", в котором при $n = 2k$ медианой называют полусумму двух центральных членов вариационного ряда $(x(k) + x(k+1))/2$. Иногда $x(k)$ называют левой медианой , а $x(k+1)$ - правой медианой [3].

Решением задачи (1) является множество $E_n(f)$, которое может быть пустым, состоять из одного или многих элементов. Выше приведен пример, когда решением является отрезок. Если $X = R^1 \setminus \{x_0\}$, $f(x,y) = (x - y)^2$, а среднее арифметическое выборки равно x_0, то $E_n(f)$ пусто.

При моделировании реальных ситуаций часто можно принять, что X состоит из конечного числа элементов, а тогда $E_n(f)$ непусто - минимум на конечном множестве всегда достигается.

Понятия случайного элемента $x = x(\omega)$ со значениями в X, его распределения, независимости случайных элементов используем согласно пункту 2 настоящей главы, т.е. справочнику Ю.В. Прохорова и Ю.А. Розанова [25]. Будем считать, что функция f измерима относительно σ-алгебры, участвующей в определении случайного элемента $x = x(\omega)$. Тогда $f(x(\omega), y)$ при фиксированном y является действительнозначной случайной величиной. Предположим, что она имеет математическое ожидание.

Определение 2. Теоретическим средним (математическим ожиданием) для случайного элемента $x = x(\omega)$ относительно меры различия *f*, обозначаемом *E(x,f),* называется решение оптимизационной задачи

$$Ef(x(\omega), y) \to \min, y \in X.$$

Это определение также согласуется с классическим. Если $X = R^1$, $f(x,y) = (x - y)^2$, то $E(x,f) = E(x)$ - обычное математическое ожидание, при этом $E\, f(x(\omega), y)$ - дисперсия случайной величины $x = x(\omega)$. Если же $X = R^1$, $f(x,y) = |x - y|$, то $E(x,f) = [a,b]$, где $a = sup\{t: F(t)\leq 0,5\}$, $b = inf\{t: F(t)\geq 0,5\}$, причем $F(t)$ - функция распределения случайной величины $x = x(\omega)$. Если график $F(t)$ имеет плоский участок на уровне $F(t) = 0,5$, то медиана - теоретическое среднее в смысле определения 2 - является отрезком. В классическом случае обычно говорят, что каждый элемент отрезка *[a; b]* является одним из возможных значений медианы. Поскольку наличие указанного плоского участка - исключительный случай, то обычно решением задачи (2) является множество из одного элемента $a = b$ - классическая медиана распределения случайной величины $x = x(\omega)$.

Теоретическое среднее E(x,f) можно определить лишь тогда, когда $Ef(x(\omega), y)$ существует при всех $y \in X$. Оно может быть пустым множеством, например, если $X = R^1 \setminus \{x_0\}$, $f(x,y) = (x - y)^2$, $x_0 = E(x)$. И то, и другое исключается, если *X* конечно. Однако и для конечных *X* теоретическое среднее может состоять не из одного, а из многих элементов. Отметим, однако, что в множестве всех распределений вероятностей на *X* подмножество тех распределений, для которых *E(x,f)* состоит более чем из одного элемента, имеет коразмерность 1, поэтому основной является ситуация, когда множество *E(x,f)* содержит единственный элемент [3].

Существование средних величин. Под существованием средних величин будем понимать непустоту множеств решений соответствующих оптимизационных задач.

Если *X* состоит из конечного числа элементов, то минимум в задачах (1) и (2) берется по конечному множеству, а потому, как уже отмечалось, эмпирические и теоретические средние существуют.

Ввиду важности обсуждаемой темы приведем доказательства. Для строгого математического изложения нам понадобятся термины из раздела математики под названием "общая топология". Топологические термины и результаты будем использовать в соответствии с классической монографией [29]. Так, топологическое пространство называется бикомпактным в том и только в том случае, когда из каждого его открытого покрытия можно выбрать конечное подпокрытие [29, с.183]..

Теорема 1. Пусть *X* - бикомпактное пространство, функция *f* непрерывна на X^2 (в топологии произведения). Тогда эмпирическое и теоретическое средние существуют.

Доказательство. Функция $f(x_i,y)$ от y непрерывна, сумма непрерывных функций непрерывна, непрерывная функция на бикомпакте достигает своего минимума, откуда и следует заключение теоремы относительно эмпирического среднего.

Перейдем к теоретическому среднему. По теореме Тихонова [29, с.194] из бикомпактности *X* вытекает бикомпактность X^2. Для каждой точки *(x, y)* из X^2

рассмотрим $\varepsilon/2$ - окрестность в X^2 в смысле показателя различия f, т.е. множество

$$U(x,y) = \{(x',y') : |f(x,y) - f(x',y')| < \varepsilon/2\}.$$

Поскольку f непрерывна, то множества $U(x,y)$ открыты в рассматриваемой топологии в X^2. По теореме Уоллеса [29, с.193] существуют открытые (в X) множества $V(x)$ и $W(y)$, содержащие x и y соответственно и такие, что их декартово произведение $V(x)$ x $W(y)$ целиком содержится внутри $U(x, y)$.

Рассмотрим покрытие X^2 открытыми множествами $V(x)$ x $W(y)$. Из бикомпактности X^2 вытекает существование конечного подпокрытия $\{V(x_i)$ x $W(y_i), i = 1,2,...,m\}$. Для каждого x из X рассмотрим все декартовы произведения $V(x_i)$ x $W(y_i)$, куда входит точка (x, y) при каком-либо y. Таких декартовых произведений и их первых множителей $V(x_i)$ конечное число. Возьмем пересечение таких первых множителей $V(x_i)$ и обозначим его $Z(x)$. Это пересечение открыто, как пересечение конечного числа открытых множеств, и содержит точку x. Из покрытия бикомпактного пространства X открытыми множествами $Z(x)$ выберем открытое подпокрытие $Z_1, Z_2, ..., Z_k$.

Покажем, что если x_1' и x_2' принадлежат одному и тому же Z_j при некотором j, то

$$\sup\{|f(x_1',y) - f(x_2',y)|, y \in X\} < \varepsilon. \qquad (3)$$

Пусть $Z_j = Z(x_0)$ при некотором x_0. Пусть $V(x_i)$ x $W(y_i)$, $i \in I$, - совокупность всех тех исходных декартовых произведений из системы $\{V(x_i)$ x $W(y_i), i = 1,2,...,m\}$, куда входят точки (x_0, y) при различных y. Покажем, что их объединение содержит также точки (x_1', y) и (x_2', y) при всех y. Действительно, если (x_0, y) входит в $V(x_i)$ x $W(y_i)$, то y входит в $W(y_i)$, а x_1' и x_2' вместе с x_0 входят в $V(x_i)$, поскольку x_1', x_2' и x_0 входят в $Z(x_0)$. Таким образом, (x_1', y) и (x_2', y) принадлежат $V(x_i)$ x $W(y_i)$, а потому согласно определению $V(x_i)$ x $W(y_i)$

$$|f(x_1',y) - f(x_i,y_i)| < \varepsilon/2, \quad |f(x_2',y) - f(x_i,y_i)| < \varepsilon/2,$$

откуда и следует неравенство (3).

Поскольку X^2 - бикомпактное пространство, то функция f ограничена на X^2, а потому существует математическое ожидание $E\ f(x(\omega), y)$ для любого случайного элемента $x(\omega)$, удовлетворяющего приведенным в предыдущем разделе условиям согласования топологии, связанной с f, и измеримости, связанной с $x(\omega)$. Если x_1 и x_2 принадлежат одному открытому множеству Z_j, то

$$|Ef(x_1,y) - |Ef(x_2,y)| < \varepsilon,$$

а потому функция

$$g(y) = E\,f(x(\omega), y) \qquad (4)$$

непрерывна на X. Поскольку непрерывная функция на бикомпактном множестве достигает своего минимума, т.е. существуют такие точки z, на которых $g(z) = inf\{g(y), y \in X\}$, то теорема 1 доказана.

В ряде интересных для приложений ситуаций X не является бикомпактным пространством. Например, если $X = R^1$. В этих случаях приходится наложить на показатель различия f некоторые ограничения, например, так, как это сделано в теореме 2.

Теорема 2. Пусть X - топологическое пространство, непрерывная (в топологии произведения) функция f: $X^2 \to R^1$ неотрицательна, симметрична (т.е.

$f(x,y) = f(y,x)$ для любых x и y из X), существует число $D>0$ такое, что при всех x,y,z из X

$$f(x,y) \le D\{f(x,z) + f(z,y)\}. \qquad (5)$$

Пусть в X существует точка x_0 такая, что при любом положительном R множество $\{x: f(x, x_0) \le R\}$ является бикомпактным. Пусть для случайного элемента $x(\omega)$, согласованного с топологией в рассмотренном выше смысле, существует $g(x_0) = Ef(x(\omega), x_0)$.

Тогда существуют (т.е. непусты) математическое ожидание $E(x,f)$ и эмпирические средние $E_n(f)$.

Замечание. Условие (5) - некоторое обобщение неравенства треугольника. Например, если g - метрика в X, а $f = g^p$ при некотором натуральном p, то для f выполнено соотношение (5) с $D = 2^p$.

Доказательство. Рассмотрим функцию $g(y)$, определенную формулой (4). Имеем

$$f(x(\omega),y) \le D\{f(x(\omega), x_0) + f(x_0,y)\}. \qquad (6)$$

Поскольку по условию теоремы $g(x_0)$ существует, а потому конечно, то из оценки (6) следует существование и конечность $g(y)$ при всех y из X. Докажем непрерывность этой функции.

Рассмотрим шар (в смысле меры различия f) радиуса R с центром в x_0:

$$K(R) = \{x : f(x, x_0) \le R\}, \quad R > 0.$$

В соответствии с условием теоремы $K(R)$ как подпространство топологического пространства X является бикомпактным. Рассмотрим произвольную точку x из X. Справедливо разложение

$$f(x(\omega), y) = f(x(\omega), y)\chi(x(\omega) \in K(R)) + f(x(\omega), y)\chi(x(\omega) \notin K(R)),$$

где $\chi(C)$ - индикатор множества C. Следовательно,

$$g(y) = Ef(x(\omega), y)\chi(x(\omega) \in K(R)) + Ef(x(\omega), y)\chi(x(\omega) \notin K(R)). \qquad (7)$$

Рассмотрим второе слагаемое в (7). В силу (5)

$$f(x(\omega), y)\chi(x(\omega) \notin K(R)) \le D\{f(x(\omega), x_0)\chi(x(\omega) \notin K(R)) + f(x_0, y)\chi(x(\omega) \notin K(R))\}. \qquad (8)$$

Возьмем математическое ожидание от обеих частей (8):

$$Ef(x(\omega), y)\chi(x(\omega) \notin K(R)) \le D\int_R^{+\infty} t dP\{f(x(\omega), x_0) \le t\} + Df(x_0, y)P(x(\omega) \notin K(R)). \qquad (9)$$

В правой части (9) оба слагаемых стремятся к 0 при безграничном возрастании R: первое - в силу того, что

$$g(x_0) = Ef(x(\omega), x_0) = \int_0^{+\infty} t dP(f(x(\omega), x_0) \le t) < \infty,$$

второе - в силу того, что распределение случайного элемента $x(\omega)$ сосредоточено на X и

$$X \setminus \bigcup_{R>0} K(R) = \varnothing.$$

Пусть $U(x)$ - такая окрестность x (т.е. открытое множество, содержащее x), для которой

$$\sup\{f(y, x), \ y \in U(x)\} < +\infty.$$

Имеем

$$f(y, x_0) \le D(f(x_0, x) + f(x, y)). \qquad (10)$$

В силу (9) и (10) при безграничном возрастании R

$$Ef(x(\omega), y)\chi(x(\omega) \notin K(R)) \to 0 \qquad (11)$$

равномерно по $y \in U(x)$. Пусть $R(0)$ таково, что левая часть (11) меньше $\varepsilon > 0$ при $R>R(0)$ и, кроме того, $y \in U(x) \subseteq K(R(0))$. Тогда при $R>R(0)$

$$|g(y)-g(x)| \leq |Ef(x(\omega),y)\chi(x(\omega)\in K(R)) - Ef(x(\omega),x)\chi(x(\omega)\in K(R))| + 2\varepsilon. \quad (12)$$

Нас интересует поведение выражения в правой части формулы (12) при $y \in U(x)$. Рассмотрим $f1$ - сужение функции f на замыкание декартова произведения множеств $U(x)$ x $K(R)$, и случайный элемент $x_1(\omega) = x(\omega)\chi(x(\omega)\in K(R))$. Тогда

$$Ef(x(\omega),y)\chi(x(\omega)\in K(R)) = Ef_1(x_1(\omega),y)$$

при $y \in U(x)$, а непрерывность функции $g_1(y) = Ef_1(x_1(\omega),y)$ была доказана в теореме 1. Последнее означает, что существует окрестность $U_1(x)$ точки x такая, что

$$|Ef_1(x_1(\omega),y) - Ef_1(x_1(\omega),x)| < \varepsilon \quad (13)$$

при $y \in U_1(x)$. Из (12) и (13) вытекает, что при $y \in U(x) \bigcap U_1(x)$

$$|g(y)-g(x)| < 3\varepsilon,$$

что и доказывает непрерывность функции $g(x)$.

Докажем существование математического ожидания $E(x,f)$. Пусть $R(0)$ таково, что

$$P(x(\omega)\in K(R(0))) > 1/2. \quad (14)$$

Пусть H - некоторая константа, значение которой будет выбрано позже. Рассмотрим точку x из множества $K(HR(0))^C$ - дополнения $K(HR(0))$, т.е. из внешности шара радиуса $HR(0)$ с центром в x_0. Пусть $x(\omega)\in K(R(0))$. Тогда имеем

$$f(x_0,x) \leq D\{f(x_0,x(\omega)) + f(x(\omega),x)\},$$

откуда

$$f(x(\omega),x) \geq \frac{1}{D} f(x_0,x) - f(x_0,x(\omega)) \geq \frac{HR(0)}{D} - R(0). \quad (15)$$

Выбирая H достаточно большим, получим с учетом условия (14), что при $x \in K(HR(0))^C$ справедливо неравенство

$$Ef(x(\omega),x) \geq \frac{1}{2}\left(\frac{HR(0)}{D} - R(0)\right). \quad (16)$$

Можно выбрать H так, чтобы правая часть (16) превосходила $g(x_0) = Ef(x(\omega),x_0)$.

Сказанное означает, что $Argmin\ g(x)$ достаточно искать внутри бикомпактного множества $K(HR(0))$. Из непрерывности функции g вытекает, что ее минимум достигается на указанном бикомпактном множестве, а потому - и на всем X. Существование (непустота) теоретического среднего $E(x,f)$ доказана.

Докажем существование эмпирического среднего $E_n(f)$. Есть искушение проводить его дословно так же, как и доказательство существования математического ожидания $E(x,f)$, лишь с заменой 1/2 в формуле (16) на частоту попадания элементов выборки x_i в шар $K(R(0))$, каковая, очевидно, стремится к вероятности попадания случайного элемента $ч = x(\omega)$ в $K(R(0))$, большей 1/2 в соответствии с (14). Однако это рассуждение показывает лишь, что вероятность непустоты $E_n(f)$ стремится к 1 при безграничном росте объема выборки. Точнее, оно показывает, что

$$\lim_{n\to\infty} P\{E_n(f) \neq \varnothing \wedge E_n(f) \subseteq K(HR(0))\} = 1.$$

Поэтому пойдем другим путем, не опирающимся к тому же на вероятностную модель выборки. Положим

$$R(1) = \max\{f(x_i,x_0), i = 1,2,...,n\}. \quad (17)$$

Если x входит в дополнение шара $K(HR(1))$, то аналогично (15) имеем

$$f(x_i, x_0) \geq \frac{HR(1)}{D} - R(1). \qquad (18)$$

При достаточно большом H из (17) и (18) следует, что

$$\sum_{i=1}^{n} f(x_i, x_0) \leq nR(1) < \sum_{i=1}^{n} f(x_i, x), \quad x \in \{K(HR(1))\}^C.$$

Следовательно, *Argmin* достаточно искать на *K(HR(1)).* Заключение теоремы 2 следует из того, что на бикомпактном пространстве *K(HR(1))* минимизируется непрерывная функция.

Теорема 2 полностью доказана.

О формулировках законов больших чисел. Пусть $x, x_1, x_2, x_3, ..., x_n$ - независимые одинаково распределенные случайные элементы со значениями в X. Закон больших чисел - это утверждение о сходимости эмпирических средних к теоретическому среднему (математическому ожиданию) при росте объема выборки n, т.е. утверждение о том, что

$$E_n(f) = E_n(x_1, x_2, x_3, ..., x_n; f) \to E(x, f) \qquad (19)$$

при $n \to \infty$. Однако и слева, и справа в формуле (19) стоят, вообще говоря, множества. Поэтому понятие сходимости в (19) требует обсуждения и определения.

В силу классического закона больших чисел при $n \to \infty$

$$\frac{1}{n}\sum_{i=1}^{n} f(x_i, y) \to Ef(x, y) \qquad (20)$$

в смысле сходимости по вероятности, если правая часть существует (теорема А.Я. Хинчина, 1923 г.).

Если пространство X состоит из конечного числа элементов, то из соотношения (20) легко вытекает (см., например, [3, с.192-193]), что

$$\lim_{n\to\infty} P\{E_n(f) \subseteq E(x, f)\} = 1. \qquad (21)$$

Другими словами, $E_n(f)$ является состоятельной оценкой $E(x, f)$.

Если $E(x, f)$ состоит из одного элемента, $E(x, f) = \{x_0\}$, то соотношение (21) переходит в следующее:

$$\lim_{n\to\infty} P\{E_n(f) = \{x_0\}\} = 1. \qquad (22)$$

Однако с прикладной точки зрения доказательство соотношений (21)-(22) не дает достаточно уверенности в возможности использования $E_n(f)$ в качестве оценки *E(x,f),* поскольку в процессе доказательства объем выборки предполагается настолько большим, что при всех $y \in X$ одновременно левые части соотношений (20) сосредотачиваются в непересекающихся окрестностях правых частей.

Замечание. Если в соотношении (20) рассмотреть сходимость с вероятностью 1, то аналогично (21) получим т.н. усиленный закон больших чисел [3, с.193-194], согласно которому с вероятностью 1 эмпирическое среднее $E_n(f)$ входит в теоретическое среднее *E(x,f),* начиная с некоторого объема выборки *n*, вообще говоря, случайного, $n = n(\omega)$. Мы не будем останавливаться на этом виде сходимости, поскольку в соответствующих постановках, подробно разобранных в монографии [3], нет принципиальных отличий от случая сходимости по вероятности.

Если X не является конечным, например, $X = R^1$, то соотношения (21) и (22) неверны. Поэтому необходимо искать иные формулировки закона больших чисел. В классическом случае сходимости выборочного среднего арифметического к математическому ожиданию, т.е. $\bar{x} \to E(x)$ можно записать закон больших чисел так: для любого $\varepsilon > 0$ справедливо предельное соотношение

$$\lim_{n\to\infty} P\{\bar{x} \in (E(x)-\varepsilon; E(x)+\varepsilon)\} = 1. \qquad (23)$$

В этом соотношении в отличие от (21) речь идет о попадании эмпирического среднего $E_n(f) = \bar{x}$ не непосредственно внутрь теоретического среднего *E(x,f),* а в некоторую *окрестность* теоретического среднего.

Обобщим эту формулировку. Как задать окрестность теоретического среднего в пространстве произвольной природы? Естественно взять его окрестность, определенную с помощью какой-либо метрики. Однако полезно обеспечить на ее дополнении до *X* *отделенность* множества значений $Ef(x(\omega),y)$ как функции y от минимума этой функции на всем *X*.

Поэтому мы сочли целесообразным определить такую окрестность с помощью самой функции $Ef(x(\omega),y)$.

Определение 3. Для любого $\varepsilon > 0$ назовем ε-пяткой функции *g(x)* множество

$$K_\varepsilon(g) = \{x : g(x) < \inf\{d(y), y \in X\}, x \in X\}.$$

Таким образом, в ε-пятку входят все те *x*, для которых значение *g(x)* либо минимально, либо отличается от минимального (или от инфимума) не более чем на ε. Так, для $X = R^1$ и функции $g(x) = x^2$ минимум равен 0, а ε-пятка имеет вид интервала $(-\sqrt{\varepsilon}; \sqrt{\varepsilon})$. В формулировке (23) классического закона больших чисел утверждается, что при любом ε>0 вероятность попадания среднего арифметического в $\sqrt{\varepsilon}$-пятку математического ожидания стремится к 1. Поскольку $\varepsilon > 0$ произвольно, то вместо $\sqrt{\varepsilon}$-пятки можно говорить о ε-пятке, т.е. перейти от (23) к эквивалентной записи

$$\lim_{n\to\infty} P\{\bar{x} \in K_\varepsilon(E(x(\omega)-x)^2\} = 1. \qquad (24)$$

Соотношение (24) допускает непосредственное обобщение на общий случай пространств произвольной природы.

СХЕМА ЗАКОНА БОЛЬШИХ ЧИСЕЛ. Пусть $x, x_{1,} x_2, x_3, ..., x_n$ - независимые одинаково распределенные случайные элементы со значениями в пространстве произвольной природы *X* с показателем различия $f: X^2 \to R^1$. *Пусть выполнены некоторые математические условия регулярности.* Тогда для любого $\varepsilon > 0$ справедливо предельное соотношение

$$\lim_{n\to\infty} P\{E_n(f) \subseteq K_\varepsilon(E(x,f))\} = 1. \qquad (25)$$

Аналогичным образом может быть сформулирована и общая идея усиленного закона больших чисел. Ниже приведены две конкретные формулировки "условий регулярности".

Законы больших чисел. Начнем с рассмотрения естественного обобщения конечного множества - бикомпактного пространства *X*.

Теорема 3. В условиях теоремы 1 справедливо соотношение (25).

Доказательство. Воспользуемся построенным при доказательстве теоремы 1 конечным открытым покрытием $\{Z_1, Z_2, ..., Z_k\}$ пространства *X* таким, что для него выполнено соотношение (3). Построим на его основе разбиение *X* на непересекающиеся множества $W_1, W_2, ..., W_m$ (объединение элементов разбиения $W_1, W_2, ..., W_m$ составляет *X*). Это можно сделать итеративно. На первом шаге из Z_1 следует вычесть $Z_2, ..., Z_k$ - это и будет W_1. Затем в качестве нового пространства надо рассмотреть разность X и W_1 , а покрытием его будет $\{Z_2, ..., Z_k\}$. И так до k-го шага, когда последнее из рассмотренных покрытий будет состоять из единственного открытого множества Z_k . Остается из построенной последовательности $W_1, W_2, ..., W_k$ вычеркнуть пустые множества, которые могли быть получены при осуществлении описанной процедуры (поэтому, вообще говоря, m может быть меньше k).

В каждом из элементов разбиения $W_1, W_2, ..., W_m$ выберем по одной точке, которые назовем центрами разбиения и соответственно обозначим $w_1, w_2, ..., w_m$. Это и есть то конечное множество, которым можно аппроксимировать бикомпактное пространство X. Пусть y входит в W_j . Тогда из соотношения (3) вытекает, что

$$\left|\frac{1}{n}\sum_{i=1}^{n} f(x_i, y) - \frac{1}{n}\sum_{i=1}^{n} f(x_i, w_j)\right| < \varepsilon. \qquad (26)$$

Перейдем к доказательству соотношения (25). Возьмем произвольное $\delta>0$. Рассмотрим некоторую точку b из $E(x,f)$. Доказательство будет основано на том, что с вероятностью, стремящейся к 1, для любого y вне $K_\delta(E(x,f))$ выполнено неравенство

$$\frac{1}{n}\sum_{i=1}^{n} f(x_i, y) > \frac{1}{n}\sum_{i=1}^{n} f(x_i, b). \qquad (27)$$

Для обоснования этого неравенства рассмотрим все элементы разбиения $W_1, W_2, ..., W_m$, имеющие непустое пересечение с внешностью δ-пятки $K_\delta(E(x,f))$. Из неравенства (26) следует, что для любого y вне $K_\delta(E(x,f))$ левая часть неравенства (27) не меньше

$$\min_j\left(\frac{1}{n}\sum_{i=1}^{n} f(x_i, w_j)\right) - \varepsilon, \qquad (28)$$

где минимум берется по центрам всех элементов разбиения, имеющим непустое пересечение с внешностью δ-пятки. Возьмем теперь в каждом таком разбиении точку v_i , лежащую вне -пятки $K_\delta(E(x,f))$. Тогда из неравенств (3) и (28) следует, что левая часть неравенства (27) не меньше

$$\min_j\left(\frac{1}{n}\sum_{i=1}^{n} f(x_i, v_j)\right) - 2\varepsilon. \qquad (29)$$

В силу закона больших чисел для действительнозначных случайных величин каждая из участвующих в соотношениях (27) и (29) средних арифметических имеет своими пределами соответствующие математические ожидания, причем в соотношении (29) эти пределы не менее

$$Ef(x(\omega), b) + \delta - 2\varepsilon,$$

поскольку точки v_i лежат вне δ-пятки $K_\delta(E(x,f))$. Следовательно, при

$$\delta - 2\varepsilon > 0$$

и достаточно большом n, обеспечивающем необходимую близость рассматриваемого конечного числа средних арифметических к их математическим ожиданиям, справедливо неравенство (27).

Из неравенства (27) следует, что пересечение $E_n(f)$ с внешностью $K_\delta(E(x,f))$ пусто. При этом точка b может входить в $E_n(f)$, а может и не входить. Во втором случае $E_n(f)$ состоит из иных точек, входящих в $K_\delta(E(x,f))$. Теорема 3 доказана.

Если X не является бикомпактным пространством, то необходимо суметь оценить рассматриваемые суммы "на периферии", вне бикомпактного ядра, которое обычно выделяется естественным путем. Один из возможных комплексов условий сформулирован выше в теореме 2.

Теорема 4. В условиях теоремы 2 справедлив закон больших чисел, т.е. соотношение (25).

Доказательство. Будем использовать обозначения, введенные в теореме 2 и при ее доказательстве. Пусть r и R, $r < R$, - положительные числа. Рассмотрим точку x в шаре $K(r)$ и точку y вне шара $K(R)$. Поскольку

$$f(x_0, y) \leq D\{f(x_0, x) + f(x < y)\},$$

то

$$f(x, y) \geq \frac{1}{D} f(x_0, y) - f(x_0, x) \geq \frac{R}{D} - r. \qquad (30)$$

Положим

$$g_n(x) = g_n(x, \omega) = \frac{1}{n} \sum_{i=1}^{n} f(x_i, x).$$

Сравним $g_n(x_0)$ и $g_n(y)$. Выборку $x_{1,} x_2, x_3, ..., x_n$ разобьем на две части. В первую часть включим те элементы выборки, которые входят в $K(r)$, во вторую - все остальные (т.е. лежащие вне $K(r)$). Множество индексов элементов первой части обозначим $I = I(n,r)$. Тогда в силу неотрицательности f имеем

$$g_n(y) \geq \frac{1}{n} \sum_{i \in I} f(x_i, y),$$

а в силу неравенства (30)

$$\sum_{i \in I} f(x_i, y) \geq \left(\frac{R}{D} - r \right) Card I(n, r),$$

где $Card\ I(n,r)$ - число элементов в множестве индексов $I(n,r)$. Следовательно,

$$g_n(y) \geq \frac{1}{n} \left(\frac{R}{D} - r \right) J, \qquad (31)$$

где $J = Card\ I(n,r)$ - биномиальная случайная величина $B(n,p)$ с вероятностью успеха $p = P\{x_i(\omega) \in K(r)\}$. По теореме Хинчина для $g_n(x_0)$ справедлив (классический) закон больших чисел. Пусть $\varepsilon > 0$. Выберем $n_1 = n_1(\varepsilon)$ так, чтобы при $n > n_1$ было выполнено соотношение

$$P\{g_n(x_0) - g(x_0) > \varepsilon\} < \varepsilon, \qquad (32)$$

где $g(x_0) = Ef(x_1, x_0)$. Выберем r так, чтобы вероятность успеха p>0,6. По теореме Бернулли можно выбрать $n_2 = n_2(\varepsilon)$ так, чтобы при $n > n_2$

$$P\{J > 0{,}5n\} > 1 - \varepsilon. \qquad (33)$$

Выберем R так, чтобы

$$\frac{1}{2} \left(\frac{R}{D} - r \right) > g(x_0) + \varepsilon.$$

Тогда

$$K_\varepsilon(g) \subseteq K(R) \qquad (34)$$

и согласно (31), (32) и (33) при $n > n_3 = \max(n_1, n_2)$ с вероятностью не менее $1 - \varepsilon$ имеем

$$g_n(y) > g_n(x_0) \qquad (35)$$

для любого y вне $K(R)$. Из (34) следует, что минимизировать g_n достаточно внутри бикомпактного шара $K(R)$, при этом $E_n(f)$ не пусто и

$$E_n(f) \subseteq K(R) \qquad (36)$$

с вероятностью не менее 1-2ε.

Пусть $g_n^{'}$ и $g^{'}$ - сужения g_n и *g(x) = Ef(x(ω), x)* соответственно на *K(R)* как функций от *x*. В силу (34) справедливо равенство $K_\varepsilon(g^{'}) = K_\varepsilon(f)$. Согласно доказанной выше теореме 3 найдется $n_4 = n_4(\omega)$ такое, что

$$P(K_0(g_n^{'}) \subseteq K_\varepsilon(g)) > 1 - \varepsilon.$$

Согласно (36) с вероятностью не менее $1 - 2\varepsilon$

$$K_0(g_n^{'}) = E_n(f)$$

при $n > n_3$. Следовательно, при $n > n_5(\varepsilon) = \max(n_3, n_4)$ имеем

$$P(E_n(f) \subseteq K_\varepsilon(g)) > 1 - 3\varepsilon,$$

что и завершает доказательство теоремы 4.

Справедливы и иные варианты законов больших чисел, полученные, в частности, в статье [27].

Асимптотическое поведение решений экстремальных статистических задач. Если проанализировать приведенные выше постановки и результаты, особенно теоремы 1 и 3, то становится очевидной возможность их обобщения. Так, доказательства этих теорем практически не меняются, если считать, что функция *f(x,y)* определена на декартовом произведении бикомпактных пространств *X* и *Y*. Тогда можно считать, что элементы выборки лежат в *X*, а *Y* - пространство параметров, подлежащих оценке. Пусть, например, выборка взята из распределения с плотностью *p(x,y)*. Если положить

$$f(x,y) = - \ln p(x,y) ,$$

то задача нахождения эмпирического среднего переходит в задачу оценивания неизвестного параметра *y* методом максимального правдоподобия, а законы больших чисел переходят в утверждения о состоятельности этих оценок в случае пространств *X* и *Y* общего вида. В случае функции *f(x,y)* общего вида можно говорить об определении и состоятельности так называемых оценок минимального контраста. Частными случаями этих оценок являются, например, устойчивые (робастные) оценки Тьюки-Хубера (см. главу 10 ниже), оценки параметров в задачах аппроксимации (параметрической регрессии) в пространствах произвольной природы.

Можно пойти и дальше в обобщении законов больших чисел. Пусть известно, что при каждом конкретном *y* при безграничном росте *n* имеет быть сходимость по вероятности

$$f_n(x(\omega), y) \to f(y).$$

В каких случаях и в каком смысле

$$Argmin\ \{f_n(x(\omega), y), y \in X\} \to Argmin\ \{ f(y), y \in X\}\ ?$$

Причем здесь можно под *n* понимать натуральное число. А можно рассматривать "сходимость по фильтру" в смысле Картана и Бурбаки [29, с.118]. В частности, описывать ситуацию вектором, координаты которого - объемы нескольких выборок, и все они безгранично растут. В классической математической статистике такие постановки рассматривать не любят.

Поскольку, как уже отмечалось, основные задачи прикладной статистики можно представить в виде оптимизационных задач, то ответ на поставленный вопрос дает возможность единообразного подхода к изучению асимптотики решений разнообразных экстремальных статистических задач. Одна из возможных формулировок дана и обоснована выше. Другая - в работе [28]. Она основана на использовании понятий асимптотической равномерной разбиваемости и координатной асимптотической равномерной разбиваемости. С помощью указанных подходов удается стандартным образом обосновывать состоятельность оценок характеристик и параметров в основных задачах прикладной статистики. К сожалению, в рамках настоящей главы нет возможности подробнее остановиться на проблеме оценивания.

Рассматриваемую тематику можно развивать дальше, в частности, рассматривать пространства X и Y, не являющиеся бикомпактными, а также изучать скорость сходимости эмпирических средних к теоретическим.

Медиана Кемени и экспертные оценки. Рассмотрим частный случай пространств нечисловой природы - пространство бинарных отношений на конечном множестве $Q=\{q_1, q_2, ..., q_k\}$, и его подпространства. Как известно, каждое бинарное отношение A можно описать матрицей $||a(i,j)||$ из 0 и 1, причем $a(i,j) = 1$ тогда и только тогда q_i и q_j находятся в отношении A, и $a(i,j) = 0$ в противном случае.

Определение 4. Расстоянием Кемени между бинарными отношениями A и B, описываемыми матрицами $||a(i,j)||$ и $||b(i,j)||$ соответственно, называется

$$d(A,B)=\sum_{i,j=1}^{k}|a(i,j)-b(i,j)|.$$

Замечание. Иногда в определение расстояния Кемени вводят множитель, зависящий от k.

Как уже отмечалось, указанное расстояние введено американским исследователем Дж. Кемени в 1950-х годах и получило в нашей стране известность благодаря монографии [24], в которой оно получено для упорядочений (т.е. ранжировок, в которых допускаются связи, или кластеризованных ранжировок - см. главу 12) исходя из некоторой системы аксиом. Некоторое время казалось, что аксиоматический подход избавляет от субъективизма в выборе расстояния, а потому - от субъективизма в выборе способа усреднения бинарных отношений. Монография [24] породила поток работ, в которых с помощью различных систем аксиом вводились те или иные расстояния в пространствах объектов нечисловой природы (в обзоре [23] на эту тему - 161 ссылка на соответствующие публикации). В итоге произвол в выборе метрик отодвинут на уровень произвола в выборе систем аксиом.

Определение 5. Медианой Кемени для выборки, состоящей из бинарных отношений, называется эмпирическое среднее, построенное с помощью расстояния Кемени.

Поскольку число бинарных отношений на конечном множестве конечно, то эмпирические и теоретические средние для произвольных показателей различия существуют и справедливы законы больших чисел, описанные формулами (21) и (22) выше.

Бинарные отношения, в частности, упорядочения, часто используются для описания мнений экспертов. Тогда расстояние Кемени измеряет близость мнений экспертов, а медиана Кемени позволяет находить итоговое усредненное мнение комиссии экспертов. Расчет медианы Кемени обычно включают в информационное обеспечение систем принятия решений с использованием оценок экспертов. Речь идет, например, о математическом обеспечении автоматизированного рабочего места "Математика в экспертизе" (АРМ "МАТЭК"), предназначенного, в частности, для использования при проведении экспертиз в задачах экологического страхования. Поэтому представляет большой практический интерес численное изучение свойств медианы Кемени при конечном объеме выборки. Такое изучение дополняет описанную выше асимптотическую теорию, в которой объем выборки предполагается безгранично возрастающим ($n \to \infty$).

Компьютерное изучение свойств медианы Кемени при конечных объемах выборок. С помощью специально разработанной программной системы В.Н. Жихаревым был проведен ряд серий численных экспериментов по изучению свойств выборочных медиан Кемени. Представление о полученных результатах дается приводимой ниже табл.1, взятой из статьи [30]. В каждой серии методом статистических испытаний определенное число раз моделировался случайный и независимый выбор экспертных ранжировок, а затем находились все медианы Кемени для смоделированного набора мнений экспертов. При этом в сериях 1-5 распределение ответа эксперта предполагалось равномерным на множестве всех ранжировок, а в серии 6 это распределение являлось монотонным относительно расстояния Кемени с некоторым центром (о понятии монотонности см. выше), т.е. вероятность выбора определенной ранжировки убывала с увеличением расстояния Кемени этой ранжировки от центра. Таким образом, серии 1-5 соответствуют ситуации, когда у экспертов нет почвы для согласия, нет группировки их мнений относительно некоторого единого среднего группового мнения, в то время как в серии 6 есть единое мнение - описанный выше центр, к которому тяготеют ответы экспертов.

Результаты, приведенные в табл.1, можно комментировать разными способами. Неожиданным явилось большое число элементов в выборочной медиане Кемени - как среднее, так и особенно максимальное. Одновременно обращает на себя внимание убывание этих чисел при росте числа экспертов и особенно при переходе к ситуации реального существования группового мнения (серия 6). Достаточно часто один из ответов экспертов входит в медиану Кемени (т.е. пересечение множества ответов экспертов и медианы Кемени непусто), а диаметр медианы как множества в пространстве ранжировок заметно меньше диаметра множества ответов экспертов. По этим показателям - наилучшее положение в серии 6. Грубо говоря, всяческие "патологии" в поведении медианы Кемени наиболее резко проявляются в ситуации, когда ее применение не имеет содержательного обоснования, т.е. когда у экспертов нет основы для согласия, их ответы равномерно распределены на множестве ранжировок.

Увеличение числа испытаний в 10 раз при переходе от серии 1 к серии 5 не очень сильно повлияло на приведенные в таблице характеристики, поэтому представляется, что суть дела выявляется при числе испытаний (в методе Монте-Карло), равном 100 или даже 50. Увеличение числа объектов или экспертов увеличивает число элементов в рассматриваемом пространстве ранжировок, а потому уменьшается частота попадания какого-либо из мнений экспертов внутрь медианы Кемени, а также отношение диаметра медианы к диаметру множества экспертов, число элементов медианы Кемени (среднее и максимальное). Можно сказать, что увеличение числа объектов или экспертов уменьшает степень дискретности задачи, приближает ее к непрерывному случаю, а потому уменьшает выраженность различных "патологий".

Есть много интересных результатов, которые мы здесь не рассматриваем. Они связанны, в частности, со сравнением медианы Кемени с другими методами усреднения мнений экспертов, например, с нахождением итогового упорядочения по методу средних рангов, а также с использованием малых окрестностей ответов экспертов для поиска входящих в медиану ранжировок, с теоретической и численной оценкой скорости сходимости в законах больших чисел.

Табл.1. Вычислительный эксперимент по изучению свойств медианы Кемени

Номер серии	1	2	3	4	5	6
Число испытаний	100	1000	50	50	1000	1000
Количество объектов	5	5	7	7	5	5
Количество экспертов	10	30	10	30	10	10
Частота непустого пересечения	0,85	0,58	0,52	0,2	0,786	0,911
Среднее отношение диаметров	0.283	0,124	0,191	0,0892	0,202	0.0437
Средняя мощность медианы	5,04	2,41	6,4	2,88	3,51	1,35
Максимальная. мощность медианы	30	14	19	11	40	12

8.5. Непараметрические оценки плотности в пространствах произвольной природы

Математический аппарат статистики объектов нечисловой природы основан не на свойстве линейности пространства и использовании разнообразных сумм элементов выборок и функций от них, как в классической статистике, а на применении показателей различия, мер близости, метрик, поэтому существенно отличается от классического. В статистике нечисловых данных выделяют общую теорию и статистику в конкретных пространствах нечисловой природы (например, статистику ранжировок). В общей теории есть два основных сюжета. Один связан со средними величинами и асимптотическим поведением решений экстремальных статистических задач, второй - с непараметрическими оценками плотности. Первый сюжет только что рассмотрен, второму посвящена заключительная часть настоящей главы.

Понятие плотности в пространстве произвольной природы X требует специального обсуждения. В пространстве X должна быть выделена некоторая специальная мера μ, относительно которой будут рассматриваться плотности, соответствующие другим мерам, например, мере ν, задающей распределение вероятностей некоторого случайного элемента ξ. В таком случае $\nu(A) = P(\xi \in A)$ для любого случайного события A. Плотность $f(x)$, соответствующая мере ν - это такая функция, что

$$\nu(A) = \int_A f(x) d\mu$$

для любого случайного события A. Для случайных величин и векторов мера μ - это объем множества A, в математических терминах - мера Лебега. Для дискретных случайных величин и элементов со значениями в конечном множестве X в качестве меры μ естественно использовать считающую меру, которая событию A ставит в соответствие число его элементов. Используют также нормированную случайную меру, когда число точек в множестве A делят на число точек во всем пространстве X. В случае считающей меры значение плотности в точке x совпадает с вероятностью попасть в точку x, т.е. $f(x) = P(\xi = x)$. Таким образом, с рассматриваемой точки зрения стирается грань между понятиями «плотность вероятности» и «вероятность (попасть в точку)».

Как могут быть использованы непараметрические оценки плотности распределения вероятностей в пространствах нечисловой природы? Например, для решения задач классификации (диагностики, распознавания образов - см. главу 5). Зная плотности распределения классов, можно решать основные задачи диагностики - как задачи выделения кластеров, так и задачи отнесения вновь поступающего объекта к одному из диагностических классов. В задачах кластер-анализа можно находить моды плотности и принимать их за центры кластеров или за начальные точки итерационных методов типа k-средних или динамических сгущений. В задачах собственно диагностики (дискриминации, распознавания образов с учителем) можно принимать решения о диагностике объектов на основе отношения плотностей, соответствующих классам. При неизвестных плотностях представляется естественным использовать их состоятельные оценки.

Методы оценивания плотности вероятности в пространствах общего вида предложены и первоначально изучены в работе [31]. В частности, в задачах диагностики объектов нечисловой природы предлагаем использовать непараметрические ядерные оценки плотности типа Парзена - Розенблатта (этот вид оценок и его название впервые были введены в статье [31]). Они имеют вид:

$$f_n(x) = \frac{1}{\eta_n(h_n, x)} \sum_{1 \le i \le n} K\left(\frac{d(x_i, x)}{h_n}\right),$$

где K: $R^1_+ \to R^1$ - так называемая ядерная функция, $x_1, x_2, \ldots, x_n \in X$ - выборка, по которой оценивается плотность, $d(x_i\ ,\ x)$ - показатель различия (метрика, расстояние, мера близости) между элементом выборки x_i и точкой x, в которой оценивается плотность, последовательность h_n показателей размытости такова, что $h_n \to 0$ и $nh_n \to \infty$ при $n \to \infty$, а $\eta_n(h_n, x)$ - нормирующий множитель, обеспечивающий выполнение условия нормировки (интеграл по всему пространству от непараметрической оценки плотности $f_n(x)$ по мере μ должен равняться 1). Ранее американские исследователи Парзен и Розенблатт использовали подобные статистики в случае $X = R^1$ с $d(x_i, x) = x_i - x$.

Введенные описанным образом ядерные оценки плотности - частный случай так называемых линейных оценок, также впервые предложенных в работе [31]. В теоретическом плане они выделяются тем, что удается получать результаты такого же типа, что в классическом одномерном случае, но, разумеется, с помощью совсем иного математического аппарата.

Свойства непараметрических ядерных оценок плотности. Рассмотрим выборку со значениями в некотором пространстве произвольного вида. В этом пространстве предполагаются заданными показатель различия d и мера μ. Одна из основных идей рассматриваемого подхода состоит в том, чтобы согласовать их между собой. А именно, на их основе построим новый показатель различия d_1 , так называемый "естественный", в терминах которого проще формулируются свойства непараметрической оценки плотности. Для этого рассмотрим шары $L_t(x) = \{y \in X : d(y,x) \le t\}$ радиуса $t \ge 0$ и их меры $F_x(t) = \mu(L_t(x))$. Предположим, что $F_x(t)$ как функция t при фиксированном x непрерывна и строго возрастает. Введем функцию $d_1(x,y) = F_x(d(x,y))$. Это - монотонное преобразование показателя различия или расстояния, а потому $d_1(x,y)$ - также показатель различия (даже если d - *метрика,* для d_1 неравенство треугольника может быть не выполнено). Другими словами, $d_1(x,y)$, как и $d(x,y)$, можно рассматривать как показатель различия (меру близости) между x и y.

Для вновь введенного показателя различия $d_1(x,y)$ введем соответствующие шары $L_{1t}(x)=\{y\in X: d_1(y,x)\le t\}$. Поскольку обратная функция $F^{-1}_x(t)$ определена однозначно, то $L_{1t}(x)=\{y\in X: d_1(y,x)\le F_x^{-1}(t)\}=L_T(x)$, где $T = F^{-1}_x(t)$. Следовательно, справедлива цепочка равенств $F^1_x(t) = \mu(L_{1t}(x)) = \mu(L_T(x)) = F_x(F^{-1}_x(t)) = t$.

Переход от d к d_1 напоминает классическое преобразование, использованное Н.В. Смирновым при изучении непараметрических критериев согласия и однородности, а именно, преобразование $\eta = F(\xi)$, переводящее случайную величину ξ с непрерывной функцией распределения $F(x)$ в случайную величину η, равномерно распределенную на отрезке [0,1]. Оба рассматриваемых преобразования существенно упрощают дальнейшие рассмотрения. Преобразование $d_1= F_x(d)$ зависит от точки x, что не влияет на дальнейшие рассуждения, поскольку ограничиваемся изучением сходимости в отдельно взятой точке.

Функцию $d_1(x,y)$, для которой мера шара радиуса t равна t, называем в соответствии с работой [31] «естественным показателем различия» или «естественной метрикой». В случае конечномерного пространства R^k и евклидовой метрики d имеем $d_1(x,y) = c_k d^k(x,y)$, где c_k - объем шара единичного радиуса в R^k.

Поскольку можно записать, что

$$K\left(\frac{d(x_i,x)}{h_n}\right)=K_1\left(\frac{d_1(x_i,x)}{h_n}\right),$$

где

$$K_1(u)=K\left(\frac{F_x^{-1}(uh_n)}{h_n}\right),$$

то переход от одного показателя различия к другому, т.е. от d к d_1 соответствует переходу от одной ядерной функции к другой, т.е. от K к K_1. Выгода от такого перехода заключается в том, что утверждения о поведении непараметрических оценок плотности приобретают более простую формулировку.

Теорема 5. *Пусть d - естественная метрика, плотность f непрерывна в точке x и ограничена на всем пространстве X, причем $f(x)>0$, ядерная функция $K(u)$ удовлетворяет простым условиям регулярности*

$$\int_0^1 K(u)du=1, \int_0^\infty (|K(u)|+K^2(u))du<\infty \cdot$$

Тогда $\eta_n(h_n, x) = nh_n$, оценка $f_n(x)$ является состоятельной, т.е. $f_n(x)\to f(x)$ по вероятности при $n\to\infty$ и, кроме того,

$$\lim_{n\to\infty}(nh_n Df_n(x)) = f(x)\int_0^{+\infty} K^2(u)du.$$

Теорема 5 доказывается методами, развитыми в работе [31]. Однако остается открытым вопрос о скорости сходимости ядерных оценок, в частности, о поведении величины $\alpha_n = M(f_n(x)-f(x))^2$ - среднего квадрата ошибки, и об оптимальном выборе показателей размытости h_n. Для того, чтобы продвинуться в решении этого вопроса, введем новые понятия. Для случайного элемента $X(\omega)$ со значениями в X рассмотрим т.н. круговое распределение $G(x,t) = P\{d(X(\omega), x)\le t\}$ и круговую плотность $g(x,t)= G'_t(x,t)$.

Теорема 6. *Пусть ядерная функция K(u) непрерывна и финитна, т.е. существует число E такое, что K(u)=0 при u>E. Пусть круговая плотность является достаточно гладкой, т.е. допускает разложение*

$$g(x,t) = f(x) + tg_t^{'}(x,0) + \frac{t^2}{2} g_{tt}^{''}(x,0) + \frac{t^3}{3!} g_{ttt}^{'''}(x,0) + ... + \frac{t^k}{k!} g_{t^{(k)}}^{(k)}(x,0) + o(h_n^k)$$

при некотором k, причем остаточный член равномерно ограничен на [0,hE]. Пусть

$$\int_0^E u^i K(u)du = 0, i = 1,2,...,k-1.$$

Тогда

$$\alpha_n = [Mf_n(x) - f(x)]^2 + Df_n(x) =$$

$$= h_n^{2k} \left(\int_0^E u^k K(u)du \right)^2 (g_{t^{(k)}}^k(x,0))^2 + \frac{f(x)}{nh_n} \int_0^E K^2(u)du + o\left(h_n^{2k} + \frac{1}{nh_n} \right).$$

Доказательство теоремы 6 проводится с помощью разработанной в статистике объектов нечисловой природы математической техники, образцы которой представлены, в частности, в работе [31]. Если коэффициенты при основных членах в правой части последней формулы не равны 0, то величина α_n достигает минимума, равного $\alpha_n = O\left(n^{-1+\frac{1}{2k+1}} \right)$, при $h_n = n^{-\frac{1}{2k+1}}$. Эти выводы совпадают с классическими результатами, полученными ранее рядом авторов для весьма частного случая прямой $X = R^1$ (см., например, монографию [32, с.316]). Заметим, что для уменьшения смещения оценки приходится применять знакопеременные ядра *K(u)*.

Непараметрические оценки плотности в конечных пространствах. В случае конечных пространств естественных метрик не существует. Однако можно получить аналоги теорем 5 и 6, переходя к пределу не только по объему выборки *n*, но и по новому параметру дискретности *m*.

Рассмотрим некоторую последовательность X_m , m = 1,2,…- конечных пространств. Пусть в X_m заданы показатели различия d_m . Будем использовать нормированные считающие меры μ_m, ставящие в соответствие каждому подмножеству *A* долю элементов всего пространства X_m , входящих в *A*. Как и ранее, рассмотрим как функцию *t* объем шара радиуса *t*, т.е. $F_{mx}(t) = \mu_m(\{y \in X_m : d_m(x,y) \le t\})$. Введем аналог естественного показателя различия $d_{1m}(x,y) = F_{mx}(d_m(x,y))$. Наконец, рассмотрим аналоги преобразования Смирнова $F_{mx}^1(t) = \mu_m(\{y \in X_m : d_{1m}(x,y) \le t\})$. Функции $F_{mx}^1(t)$, в отличие от ситуации предыдущего раздела, уже не совпадают тождественно с *t*, они кусочно-постоянны и имеют скачки в некоторых точках t_i , i =1,2,…, причем в этих точках $F_{mx}^1(t_i) = t_i$.

Теорема 7. *Пусть точки скачков равномерно сближаются, т.е.* $\max(t_i - t_{i-1}) \to 0$ *при* $m \to \infty$ *(другими словами,* $\sup | F_{mx}^1(t) - t| \to 0$ *при* $m \to \infty$*). Тогда существует последовательность параметров дискретности* m_n *такая, что при предельном переходе* $n \to \infty, m \to \infty, m \ge m_n$ *справедливы заключения теорем 5 и 6.*

Пример 1. Пространство $X_m = 2^{\sigma(m)}$ всех подмножеств конечного множества $\sigma(m)$ из m элементов допускает (см. монографию [3]) аксиоматическое введение метрики $d(A,B) = card(A\Delta B)/2^m$, где Δ - символ симметрической разности множеств. Рассмотрим непараметрическую ядерную оценку плотности типа Парзена - Розенблатта

$$f_{nm}(A) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{1}{h_n}\Phi\left(\frac{2card(A\Delta X_i) - m}{\sqrt{m}}\right)\right),$$

где $\Phi(\cdot)$ - функция нормального стандартного распределения. Можно показать, что эта оценка удовлетворяет условиям теоремы 7 с $m_n = (\ln n)^6$.

Пример 2. Рассмотрим пространство функций $f: Y_r \to Z_q$, определенных на конечном множестве $Y_r = \{1/r,\ 2/r,...,(r-1)/r,\ 1\}$, со значениями в конечном множестве $Z_q = \{0,\ 1/q, 2/q,...,\ (q-1)/q,\ 1\}$.

Это пространство можно интерпретировать как пространство нечетких множеств (см. о нечетких множествах, например, монографии [3,10]), а именно, Y_r - носитель нечеткого множества, а Z_q - множество значений функции принадлежности. Очевидно, число элементов пространства X_m равно $(q+1)^r$. Будем использовать расстояние $d(f,g) = \sup|f(y) - g(y)|$. Непараметрическая оценка плотности имеет вид:

$$f_{nm}(x) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{[2\sup_y |x(y) - x_i(y)| + 1/q]^r}{h_n(1+1/q)^r}\right).$$

Если $r = n^\alpha$, $q = n^\beta$, то при $\beta > \alpha$ выполнены условия теоремы 7, а потому справедливы теоремы 5 и 6.

Пример 3. Рассматривая пространства ранжировок *m* объектов, в качестве расстояния *d(A,B)* между ранжировками *A* и *B* примем минимальное число инверсий, необходимых для перехода от А к В. Тогда $\max(t_i - t_{i-1})$ не стремится к 0 при $m \to \infty$, условия теоремы 7 не выполнены.

Пример 4. В прикладных работах наиболее распространенный пример объектов нечисловой природы – вектор разнотипных данных: реальный объект описывается вектором, часть координат которого - значения количественных признаков, а часть - качественных (номинальных и порядковых). Для пространств разнотипных признаков, т.е. декартовых произведений непрерывных и дискретных пространств, возможны различные постановки. Пусть, например, число градаций качественных признаков остается постоянным. Тогда непараметрическая оценка плотности сводится к произведению частоты попадания в точку в пространстве качественных признаков на классическую оценку Парзена-Розенблатта в пространстве количественных переменных. В общем случае расстояние *d(x,y)* можно, например, рассматривать как сумму трех расстояний. А именно, евклидова расстояния d_1 между количественными факторами, расстояния d_2 между номинальными признаками ($d_2(x,y) = 0$, если $x = y$, и $d_2(x,y) = 1$, если $x \neq y$) и расстояния d_3 между порядковыми переменными (если *x* и *y* - номера градаций, то $d_3(x,y) = |x - y|$). Наличие количественных факторов приводит к непрерывности и строгому возрастанию функции $F_{mx}(t)$, а потому для непараметрических оценок плотности в пространствах разнотипных признаков верны теоремы 5 - 6.

Статистика объектов нечисловой природы как часть эконометрики продолжает бурно развиваться. Увеличивается количество ее практически полезных применений при анализе конкретных экономических данных - в маркетинговых исследованиях, контроллинге, при управлении предприятием и др.

Цитированная литература

1. Шубкин В.П. Социологические опыты. - М.: Мысль,1970.-256 с.
2. Щукина Г.И. Проблема познавательного интереса в педагогике. - М.: Педагогика, 1971.-352 с.
3. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука,1979.-296 с.
4. **Орлов А.И. Статистика объектов нечисловой природы (Обзор). - Журнал «Заводская лаборатория». 1990. Т.56. No.3. С.76-83.**
5. **Орлов А.И. Объекты нечисловой природы. - Журнал «Заводская лаборатория». 1995. Т.61. No.3. С.43-52.**
6. Кендэл М. Ранговые корреляции. - М.:Статистика,1975. - 216 с.
7. Беляев Ю.К. Вероятностные методы выборочного контроля. - М.: Наука, 1975. - 408 с.
8. Лумельский Я.П. Статистические оценки результатов контроля качества. - М.: Изд-во стандартов, 1979. - 200 с.
9. Дэвид Г. Метод парных сравнений. - М.: Статистика, 1978.- 144 с.
10. Орлов А.И. Задачи оптимизации и нечеткие переменные. - М.: Знание, 1980. - 64с.
10. Кендалл М.Дж., Стьюарт А., Статистические выводы и связи. М.: Наука, 1973. - 900 с.
11. Себер Дж. Линейный регрессионный анализ. - М.: Мир, 1980. - 456 с.
12. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Основы моделирования и первичная обработка данных. - М.: Финансы и статистика. 1983. - 472 с.
13. Борель Э. Вероятность и достоверность. - М.: ГИФМЛ, 1961. - 120 с.
14. **Орлов А.И. Вероятностные модели конкретных видов объектов нечисловой природы. - Журнал «Заводская лаборатория». 1995. Т.61. No.5. С.43-51.**
15. Вероятность и математическая статистика: Энциклопедия / Гл. ред. Ю.В. Прохоров. - М.: Большая Российская энциклопедия, 1999. - 910 с.
16. Орлов А.И. Статистика объектов нечисловой природы и экспертные оценки. – В сб.: Экспертные оценки / Вопросы кибернетики. Вып.58. - М.: Научный Совет АН СССР по комплексной проблеме "Кибернетика", 1979. - С.17-33.
17. Орлов А.И. Случайные множества с независимыми элементами (люсианы) и их применения. – В сб.: Алгоритмическое и программное обеспечение прикладного статистического анализа. Ученые записки по статистике, т.36. - М.: Наука, 1980. - С. 287-308.
18. **Орлов А.И. Парные сравнения в асимптотике Колмогорова. - В сб.: Экспертные оценки в задачах управления. - М.: Изд-во Института проблем управления АН СССР, 1982. - С. 58-66.**

19. Орлов А.И. Логистическое распределение. – В сб.: Математическая энциклопедия. Т.3. - М.: Советская энциклопедия, 1982. - С.414.
20. Орлов А.И. О нецелесообразности использования итеративных процедур нахождения оценок максимального правдоподобия. - Журнал «Заводская лаборатория». 1986. Т.52. No.5. С.67-69.
21. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. - М.: Наука, 1983 (3-е изд.). - 474 с.
22. Тюрин Ю.Н., Василевич А.П., Андрукович П.Ф. Статистические модели ранжирования. - В сб.: Статистические методы анализа экспертных оценок. - М.: Наука, 1977. - С.30-58.
23. Раушенбах Г.В. Меры близости и сходства. - В сб.: Анализ нечисловой информации в социологических исследованиях. - М.: Наука, 1985. - С.169-203.
24. Кемени Дж., Снелл Дж. Кибернетическое моделирование: Некоторые приложения. - М.: Советское радио, 1972. - 192 с.
25. Прохоров Ю.В., Розанов Ю.А. Теория вероятностей (Основные понятия. Предельные теоремы. Случайные процессы). - М.: Наука, 1973.- 496 с.
26. Кривцов В.С., Орлов А.И., Фомин В.Н. Современные статистические методы в стандартизации и управлении качеством продукции. – Журнал «Стандарты и качество». 1988. No.3. С.32-36.
27. Орлов А.И. Асимптотика решений экстремальных статистических задач. - В сб.: Анализ нечисловых данных в системных исследованиях. Сборник трудов. Вып.10. - М.: Всесоюзный научно-исследовательский институт системных исследований, 1982. - С. 4-12.
28. Орлов А.И. Асимптотическое поведение статистик интегрального типа. - В сб.: Вероятностные процессы и их приложения. Межвузовский сборник. - М.: МИЭМ, 1989. С.118-123.
29. Келли Дж. Общая топология. - М.: Наука, 1968. - 384 с.
30. Жихарев В.Н., Орлов А.И. Законы больших чисел и состоятельность статистических оценок в пространствах произвольной природы. - В сб.: Статистические методы оценивания и проверки гипотез. Межвузовский сборник научных трудов. - Пермь: Изд-во Пермского государственного университета, 1998. С.65-84.
31. Орлов А.И. Непараметрические оценки плотности в топологических пространствах. - В сб.: Прикладная статистика. Ученые записки по статистике, т.45. - М.: Наука, 1983. - С. 12-40.
32. Ибрагимов И.А., Хасьминский Р.З. Асимптотическая теория оценивания. - М.: Наука, 1979. - 528 с.

Глава 9. Статистика интервальных данных

В статистике интервальных данных, как части статистики нечисловых данных, элементы выборки - не числа, а интервалы. Это приводит к алгоритмам и выводам, принципиально отличающимся от классических. В главе 9 рассмотрены основные идеи и подходы асимптотической статистики интервальных данных, приведены результаты, связанные с основополагающими в рассматриваемой области эконометрики понятиями нотны и рационального объема выборки.

9.1. Основные идеи статистики интервальных данных

Перспективная и быстро развивающаяся область статистических исследований последних лет - статистика интервальных данных. Речь идет о развитии эконометрических методов в ситуации, когда статистические данные - не числа, а интервалы, в частности, порожденные наложением ошибок измерения на значения случайных величин.

В настоящее время признается необходимым изучение устойчивости (робастности) оценок параметров к малым отклонениям исходных данных и предпосылок модели. Однако популярная среди теоретиков (см. ниже в главе 10) модель засорения (Тьюки-Хьюбера) представляется не вполне адекватной. Эта модель нацелена на изучение влияния больших "выбросов". Поскольку любые реальные измерения лежат в некотором фиксированном диапазоне, а именно, заданном в техническом паспорте средства измерения, то зачастую выбросы не могут быть слишком большими. Поэтому представляются полезными иные, более общие схемы устойчивости, в частности, рассмотренные в главе 10 ниже, в которых, например, учитываются отклонения распределений результатов наблюдений от предположений модели.

В одной из таких схем изучается влияние интервальности исходных данных на статистические выводы. Необходимость такого изучения стала для нас очевидной следующим образом. В государственных стандартах СССР по прикладной статистике в обязательном порядке давалось справочное приложение "Примеры применения правил стандарта". При разработке ГОСТ 11.011-83 (см. издание [1]) нам были переданы для анализа реальные данные о наработке резцов до предельного состояния (в часах). Оказалось, что все эти данные представляли собой либо целые числа, либо полуцелые (т.е. после умножения на 2 становящиеся целыми). Ясно, что исходная длительность наработок резцов до отказа искажена. Необходимо учесть в статистических процедурах наличие такого искажения исходных данных. Как это сделать?

Первое, что приходит в голову - модель группировки данных, согласно которой для истинного случайного значения X (мысленно) проводится замена на ближайшее число из множества $\{0{,}5n,\ n=1,2,3,...\}$. Однако эту модель нельзя принимать без обсуждения, ее целесообразно подвергнуть сомнению, а также рассмотреть иные модели. Так, возможно, что X надо приводить к ближайшему сверху элементу указанного множества - если проверка качества поставленных на испытание резцов проводилась раз в полчаса. Другой вариант модели: если расстояния от X до двух ближайших элементов множества $\{0{,}5n,\ n=1,2,3,...\}$ примерно равны, то естественно ввести рандомизацию при выборе заменяющего числа, и т.д.

Наиболее адекватной представляется новая эконометрическая модель, согласно которой **результаты наблюдений - не числа, а интервалы**. Например, если в таблице приведено значение 53,5, то это значит, что реальное значение - какое-то число от 53,0 до 54,0, т.е. какое-то число в интервале [53,5-0,5; 53,5+0,5], где 0,5 - максимально возможная погрешность. Принимая эту модель, мы попадаем в научную область под названием "статистика интервальных данных". Она идейно связана с интервальной математикой, в которой в роли чисел выступают интервалы (см., например, монографию [2] академика РАН Ю.И. Шокина). Это направление математики является дальнейшим развитием всем известных правил приближенных вычислений, посвященных выражению погрешностей суммы, разности, произведения, частного через погрешности тех чисел, над которыми осуществляются перечисленные операции. Как видно из сборника трудов Международной конференции по интервальным и стохастическим методам в науке и технике (ИНТЕРВАЛ-92), к настоящему времени удалось решить, в частности, ряд задач теории интервальных дифференциальных уравнений, в которых коэффициенты, начальные условия и решения описываются с помощью интервалов. По мнению ряда специалистов, статистика интервальных данных является частью интервальной математики [7]. Впрочем, есть другая точка зрения, согласно которой такое включение нецелесообразно, поскольку статистика интервальных данных использует несколько иные подходы к алгоритмам анализа реальных данных, чем сложившиеся в интервальной математике (подробнее см. ниже).

Общее описание направлений статистического анализа интервальных данных. Ниже развиваются асимптотические методы статистического анализа интервальных данных при больших объемах выборок и малых погрешностях измерений. В отличие от классической математической статистики, сначала устремляется к бесконечности объем выборки и только потом - уменьшаются до нуля погрешности. В частности, еще в начале 1980-х годов с помощью такой асимптотической теории были сформулированы правила выбора метода оценивания параметров гамма-распределения в ГОСТ 11.011-83 [1].

Разработана общая схема исследования, включающая расчет нотны (максимально возможного отклонения статистики, вызванного интервальностью исходных данных) и рационального объема выборки (превышение которого не дает существенного повышения точности оценивания). Она применена к оцениванию математического ожидания и дисперсии, медианы и коэффициента вариации, параметров гамма-распределения и характеристик аддитивных статистик, при проверке гипотез о параметрах нормального распределения, в т.ч. с помощью критерия Стьюдента, а также гипотезы однородности с помощью критерия Смирнова. Изучено асимптотическое поведение оценок метода моментов и оценок максимального правдоподобия (а также более общих - оценок минимального контраста), проведено асимптотическое сравнение этих методов в случае интервальных данных, найдены общие условия, при которых, в отличие от классической математической статистики, метод моментов дает более точные оценки, чем метод максимального правдоподобия. Разработаны подходы к рассмотрению интервальных данных в основных постановках регрессионного, дискриминантного и кластерного анализов. В частности, изучено влияние погрешностей измерений и наблюдений на свойства алгоритмов регрессионного анализа, разработаны способы расчета нотн и рациональных объемов выборок, введены и исследованы новые понятия многомерных и асимптотических нотн, доказаны соответствующие предельные теоремы. Начата разработка интервального дискриминантного анализа, в частности, рассмотрено влияние интервальности данных на показатель качества классификации.

Как показала, в частности, международная конференция ИНТЕРВАЛ-92, в области асимптотической математической статистики интервальных данных российская научная школа имеет мировой приоритет. По нашему мнению, со временем во все виды статистического программного обеспечения должны быть включены алгоритмы интервальной статистики, "параллельные" обычно используемым алгоритмам прикладной математической статистики. Это позволит в явном виде учесть наличие погрешностей у результатов наблюдений, сблизить позиции метрологов и статистиков.

Многие из утверждений статистики интервальных данных весьма отличаются от аналогов из классической математической статистики. В частности, не существует состоятельных оценок; средний квадрат ошибки оценки, как правило, асимптотически равен сумме дисперсии оценки, рассчитанной согласно классической теории, и некоторого положительного числа (равного квадрату т.н. нотны - максимально возможного отклонения значения статистики из-за погрешностей исходных данных) - в результате метод моментов оказывается иногда точнее метода максимального правдоподобия; нецелесообразно увеличивать объем выборки сверх некоторого предела (называемого рациональным объемом выборки) - вопреки классической теории, согласно которой чем больше объем выборки, тем точнее выводы.

История развития статистики интервальных данных противоречива. Так, в стандарт [1] был включен специальный раздел 5, посвященный выбору метода оценивания при неизвестных параметрах формы и масштаба и известном параметре сдвига, он был основан на концепциях статистики интервальных данных. Однако теоретическое обоснование этого раздела стандарта было опубликовано лишь через 5 лет. Следует отметить, что хотя в 1982 г. при разработке стандарта [1] уже были найдены основные идеи статистики интервальных данных, однако они не были полностью реализованы в нормативном документе (ГОСТ 11.011-83), и этот стандарт написан в основном в классической манере. Развитие идей статистики интервальных данных продолжается уже в течение 20 лет, и еще много чего надо сделать! Большое значение статистики интервальных данных для современной прикладной статистики обосновано в статье [3].

Одна из ведущая научная школа в области статистики интервальных данных - это школа проф. А.П. Вощинина, активно работающая с конца 70-х годов. Полученные результаты отражены в ряде монографий (см., в частности, [4-6]), статей [7], научных докладов, в том числе в трудах Международной конференции ИНТЕРВАЛ-92, диссертаций. В частности, изучены проблемы регрессионного анализа, планирования эксперимента, сравнения альтернатив и принятия решений в условиях интервальной неопределенности. Рассмотренное ниже направление исследований отличается нацеленностью на асимптотические результаты, полученные при больших объемах выборок и малых погрешностях измерений, поэтому оно и названо **асимптотической статистикой интервальных данных.**

Сформулируем сначала основные идеи асимптотической математической статистики интервальных данных, а затем рассмотрим реализацию этих идей на некоторых из перечисленных выше примеров. Следует сразу подчеркнуть, что основные идеи достаточно просты, в то время как их проработка в конкретных ситуациях зачастую оказывается достаточно трудоемкой.

Основные понятия асимптотической математической статистики интервальных данных. Пусть существо реального явления описывается выборкой x_1 , x_2 , ..., x_n . В вероятностной теории математической статистики, из которой мы исходим (см. приложение 1 в конце книги), выборка - это набор независимых в совокупности одинаково распределенных случайных величин. Однако беспристрастный и тщательный анализ подавляющего большинства реальных задач показывает, что статистику известна отнюдь не выборка x_1 , x_2 , ..., x_n , а величины

$$y_j = x_j + \varepsilon_j \ , \ j = 1, 2, \dots , n ,$$

где $\varepsilon_1, \varepsilon_2, ..., \varepsilon_n$ – некоторые погрешности измерений, наблюдений, анализов, опытов, исследований (например, инструментальные ошибки).

Одна из причин появления погрешностей - запись результатов наблюдений с конечным числом значащих цифр. Дело в том, что для случайных величин с непрерывными функциями распределения событие, состоящее в попадании хотя бы одного элемента выборки в множество рациональных чисел, согласно правилам теории вероятностей имеет вероятность 0, а такими событиями в теории вероятностей принято пренебрегать. Поэтому при рассуждениях о выборках из нормального, логарифмически нормального, экспоненциального, равномерного, гамма - распределений, распределения Вейбулла-Гнеденко и др. приходится принимать, что эти распределения имеют элементы исходной выборки x_1 , x_2 , ..., x_n , в то время как статистической обработке доступны лишь искаженные значения $y_j = x_j + \varepsilon_j$.

Введем обозначения

$$x = (x_1, x_2, ..., x_n), \ y = (y_1, y_2, ..., y_n), \ \varepsilon = \varepsilon_1 + \varepsilon_2 + ... + \varepsilon_n .$$

Пусть статистические выводы основываются на статистике $f : R^n \to R^1$, используемой для оценивания параметров и характеристик распределения, проверки гипотез и решения иных статистических задач. Принципиально важная для статистики интервальных данных идея такова: СТАТИСТИК ЗНАЕТ ТОЛЬКО $f(y)$, НО НЕ $f(x)$.

Очевидно, в статистических выводах необходимо отразить различие между $f(y)$ и $f(x)$. Одним из двух основных понятий статистики интервальных данных является понятие нотны.

Определение. Величину максимально возможного (по абсолютной величине) отклонения, вызванного погрешностями наблюдений ε, известного статистику значения $f(y)$ от истинного значения $f(x)$, т.е.

$$N_f(x) = \sup | f(y) - f(x) | ,$$

где супремум берется по множеству возможных значений вектора погрешностей ε (см. ниже), будем называть НОТНОЙ. .

Если функция f имеет частные производные второго порядка, а ограничения на погрешности имеют вид

$$| \varepsilon_i | \le \Delta, \ i = 1,2,..., n, \qquad \textbf{(1)}$$

причем Δ мало, то можно показать, что *нотна* с точностью до бесконечно малых более высокого порядка имеет вид

$$N_f(x) = (\sum_{i=1}^{n} | \partial f / \partial x_i |) \Delta .$$

Условие (1) означает, что исходные данные представляются статистику в виде интервалов $[y_i - \Delta; y_i + \Delta], i = 1,2,..., n$ (отсюда и название этого научного направления). Ограничения на погрешности могут задаваться разными способами - кроме абсолютных ошибок используются относительные или иные показатели различия между x и y.

Основные результаты в вероятностной модели. В классической вероятностной модели имеют элементы исходной выборки x_1, x_2, ..., x_n рассматриваются как независимые одинаково распределенные случайные величины. Как правило, существует некоторая константа $C > 0$ такая, что в смысле сходимости по вероятности

$$\lim_{n\to\infty} N_f(x) = C\Delta. \qquad (2)$$

Соотношение (2) доказывается отдельно для каждой конкретной задачи.

При использовании классических эконометрических методов в большинстве случаев используемая статистика $f(x)$ является асимптотически нормальной. Это означает, что существуют константы a и σ^2 такие, что

$$\lim_{n\to\infty} P\left(\sqrt{n}\,\frac{f(x)-a}{\sigma} < x\right) = \Phi(x),$$

где $\Phi(x)-$ функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. При этом обычно оказывается, что

$$\lim_{n\to\infty} \sqrt{n}(Mf(x)-a) = 0$$

и

$$\lim_{n\to\infty} nDf(x) = \sigma^2,$$

а потому в классической эконометрике средний квадрат ошибки статистической оценки равен

$$M(f(x)-a)^2 = (Mf(x)-a)^2 + Df(x) = \frac{\sigma^2}{n}$$

с точностью до членов более высокого порядка.

В статистике интервальных данных ситуация совсем иная - обычно можно доказать, что средний квадрат ошибки равен

$$\max_{\{\varepsilon\}} M(f(y)-a)^2 = \frac{\sigma^2}{n} + N_f^2(y) + o(\Delta^2 + \frac{1}{n}). \qquad (3)$$

Из соотношения (3) можно сделать ряд важных следствий. Прежде всего отметим, что правая часть этого равенства, в отличие от правой части соответствующего классического равенства, не стремится к 0 при безграничном возрастании объема выборки. Она остается больше некоторого положительного числа, а именно, квадрат нотны. Следовательно, статистика $f(x)$ не является состоятельной оценкой параметра a. Более того, состоятельных оценок вообще не существует.

Пусть доверительным интервалом для параметра a, соответствующим заданной доверительной вероятности γ, в классической математической статистике является интервал $(c_n(\gamma); d_n(\gamma))$. В статистике интервальных данных аналогичный доверительный интервал является более широким. Он имеет вид $(c_n(\gamma) - N_f(y); d_n(\gamma) + N_f(y))$. Таким образом, его длина увеличивается на две нотны. Следовательно, при увеличении объема выборки длина доверительного интервала не может стать меньше, чем $2C\Delta$ (см. формулу (2)).

В статистике интервальных данных методы оценивания параметров имеют другие свойства по сравнению с классической математической статистикой. Так, при больших объемах выборок метод моментов может быть заметно лучше, чем метод максимального правдоподобия (т.е. иметь меньший средний квадрат ошибки - см. формулу (3)), в то время как в классической математической статистике второй из названных методов всегда не хуже первого.

Рациональный объем выборки. Анализ формулы (3) показывает, что в отличие от классической математической статистики нецелесообразно безгранично увеличивать объем выборки, поскольку средний квадрат ошибки остается всегда большим квадрата нотны. Поэтому представляется полезным ввести понятие "рационального объема выборки" n_{rat}, при достижении которого продолжать наблюдения нецелесообразно.

Как установить "рациональный объем выборки"? Можно воспользоваться идеей "принципа уравнивания погрешностей", выдвинутой в монографии [8]. Речь идет о том, что вклад погрешностей различной природы в общую погрешность должен быть примерно одинаков. Этот принцип дает возможность выбирать необходимую точность оценивания тех или иных характеристик в тех случаях, когда это зависит от исследователя. В статистике интервальных данных в соответствии с "принципом уравнивания погрешностей" предлагается определять рациональный объем выборки n_{rat} из условия равенства двух величин - метрологической составляющей, связанной с нотной, и статистической составляющей - в среднем квадрате ошибки (3), т.е. из условия

$$\frac{\sigma^2}{n_{rat}} = N_f^2(y), \quad n_{rat} = \frac{\sigma^2}{N_f^2(y)}.$$

Для практического использования выражения для рационального объема выборки неизвестные теоретические характеристики необходимо заменить их оценками. Это делается в каждой конкретной задаче по-своему.

Исследовательскую программу в области статистики интервальных данных можно "в двух словах" сформулировать так: для любого эконометрического алгоритма анализа данных (алгоритма прикладной статистики) необходимо вычислить нотну и рациональный объем выборки (или иные величины из того же понятийного ряда, возникающие в многомерном случае, при наличии нескольких выборок и при иных обобщениях описываемой здесь простейшей схемы). Затем проследить влияние погрешностей исходных данных на точность оценивания, доверительные интервалы, значения статистик критериев при проверке гипотез, уровни значимости и другие характеристики статистических выводов. Очевидно, классическая математическая статистика является частью статистики интервальных данных, выделяемой условием $\Delta = 0$.

9.2. Примеры статистического анализа интервальных данных

Поясним теоретические концепции статистики интервальных данных на простых примерах.

Пример 1. Оценивание математического ожидания. Пусть необходимо оценить математическое ожидание случайной величины с помощью обычной оценки (см. главу 4) - среднего арифметического результатов наблюдений, т.е.

$$f(x) = \frac{x_1 + x_2 + \ldots + x_n}{n}.$$

Тогда $N_f(x) = \Delta$. Таким образом, нотна полностью известна и не зависит от многомерной точки, в которой берется. Вполне естественно: если каждый результат наблюдения известен с точностью до Δ, то и среднее арифметическое известно с той же точностью. Ведь возможна систематическая ошибка - если к каждому результату наблюдению добавить Δ, то и среднее арифметическое увеличится на Δ.

Поскольку

$$D(\bar{x}) = \frac{D(x_1)}{n},$$

то в обозначениях предыдущего пункта

$$\sigma^2 = D(x_1).$$

Следовательно, рациональный объем выборки равен

$$n_{rat} = \frac{D(x_1)}{\Delta^2}.$$

Для практического использования полученной формулы надо оценить дисперсию результатов наблюдений. Можно доказать, что, поскольку Δ мало, это можно сделать обычным способом, например, с помощью несмещенной выборочной оценки дисперсии

$$s^2(y) = \frac{1}{n-1} \sum_{1 \le i \le n} (y_i - \bar{y})^2.$$

Здесь и далее рассуждения часто идут на двух уровнях. Первый - это уровень "истинных" случайных величин, обозначаемых *"x"*, описывающих реальность, но неизвестных эконометрику. Второй - уровень известных эконометрику величин *"y"*, отличающихся погрешностями от истинных. Погрешности малы, поэтому функции от x отличаются от функций от y на некоторые бесконечно малые величины. Эти соображения и позволяют нам использовать $s^2(y)$ как оценку $D(x_1)$.

Итак, выборочной оценкой рационального объема выборки является

$$n_{sample-rat} = \frac{s^2(y)}{\Delta^2}.$$

Уже на этом первом рассматриваемом примере видим, что рациональный объем выборки находится не где-то вдали, а непосредственно рядом с теми объемами, с которыми имеет дело любой практически работающий эконометрик. Например, если статистик знает, что $\Delta = \frac{\sigma}{6}$, то n_{rat} = 36. А именно такова погрешность контрольных шаблонов во многих технологических процессах! Поэтому, занимаясь эконометрикой качества (см. главу 13), обратите внимание и на действующую на предприятии систему измерений.

По сравнению с главой 4 доверительный интервал для математического ожидания (для заданной доверительной вероятности γ) имеет другой вид:

$$(\bar{y} - \Delta - u(\gamma)\frac{s}{\sqrt{n}}; \bar{y} + \Delta + u(\gamma)\frac{s}{\sqrt{n}}), \qquad (4)$$

где $u(\gamma)$ - квантиль порядка (1+ γ)/2 стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1..

По поводу формулы (4) была довольно жаркая дискуссия среди специалистов. Отмечалось, что она получена на основе Центральной Предельной Теоремы теории вероятностей и может быть использована при любом распределении результатов наблюдений (с конечной дисперсией). Если же имеется дополнительная информация, то, по мнению отдельных специалистов, формула (4) может быть уточнена. Например, если известно, что распределение x_i является нормальным, в качестве $u(\gamma)$ целесообразно использовать квантиль распределения Стьюдента. К этому надо добавить, что по небольшому числу наблюдений нельзя надежно установить нормальность, а при росте объема выборки квантили распределения Стьюдента приближаются к квантилям нормального распределения. Вопрос о том, часто ли результаты наблюдений имеют нормальное распределение, подробно обсуждался в начале главы 4.

Пример 2. Оценивание дисперсии. Для статистики $f(y) = s^2(y)$, где $s^2(y)$ - выборочная дисперсия (несмещенная оценка теоретической дисперсии), имеем

$$N_f(y) = \frac{2\Delta}{n-1}\sum_{i=1}^{n} | y_i - \bar{y} | + O(\Delta^2).$$

Можно показать, что нотна $N_f(y)$ сходится к

$$2\Delta M \mid x_1 - M(x_1) \mid$$

по вероятности с точностью до $o(\Delta)$, когда n стремится к бесконечности. Это же предельное соотношение верно и для нотны $N_f(x)$, вычисленной для исходных данных. Таким образом, в данном случае справедлива формула (2) с

$$C = 2M \mid x_1 - M(x_1) \mid .$$

Известно что случайная величина

$$\frac{s^2 - \sigma^2}{\sqrt{n}}$$

является асимптотически нормальной с математическим ожиданием 0 и дисперсией $D(x_1^2)$. Этот факт использовался в главе 4 для построения асимптотического доверительного интервала для дисперсии.

Из сказанного вытекает, что в статистике интервальных данных асимптотический доверительный интервал для дисперсии σ^2 (соответствующий доверительной вероятности γ) имеет вид

$$(s^2(y) - A; \quad s^2 + A),$$

где

$$A = \frac{u(\gamma)}{\sqrt{n(n-1)}}\sqrt{\sum_{i=1}^{n}(y_i^2 - \frac{1}{n}\sum_{j=1}^{n} y_j^2)^2} + \frac{2\Delta}{n-1}\sum_{i=1}^{n} | y_i - \bar{y} |,$$

где $u(\gamma)$ обозначает тот же самый квантиль стандартного нормального распределения, что и выше в случае оценивания математического ожидания.

Рациональный объем выборки для дисперсии равен

$$n_{rat} = \frac{D(x_1^2)}{4\Delta^2 (M \mid x_1 - M(x_1) \mid)^2},$$

а выборочную оценку рационального объема выборки $n_{sample-rat}$ можно вычислить, заменяя теоретические моменты на соответствующие выборочные и используя доступные эконометрику результаты наблюдений, содержащие погрешности.

Что можно сказать о численной величине рационального объема выборки? Как и в случае оценивания математического ожидания, она отнюдь не выходит за пределы обычно используемых объемов выборок. Так, если распределение результатов наблюдений x_i является нормальным с математическим ожиданием 0 и дисперсией σ^2, то в результате вычисления моментов случайных величин в предыдущей формуле получаем, что

$$n_{rat} = \frac{\sigma^2}{\pi\Delta^2}.$$

Например, если $\Delta = \sigma / 6$, то $n_{rat} = 11$. Это меньше, чем при оценивании математического ожидания в предыдущем примере.

9.3. Статистика интервальных данных и оценки погрешностей характеристик финансовых потоков инвестиционных проектов

Методы статистики интервальных данных оказываются полезными не только в традиционных эконометрических задачах, но и во многих других областях экономики и менеджмента, например, в инновационном менеджменте.

Основная идея формулируется так. Все знают, что любое инженерное измерение проводится с некоторой погрешностью. Эту погрешность обычно приводят в документации и учитывают при принятии решений. Ясно, что и любое экономическое измерение также проводится с погрешностью. А вот какова она? Необходимо уметь ее оценивать, поскольку ошибки при принятии экономических решений обходятся дорого.

Например, как принимать решение о выгодности или невыгодности инвестиционного проекта? Как сравнивать инвестиционныепроекты между сообой? Как известно, для решения этих задач используют такие экономические характеристики, как NPV (Net Present Value) - чистая текущая стоимость (этот термин переводится с английского также как чистый дисконтированный доход, чистое приведенное значение и др.), внутренняя норма доходности, срок окупаемости, показатели рентабельности и др.

С экономической точки зрения инвестиционные проекты описываются финансовыми потоками, т.е. функциями от времени, значениями которых являются платежи (и тогда значения этих функций отрицательны) и поступления (значения функций положительны). Сравнение инвестиционных проектов - это сравнение функций от времени с учетом внешней среды, проявляющейся в виде дисконт-функции (как результата воздействия СТЭП-факторов), и представлений законодателя или инвестора - обычно ограничений на финансовые потоки платежей и на горизонт планирования. Основная проблема при сравнении инвестиционных проектов такова: *что лучше - меньше, но сейчас, или больше, но потом*? Как правило, чем больше вкладываем сейчас, тем больше получаем в более или менее отдаленном будущем. Вопрос в том, достаточны ли будущие поступления, чтобы покрыть нынешние платежи и дать приемлемую для инвестора прибыль?

В настоящее время широко используются различные теоретические подходы к сравнению инвестиционных проектов и облегчающие расчеты компьютерные системы, в частности, Project Expert, COMFAR, PROPSIN, Альт-Инвест, ТЭО-ИНВЕСТ. Однако ряд важных моментов в них не учтен.

Введем основные понятия. Дисконт-функция как функция от времени показывает, сколько стоит для фирмы 1 руб. в заданный момент времени, если его привести к начальному моменту. Если дисконт-функция - константа для разных отраслей, товаров и проектов, то эта константа называется дисконт-фактором, или просто дисконтом. Дисконт-функция определяется совместным действием различных факторов, в частности, реальной процентной ставки и индекса инфляции. Реальная процентная ставка описывает "нормальный" рост экономики (т.е. без инфляции). В стабильной ситуации доходность от вложения средств в различные отрасли, в частности, в банковские депозиты, примерно одинакова.

Сейчас она, по оценке ряда экспертов, около 12%. Итак, нынешний 1 руб. превращается в 1,12 руб. через год, а потому 1 руб. через год соответствует 1/1,12 = 0,89 руб. сейчас - это и есть максимум дисконта.

Обозначим дисконт буквой *C*. Если *q* - банковский процент (плата за депозит), т.е. вложив в начале года в банк 1 руб., в конце года получим (1+ *q*) руб., то дисконт определяется по формуле C=1/(1+q). При таком подходе полагают, что банковские проценты одинаковы во всех банках. Более правильно было бы считать *q*, а потому и *C*, нечисловыми величинами, а именно, интервалами $[q_1;\ q_2]$ и $[C_1;\ C_2]$. Следовательно, экономические выводы должны быть исследованы на *устойчивость* (применяют и термин *"чувствительность"*) по отношению к возможным отклонениям.

Как функцию времени *t* дисконт-функцию обозначим *C(t)*. При постоянстве дисконт-фактора имеем $C(t) = C^t$. Если *q* = 0,12, *C* = 0,89, то 1 руб. за 2 года превращается в $1{,}12^2 = 1{,}2544$, через 3 - в 1,4049. Итак, 1 руб., получаемый через 2 года, соответствует 1/1,2544=0,7972 руб., т.е. 79,72 коп. сейчас, а 1 руб., обещанный через 3 года, соответствует 0,71 руб. сейчас. Другими словами, *C(2)* = 0,80, а *C(3)* = 0,71. Если дисконт-фактор зависит от времени, в первый год равен C_1, во второй - C_2, в третий - C_3,..., в *t*-ый год - C_t, то $C(t)=C_1C_2C_3...C_t$. .

Рассмотрим характеристики потоков платежей. Срок окупаемости - тот срок, за который доходы покроют расходы. Обычно предполагается, что после этого проект приносит только прибыль. Это верно не всегда. Простейший вариант, для которого не возникает никаких парадоксов, состоит в том, что все инвестиции (капиталовложения) делаются сразу, в начале, а затем инвестор получает только доход. Сложности возникают, если проект состоит из нескольких очередей, вложения распределены во времени. Тогда, например, понятие "срок окупаемости" может быть денежных единиц со временем, т.е. не учитывает дисконтирование. Если неоднозначно - вслед за окупаемостью первой очереди может придти очередь затрат на вторую очередь проекта…

Примитивный способ расчета срока окупаемости состоит в делении объема вложений *A* на ожидаемый ежегодный доход *B*. Тогда срок окупаемости равен *A/B*. Этот способ падение стоимости дисконт-фактор равен *C,* то максимально возможный суммарный доход равен

$$BC+BC^2+BC^3+BC^4+BC^5+...=BC(1+C+C^2+C^3+C^4+...) = BC / (1-C).$$

Если *A/B* меньше *C/(1-C),* то можно рассчитать срок окупаемости проекта, но он будет больше, чем *A/B*. Если же *A/B* больше или равно *C/(1-C),* то проект не окупится никогда. Поскольку максимум *C* равен 0,89, то проект не окупится никогда, если *A/B* не меньше 8,09.

Пусть вложения равны 1 млн. руб., ежегодная прибыль составляет 500 тыс., т.е. *A/B*=2, дисконт-фактор *C*=0.8. При примитивном подходе (при *C*=1) срок окупаемости равен 2 годам. А на самом деле? За *k* лет будет возвращено

$$BC(1+C+C^2+C^3+C^4+...+C^k)=BC(1-C^{k+1}) / (1-C).$$

Срок окупаемости *k* получаем из уравнения $1=0{,}5x0{,}8(1-0{,}8^{\ k+1})/(1-0{,}8)$, откуда k= 2,11. Он оказался равным 2,11 лет, т.е. увеличился примерно на 4 недели. Это немного. Однако если *B* = 0,2, то имеем уравнение $1=0{,}2x0{,}8(1-0{,}8^{k+1})/(1-0{,}8)$. У этого уравнения нет корней, поскольку *A/B*=5>*C/(1-C)*=0.8/(1-0,8)=4. Проект не окупится никогда. Прибыль можно ожидать лишь при *A/B*<4. Рассмотрим промежуточный случай, *B*=0,33, с "примитивным" сроком окупаемости 3 года. Тогда имеем уравнение $1=0{,}33x0{,}8\ (1-0{,}8^{\ k+1})/(1-0{,}8)$, откуда *k* = 5,40.

Рассмотрим финансовый поток *a(0), a(1), a(2), a(3), ... , a(t), ...* (для простоты примем, что платежи или поступления происходят раз в год). Выше рассмотрен поток с одним платежом $a(0)=(-A)$ и дальнейшими поступлениями $a(1) = a(2) = a(3) = ... = a(t) = ... = B$. Чистая текущая стоимость (Net Present Value, сокращенно NPV), рассчитывается для финансового потока путем приведения затрат и поступлений к начальному моменту времени:

$$NPV = a(0) + a(1)C(1)+a(2)C(2)+a(3)C(3)+...+ a(t)C(t) + ...,$$

где $C(t)$ - дисконт-функция. В простейшем случае, когда дисконт-фактор не меняется год от года и имеет вид $C=1/(1+q)$, формула для NPV конкретизируется:

$$NPV=NPV(q)=a(0)+a(1)/(1+q) + a(2)/(1+q)^2 +a(3)/(1+q)^3 +...+a(t)/(1+q)^t + ...$$

Пусть, например, $a(0)= -10,\ a(1)=3,\ a(2)=4,\ a(3)=5$. Пусть q=0,12, тогда

$$NPV(0,12)=-10+3x0,89+4x0.80+5x0,71=-10+2,67 + 3,20+3,55= -0,58.$$

Итак, проект невыгоден для вложения капитала, поскольку $NPV(0,12)$ отрицательно. При отсутствии дисконтирования (при $C = 1,\ q = 0$) вывод иной:

$$NPV(0) = - 10 + 3 + 4 + 5 = 2,$$

проект выгоден.

Срок окупаемости и сам вывод о прибыльности проекта зависят от неизвестного дисконт-фактора C или даже от неизвестной дисконт-функции - ибо какие у нас основания считать будущую дисконт-функцию постоянной? Экономическая история России последних лет показывает, что банки часто меняют проценты платы за депозит. Часто предлагают использовать норму дисконта, равную *приемлемой для инвестора норме дохода на капитал.* Это значит, что экономисты явным образом обращаются к инвестору как к эксперту, который должен назвать им некоторое число исходя из своего опыта и интуиции (т.е. экономисты перекладывают свою работу на инвестора). Кроме того, при этом игнорируется изменение указанной нормы во времени,

Приведем пример исследования NPV на устойчивость (чувствительность) к малым отклонениям значений дисконт-функции. Для этого надо найти максимально возможное отклонение NPV при допустимых отклонениях значений дисконт-функции (или, если угодно, значений банковских процентов). В качестве примера рассмотрим

$$NPV = NPV (a(0), a(1), C(1), a(2), C(2), a(3), C(3))=$$
$$= a(0) + a(1)C(1) + a(2)C(2) + a(3)C(3).$$

Предположим, что изучается устойчивость (чувствительность) для ранее рассмотренных значений

$$a(0)=-10,\ a(1)=3,\ a(2)=4,\ a(3)=5,\ C(1)=0,89,\ C(2)=0,80,\ C(3)=0,71.$$

Пусть максимально возможные отклонения $C(1),\ C(2),\ C(3)$ равны $\pm 0,05$. Тогда, максимум значений NPV равен

$$NPV_{max} = -10+3x0,94+4x0.85+5x0,76 = -10+ 2,82 + 3,40 + 3,80 = 0,02,$$

в то время как минимум значений NPV есть

$$NPV_{min} = -10+3x0,84+4x0.75+5x0,66 = -10 +2,52 +3,00+3,30 = -1,18.$$

Для NPV получаем интервал от (-1,18) до (+0,02). В нем есть и положительные, и отрицательные значения. Следовательно, нет однозначного заключения - проект убыточен или выгоден. Для принятия решения не обойтись без экспертов.

Для иных характеристик, например, внутренней нормы доходности, выводы аналогичны. Дополнительные проблемы вносит неопределенность горизонта планирования, а также будущая инфляция (см. главу 7). Если считать, что финансовый поток должен учитывать инфляцию, то это означает, что до принятия решений об инвестициях необходимо на годы вперед спрогнозировать рост цен, а это до сих пор еще не удавалось ни одной государственной или частной исследовательской структуре. Если же рост цен не учитывать, то отдаленные во времени доходы могут "растаять" в огне инфляции. На практике риски учитывают, увеличивая q на десяток-другой процентов.

Следующая глава 10 посвящена более подробному рассмотрению проблем исследования *устойчивости* эконометрических выводов по отношению к возможным отклонениям исходных данных и предпосылок моделей.

Цитированная литература

1. ГОСТ 11.011-83. Прикладная статистика. Правила определения оценок и доверительных границ для параметров гамма-распределения. - М.: Изд-во стандартов, 1984. - 53 с.

2. Шокин Ю.И. Интервальный анализ. Новосибирск: Наука, 1981, 112 с.

3. Орлов А.И. Современная прикладная статистика. - Заводская лаборатория. 1998. Т.64. № 3. - С.52-60.

4. Вощинин А.П. Метод оптимизации объектов по интервальным моделям целевой функции. - М.: МЭИ, 1987. - 109 с.

5. Вощинин А.П., Сотиров Г.Р. Оптимизация в условиях неопределенности. - М.: МЭИ - София: Техника, 1989. - 224 с.

6. Вощинин А.П., Акматбеков Р.А. Оптимизация по регрессионным моделям и планирование эксперимента. - Бишкек: Илим, 1991. - 164 с.

7. Вощинин А.П. Метод анализа данных с интервальными ошибками в задачах проверки гипотез и оценивания параметров неявных линейно параметризованных функций. - Заводская лаборатория. 2000. Т.66. № 3. - С.51-65.

8. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979. - 296 с.

9..Орлов А.И. Интервальный статистический анализ. – В сб.: Статистические методы оценивания и проверки гипотез. Межвузовский сборник научных трудов. – Пермь: Изд-во Пермского государственного университета, 1993. - С.149-158.

Глава 10. Проблемы устойчивости эконометрических процедур

В настоящей главе после обсуждения актуальности проблем изучения устойчивости в различных эконометрических задачах и разбора общей схемы такого изучения рассматриваются три конкретные направления - робастность статистических процедур, устойчивость отношению к объему выборки и устойчивость по отношению к горизонту планирования.

10.1. Общая схема устойчивости

Проблемам познания, в том числе в социально-экономической области, посвящено огромное количество работ. Однако это не значит, что обо всем в этой области уже все сказано. А о некоторых положениях целесообразно говорить еще и еще раз, пока они ни станут общеизвестными.

В настоящей книге предлагаются, изучаются и обсуждаются эконометрические модели социально-экономических явлений и процессов, а также рассматриваются общие требования, которые естественно предъявлять к подобным моделям. В идеале каждую такую модель следовало бы рассматривать как аксиоматическую теорию. В этом идеальном случае создание и использование модели происходит в соответствии с известной триадой "практика - теория - практика". А именно, сначала вводятся некоторые математические объекты, соответствующие интересующим исследователя реальным объектам, и на основе представлений о свойствах реальных объектов формулируются необходимые для успешного моделирования свойства математических объектов, которые и принимаются в качестве аксиом. Затем аксиоматическая теория развивается как часть математики, вне связи с представлениями о реальных объектах. На заключительном этапе полученные в математической теории результаты интерпретируются содержательно. Получаются утверждения о реальных объектах, являющиеся следствиями тех и только тех их свойств, которые ранее были аксиоматизированы.

Рассматриваемые в настоящей книге эконометрические модели также выражены на математическом языке, исследование их ведется средствами математики без привлечения содержательных социально-экономических соображений, а выводы интерпретируются на языке соответствующей предметной области, т.е. содержательно.

После построения математической модели реального явления или процесса встает вопрос о ее адекватности. Иногда ответ на этот вопрос может дать эксперимент. Рассогласование модельных и экспериментальных данных следует интерпретировать как признак неадекватности некоторых из принятых аксиом. Однако для проверки адекватности социально-экономических моделей зачастую невозможно поставить решающий эксперимент в отличие, скажем, от физических моделей. С другой стороны, для одного и того же социально-экономического явления или процесса, как правило, можно составить много возможных моделей, если угодно, много разновидностей одной базовой модели. Поэтому необходимы какие-то дополнительные условия, которые позволяли бы их множества возможных моделей и эконометрических методов анализа данных выбрать наиболее подходящие. В настоящей главе в качестве одного из подобных условий выдвигается требование *устойчивости* модели и метода анализа данных относительно допустимых отклонений исходных данных и предпосылок модели или условий применимости метода.

Отметим, что в большинстве случаев исследователей и практических работников интересуют не столько сами модели и методы, сколько решения, которые с их помощью принимаются. Ведь модели и методы для того и разрабатываются, чтобы подготавливать решения. Вместе с тем очевидно, что решения, как правило, принимаются в условиях неполноты информации. Так, любые числовые параметры известны лишь с некоторой точностью. Введение в рассмотрение возможных неопределенностей исходных данных требует каких-то заключений относительно устойчивости принимаемых решений по отношению к этим допустимым неопределенностям.

Введем основные понятия согласно монографии [1].. Будем считать, что имеются *исходные данные,* на основе которых принимаются *решения.* Способ переработки (отображения) исходных данных в решение назовем *моделью.* Таким образом, с общей точки зрения модель - это функция, переводящая исходные данные в решение, т.е. способ перехода значения не имеет. Очевидно, любая рекомендуемая для практического использования модель должна быть исследована *на устойчивость* относительно допустимых отклонений исходных данных. Укажем некоторые возможные применения результатов подобного исследования:

- заказчик научно-исследовательской работы получает представление о точности предлагаемого решения;

- удается выбрать из многих моделей наиболее адекватную;

- по известной точности определения отдельных параметров модели удается указать необходимую точность нахождения остальных параметров;

- переход к случаю "общего положения" позволяет получать более сильные с математической точки зрения результаты.

Примеры. По каждому из четырех перечисленных возможных применений в настоящей книге уже приведены различные примеры. В эконометрике точность предлагаемого решения связана с разбросом исходных данных и с объемом выборки, и способы оценки точности решения для различных задач расписаны выше. Выбору наиболее адекватной модели посвящены многие рассмотрения в главах 4 и 5, связанные с обсуждением моделей однородности и регрессии. Рациональный объем выборки в статистике интервальных данных (глава 9) исходит из принципа уравнивания погрешностей, основанного на том, что по известной точности определения отдельных параметров модели удается указать необходимую точность нахождения остальных параметров. Другим примером применения той же концепции является нахождение необходимой точности оценивания параметров в моделях логистики, рассмотренных в главе 5 монографии [1]. Наконец, переходом к случаю "общего положения" в эконометрике является, в частности, переход к непараметрической статистике, необходимый из-за невозможности обосновать принадлежность результатов наблюдений к тем или иным параметрическим семействам.

Специалисты по моделированию и теории управления считают устойчивость одной из важных характеристик социально-экономических моделей. Достаточно глубокие исследования ведутся по ряду направлений.

Первоначальное изучение влияния малого изменения одного параметра обычно называют *анализом чувствительности*. Оно обычно описывается значением частной производной. Если модель задается дифференцируемой функцией, то итог анализа чувствительности - вектор значений частных производных в анализируемой точке.

Теория устойчивости решений дифференциальных уравнений развивается по крайней мере с XIX в. Выработаны соответствующие понятия - устойчивость по Ляпунову, корректность, доказаны глубокие теоремы. Для решения некорректных задач академиком АН СССР А.Н. Тихоновым в начале 1960-х годов был предложен метод регуляризации. Модели социально-экономических явлений и процессов, выражаемые с помощью дифференциальных уравнений, могут быть исследованы на устойчивость путем применения хорошо разработанного математического аппарата.

Вопросы устойчивости изучались практически во всех направлениях экономико-математических методов - и в математическом программировании, и в теории массового обслуживания (теории очередей), и в эколого-экономических моделях, и в различных областях эконометрики.

Прежде чем переходить к конкретным постановкам, обсудим "общую схему устойчивости", дающую понятийную базу для обсуждения проблем устойчивости в различных предметных областях.

Определение 1. *Общей схемой устойчивости называется объект*

$$\{\mathrm{A}, \mathrm{B}, d, f, \mathrm{E}\}.$$

Здесь A - множество, называемое (и интерпретируемое) пространством исходных данных; B - множество, называемое пространством решений. Однозначное отображение $f : \mathrm{A} \to \mathrm{B}$ называется моделью. Об этих трех составляющих общей схемы устойчивости уже шла речь выше.

Оставшиеся два понятия нужны для уточнения понятий близости в пространстве исходных данных и пространстве решений. Подобные уточнения могут быть сделаны разными способами. Самое "слабое" уточнение - на языке топологических пространств. Тогда возможны качественные выводы (сходится - не сходится), но не количественные расчеты. Самое "сильное" уточнение - на языке метрических пространств. Промежуточный вариант - используются показатели различия (отличаются от метрик тем, что не обязательно выполняются неравенства треугольника) или вводимые ниже понятия.

Пусть d -показатель устойчивости, т.е. неотрицательная функция, определенная на подмножествах *У* множества B и такая, что из $Y_1 \subseteq Y_2$ вытекает $d(Y_1) \le d(Y_2)$. Часто показатель устойчивости *d(Y)* определяется с помощью метрики, псевдометрики или показателя различия (меры близости) ρ как диаметр множества *У,* т.е.

$$d(Y) = \sup\{\rho(y_1, y_2), y_1 \in Y, y_2 \in Y\}.$$

Таким образом, говоря попросту, в пространстве решений с помощью показателя устойчивости вокруг образа исходных данных может быть сформирована система окрестностей. Но сначала надо такую систему сформировать в пространстве исходных данных.

Пусть $\mathrm{E} = \{E(x,\alpha), x \in \mathrm{A}, \alpha \in \Theta\}$ - совокупность допустимых отклонений, т.е. система подмножеств множества A такая, что каждому элементу множества исходных данных $x \in \mathrm{A}$ и каждому значению параметра α из некоторого множества параметров Θ соответствует подмножество $E(x,\alpha)$ множества исходных данных, называемое множеством допустимых отклонений в точке x при значении параметра, равном α. Наглядно можно представить себе, что вокруг точки x взята окрестность радиуса α.

Определение 2. *Показателем устойчивости в точке x при значении параметра, равном* α*, называется число*

$$\beta(x, E(x,\alpha)) = d(f(E(x,\alpha))).$$

Другими словами, это - диаметр образа множества допустимых колебаний при рассматриваемом в качестве модели отображении. Очевидно, что этот показатель устойчивости зависит как от исходных данных, так и от диаметра множества возможных отклонений в исходном пространстве. Для непрерывных функций показатель устойчивости обычно называется модулем непрерывности.

Естественно посмотреть, насколько сузится образ окрестности возможных отклонений при максимально возможном сужении этой окрестности.

Определение 3. *Абсолютным показателем устойчивости в точке x называется число*

$$\beta(x, \mathrm{E}) = \inf\{\beta(x, E(x,\alpha), \alpha \in \Theta\}.$$

Если функция f непрерывна, а окрестности - именно те, о которых идет речь в математическом анализе, то максимальное сужение означает сужение к точке и абсолютный показатель устойчивости равен 0. Но в главах 3 и 9 мы сталкивались с совсем иными ситуациями. В главе 3 окрестностью исходных данных были все те вектора, что получались из исходного путем преобразования координат с помощью допустимого преобразования шкалы, а допустимое преобразование шкалы бралось из соответствующей группы допустимых преобразований. В главе 9 под окрестностью исходных данных естественно было понимать - при описании выборки - куб с ребрами 2Δ и центром в исходном векторе. И в том, и в другом случае максимальное сужение не означает сужение к точке.

Естественным является желание ввести характеристики устойчивости на всем пространстве. Не вдаваясь в математические тонкости (см. о них монографию [1]), рассмотрим меру μ на пространстве A такую, что мера всего пространства равна 1 (т.е. $\mu(\mathrm{A}) = 1$).

Определение 4. *Абсолютным показателем устойчивости на пространстве исходных данных* A *по мере* μ *называется число*

$$\gamma(\mu) = \int_{\mathrm{A}} \beta(x, \mathrm{E}) d\mu.$$

Здесь имеется в виду так называемый интеграл Лебега. Интегрирование проводится по (абстрактному) пространству исходных данных A по мере μ. Естественно, должны быть выполнены некоторые внутриматематические условия, думать о которых эконометрику ни к чему. Читателю, незнакомому с интегрированием по Лебегу, достаточно мысленно заменить в предыдущей формуле интеграл на сумму (а пространство A считать конечным, хотя и состоящим из большого числа элементов).

Определение 5. *Максимальным абсолютным показателем устойчивости называется*

$$\gamma = \sup\{\beta(x, \mathrm{E}), x \in \mathrm{A}\}.$$

Легко видеть, что $\gamma = \sup \gamma(\mu)$, где супремум берется по всем описанным выше мерам.

Итак, построена иерархия показателей устойчивости эконометрических и экономико-математических моделей. Она с успехом использовалась в исследованиях, подробно развивалась, в частности, в монографии [1]. В частности, полезным оказалось следующее определение.

Определение 6. *Модель f называется абсолютно ε-устойчивой, если $\gamma \le \varepsilon$, где γ - максимальный абсолютный показатель устойчивости.*

Пример. Если показатель устойчивости формируется с помощью метрики ρ, совокупность допустимых отклонений E - это совокупность всех окрестностей всех точек пространства исходных данных A, то 0-устойчивость модели f 'эквивалентна непрерывности модели *f* на множестве A.

Основная проблема в общей схеме устойчивости - *проверка ε-устойчивости данной модели f относительно данной системы допустимых отклонений* E.

Часто оказываются полезными следующие два обобщения основной проблемы.

Проблема А (характеризации устойчивых моделей). *Даны пространство исходных данных* A, *пространство решений* B, *показатель устойчивости d, совокупность допустимых отклонений* E *и неотрицательное число* ε. *Описать достаточно широкий класс ε-устойчивых моделей f. Или: найти все ε-устойчивые модели среди моделей, обладающих данными свойствами, т.е. входящих в данное множество моделей.*

Проблема Б (характеризации систем допустимых отклонений). *Даны пространство исходных данных* A, *пространство решений* B, *показатель устойчивости d, модель f и неотрицательное число* ε. *Описать достаточно широкий класс систем допустимых отклонений* E, *относительно которых модель f является ε-устойчивой. Или: найти все такие системы допустимых отклонений* E *среди совокупностей допустимых отклонений, обладающих данными свойствами, т.е. входящих в данное множество совокупностей допустимых отклонений.*

Ясно, что проблемы А и Б можно рассматривать не только для показателя устойчивости γ, но и для других только что введенных показателей устойчивости, а именно, $\gamma(\mu), \beta(x, \mathrm{E}), \beta(x, E(x,\alpha))$.

Язык общей схемы устойчивости позволяет описывать конкретные задачи специализированных теорий устойчивости в различных областях исследований, выделять в основные элементы в них, ставить проблемы типа А и Б. На этом языке легко формулируются задачи теории устойчивости оеешений жифференциальных уравнений, теории робастности статистических процедур, проблемы адекватности теории измерений (см. главу 3), достигаемой точности расчетов в статистике интервальных данных (см. главу 11) и в логистике (см. монографию [1]), и т.д.

Для примера рассмотрим определение устойчивости по Ляпунову решения $\varphi(t,x)$ нормальной автономной системы дифференциальных уравнений $\dot{y} = g(y)$ с начальными условиями $\varphi(0,x) = x$. Здесь пространство исходных данных А - конечномерное евклидово пространство, множество допустимых отклонений $E(x,\alpha) -$ окрестность радиуса α точки $x \in \mathrm{A}$, пространство решений В - множество функций на луче $[0;+\infty)$ с метрикой

$$\rho(y_1, y_2) = \sup_{t\geq 0} | y_1(t) - y_2(t) | .$$

Модель f - отображение, переводящее начальные условия x в решение системы дифференциальных уравнений с этими начальными условиями $\varphi(t,x)$.

В терминах общей схемы устойчивости положение равновесия a называется *устойчивым по Ляпунову*, если $\beta(a, \mathrm{E}) = 0$. Для формулировки определения асимптотической устойчивости по Ляпунову надо ввести в пространстве решений В псевдометрику

$$\rho_1(y_1, y_2) = \overline{\lim_{t\to\infty}} | y_1(t) - y_2(t) | .$$

Положение равновесия a называется асимптотически устойчивым, если $\beta_1(a, E(a,\varepsilon)) = 0$ для некоторого $\varepsilon > 0$, где показатель устойчивости β_1 рассчитан с использованием псевдометрики ρ_1.

Таким образом, общая схема устойчивости естественным образом включает в себя классические понятия теории устойчивости по Ляпунову. Вместе с тем стоит отметить, что эта схема дает общий подход к различным проблемам устойчивости, прежде всего в эконометрических и экономико-математических постановках, дает систему понятий, которые в каждом конкретном случае должны приспосабливаться к решаемой задаче.

До настоящего момента для определенности речь шла о допустимых отклонениях в пространстве исходных данных. Часто оказывается необходимым говорить и об отклонениях от предпосылок модели. С чисто формальной точки зрения для этого достаточно расширить понятие "исходные данные" до пары *(x, f),* т.е. включив "прежнюю" модель в качестве второго элемента пары. Все остальные определения остаются без изменения. Теперь отклонения в пространстве решений вызываются не только отклонениями в исходных данных *x*, но и отклонениями от предпосылок модели, т.е. отклонениями *f.* Это соображение нам понадобится в следующем пункте настоящей главы, посвященном робастности статистических процедур.

Различные асимптотические постановки в эконометрической теории (третий пункт настоящей главы) также естественно рассматривать как задачи устойчивости. Если при безграничном возрастании объема выборки некоторая величина стремится к пределу, то в терминах общей схемы устойчивости это означает, что она 0-устойчива в соответствующей псевдометрике (см. выше обсуждение асимптотической устойчивости по Ляпунову). С содержательной точки зрения употребление термина "устойчивость" в такой ситуации представляется вполне оправданным, поскольку рассматриваемая величина мало меняется при изменении объема выборки.

Для стратегического менеджмента весьма важна проблема горизонта планирования (подробнее см. учебное пособие [2]). Очевидно, что вид оптимальных решений зависит от заранее заданной длины интервала, для которого строится оптимальных план (т.е. от горизонта планирования). Это означает, что необходимо обосновать выбор горизонта планирования. Принять его бесконечным нерационально, поскольку совершенно ясно, что через каких-нибудь 100 лет производительные силы и производственные отношения будут совсем иные, чем в настоящее время, и пытаться их учитывать для принятия решений в настоящее время нецелесообразно. Как же быть? Об этом - в четвертом пункте настоящей главы.

10.2. Робастность статистических процедур

Термин "робастность" (*robustness* - англ.) образован от *robust* - крепкий, грубый (англ.). Сравните с названием одного из сортов кофе - *robusta*. Имеется в виду, что робастные статистические процедуры должны "выдерживать" ошибки, которые теми или иными способами могут попадать в исходные данные или искажать предпосылки используемых вероятностно-статистических моделей.

Термин "робастный" стал популярным в нашей стране в 1970-е годы. Сначала он использовался фактически как сужение термина "устойчивый" на алгоритмы статистического анализа данных классического типа (не включая теорию измерений, статистику нечисловых и интервальных данных). Затем реальная сфера его применения сузилась.

Пусть исходные данные - это выборка, т.е. совокупность независимых одинаково распределенных случайных величин с одной и той же функцией распределения *F(x).* Наиболее простая модель изучения устойчивости - это модель засорения

$$F(x) = (1-\varepsilon)F_0(x) + \varepsilon H(x). \qquad (1)$$

Эта модель имеются также моделью Тьюки-Хубера. (Джон Тьюки - американский исследователь, П.Хубер, или Хьюбер - швейцарский ученый) Модель (1) показывает, что с близкой к 1 вероятностью, а именно, с вероятностью $(1-\varepsilon)$, наблюдения берутся из совокупности с функцией распределения $F_0(x)$, которая предполагается обладающей "хорошими" свойствами. Например, она имеет известный эконометрику вид (хотя бы с точностью до параметров), у нее существуют все моменты, и т.д. Но с малой вероятностью ε появляются наблюдения из совокупности с "плохим" распределением, например, взятые из распределения Коши, не имеющего математического ожидания, резко выделяющиеся аномальные наблюдения, выбросы.

Актуальность модели (1) не вызывает сомнений. Наличие засорений (выбросов) может сильно исказить результаты эконометрического анализа данных. Ясно, что если функция распределения элементов выборки имеет вид (1), где первое слагаемое соответствует случайной величине с конечным математическим ожиданием, а второе - такой, для которого математического ожидания не существует (например, если *H(x)* - функция распределения Коши), то для итоговой функций распределения (1) также не существует математического ожидания. Исследователя обычно интересуют характеристики первого слагаемого, но найти их, т.е. освободиться от влияния засорения, не так-то просто. Например, среднее арифметическое результатов наблюдений не будет иметь никакого предела (это - строгое математическое утверждение, вытекающее из того, что математическое ожидание не существует [3]).

Существуют различные способы борьбы с засорением. Эмпирическое правило "борьбы с засорениями" при подведении итогов работы команды судей найдено в фигурном катании: наибольшая и наименьшая оценки отбрасываются, а по остальным рассчитывается средняя арифметическая (см. главу 12). Ясно, что "засорение" окажется среди отброшенных оценок.

Оценивать характеристики и параметры, проверять статистические гипотезы, вообще осуществлять эконометрический анализ данных все чаще рекомендуют на основе эмпирических квантилей (другими словами, порядковых статистик, членов вариационного ряда), отделенных от концов вариационного ряда. Речь идет об использовании статистик типа

$$ax(0{,}1n) + bx(0{,}3n) + cx(0{,}5n) + dx(0{,}7n) + ex(0{,}9n).$$

Ценой небольшой потери в эффективности избавляемся от засоренности типа описанной в модели (1).

Вариантом этого подхода является переход к сгруппированным данным. Прямая разбивается на интервалы, и вместо количественных значений эконометрик подсчитывает лишь, сколько наблюдений попало в те или иные интервалы. Особое значение приобретают крайние интервалы - к ним относят все наблюдения, которые больше некоторого верхнего порога и меньше некоторого нижнего порога. Любым методам анализа сгруппированных данных резко выделяющиеся наблюдения не страшны.

Можно поставить под сомнение и саму опасность засорения. Дело в том, что практически все реальные величины ограничены. Все лежат на каком-то интервале - от и до. Это совершенно ясно, если речь идет о физическом измерении - все укладывается в шкалу прибора. По-видимому, и для эконометрических измерений наибольшие сложности создают не сверхбольшие помехи, а не засорения, что находятся "на грани" между "интуитивно возможным" и "интуитивно невозможным".

Что же это означает? Если элементы выборки по абсолютной величине не превосходят числа A, то все засорение может сдвинуть среднее арифметическое на величину εA. Если засорение невелико, то и сдвиг мал.

Построена достаточно обширная и развитая теория, посвященная разработке и изучению методов анализа данных в модели (1). С ней можно познакомиться по монографиям [4-6]. К сожалению, в теории обычно предполагается известной степень засорения ε, а на практике эта величина неизвестна. Кроме того, теория обычно направлена на защиту от воздействий, якобы угрожающих из бесконечности, а на самом деле реальные данные финитны (сосредоточены на конечных отрезках). Все это объясняет, почему теория робастности, исходящая из модели (1), популярна среди теоретиков, но мало интересна тем, кто анализирует реальные экономические данные.

Рассмотрим несколько более сложную модель. Пусть наблюдаются реализации независимых случайных величин $x_1, x_2, ..., x_n$ с функциями распределения $F_1(x), F_2(x), ..., F_n(x)$ соответственно. Эта модель соответствует гипотезе о том, что в процессе наблюдения (измерения) условия несколько менялись. Естественной представляется модель малых отклонений функций распределений наблюдаемых случайных величин от некоторой "базовой" функции распределения $F_0(x)$. Множество возможных значений функций распределений наблюдаемых случайных величин описывается следующим образом:

$$L((F_1, F_2, ..., F_n); \varepsilon) = \{(F_1, F_2, ..., F_n) : \sup_x | F_i(x) - F_0(x) | < \varepsilon, i = 1,2,..., n].$$

Следующий тип моделей - это введение малой (т.е. слабой) зависимости между рассматриваемыми случайными величинами (см., например, монографию [7]). Ограничения на взаимную зависимость можно задать разными способами. Пусть $F(x_1, x_2, ..., x_n)$ - совместная функция распределения, $\rho(i, j)$ – коэффициент корреляции между i-ой и j-ой случайными величинами. Множество возможных совместных функций распределения описывается следующим образом:

$$Z\big(F(x_1, x_2, ..., x_n); \varepsilon\big) = \{F(x_1, x_2, ..., x_n) : | \rho(i, j) | \le \varepsilon, 1 \le i < j \le n\}.$$

Есть еще целый ряд постановок задач робастности. Если накладывать погрешности непосредственно на результаты наблюдений (измерений), то получаем постановки задач статистики интервальных данных (см. главу 11), поскольку каждый результат наблюдения превращается в интервал - исходное значение плюс-минус погрешность.

Разработано много вариантов робастных методов анализа статистических данных (см. монографии [1,4-6]). Иногда говорят, что робастные методы позволяют использовать информацию о том, что реальные наблюдения лежат "около" тех или иных параметрических семейств, например, нормальных. В этом, дескать, их преимущество по сравнению с непараметрическими методами, которые предназначены для анализа данных из всех возможных распределений. Однако количественных подтверждений этих уверений любителей робастных методов обычно не удается найти.

10.3. Устойчивость по отношению к объему выборки

В настоящем пункте рассматривается проблема и методы оценки близости предельных распределений статистик и распределений, соответствующих конечным объемам выборок. При каких объемах выборок уже можно пользоваться предельными распределениями? Каков точный смысл термина "можно" в предыдущей фразе? Основное внимание уделяется переходу от точных формул допредельных распределений к пределу и применению метода статистических испытаний (Монте-Карло).

Асимптотическая математическая статистика и практика анализа статистических данных. Как обычно подходят к обработке реальных данных в конкретной эконометрической задаче? Первым делом строят статистическую модель. Если хотят перенести выводы с совокупности результатов наблюдений на более широкую совокупность, например, предсказать что-либо (см. главу 14), то рассматривают, как правило, вероятностно-статистическую модель. Например, традиционную модель выборки, в которой результаты наблюдений - реализации независимых (в совокупности) одинаково распределенных случайных величин. Очевидно, ***любая модель лишь приближенно соответствует реальности.*** В частности, естественно ожидать, что распределения результатов наблюдений несколько отличаются друг от друга, а сами результаты связаны между собой, хотя и слабо (см. предыдущий пункт).

Итак, первый этап - переход от реальной ситуации к математической модели. Далее - неожиданность: на настоящем этапе своего развития математическая теория эконометрики и статистики зачастую не позволяет провести необходимые исследования для имеющихся объемов выборок. Более того, отдельные математики пытаются оправдать свой отрыв от практики соображениями о структуре этой теории, на первый взгляд убедительными. Неосторожная давняя фраза Б.В. Гнеденко и А.Н.Колмогорова: "Познавательная ценность теории вероятностей раскрывается только предельными теоремами" (см. классическую монографию [8], одну из наиболее ценных математических книг XX в.) взята на вооружение и более близкими к нам по времени авторами. Так, И.А. Ибрагимов и Р.З. Хасьминский пишут: "Решение неасимптотических задач оценивания, хотя и весьма важное само по себе, как правило, не может являться объектом достаточно общей математической теории. Более того, соответствующее решение часто зависит от конкретного типа распределения, объема выборки и т.д. Так, теория малых выборок из нормального закона будет отличаться от теории малых выборок из закона Пуассона" (см. напичканную формулами монографию [9, с.7]).

Согласно цитированным и подобным им авторам, основное содержание математической теории статистики - предельные теоремы, полученные в предположении, что объемы рассматриваемых выборок стремятся к бесконечности. Эти теоремы опираются на предельные соотношения теории вероятностей, типа Закона Больших Чисел и Центральной Предельной Теоремы. Ясно, что сами по себе подобные утверждения относятся к математике, т.е. к сфере чистой абстракции, и не могут быть непосредственно применены для анализа реальных данных. Их практическое использование, о котором "чистые" математики предпочитают не думать, опирается на важное предположение: *"При данном объеме выборки достаточно точными являются асимптотические формулы."*

Конечно, в качестве первого приближения представляется естественным воспользоваться асимптотическими формулами, не тратя сил на анализ их точности. Но это - лишь начало долгой цепи исследований. Как же обычно преодолевают разрыв между результатами асимптотической математической статистики и потребностями практики эконометрического и статистического анализа данных? Какие "подводные камни" подстерегают на этом пути? Обсуждению этих вопросов и посвящен настоящий пункт.

Точные формулы и асимптотика. Начнем с наиболее продвинутой в математическом плане ситуации, когда для статистики известны как предельное распределение, так и распределения при конечных объемах выборки.

Примером является двухвыборочная односторонняя статистика Н.В.Смирнова. Рассмотрим две независимые выборки объемов m и n из непрерывных функций распределения $F(x)$ и $G(x)$ соответственно. Для проверки гипотезы однородности двух выборок (ср. главу 4)

$$H_0 : F(x) = G(x) \text{ для всех действительных чисел } x$$

в 1939 г. Н.В. Смирнов в статье [10] предложил использовать статистику

$$D^+(m,n) = sup\ (F_m(x) - G_n(x))\ ,$$

где $F_m(x)$ - эмпирическая функция распределения, построенная по первой выборке, $G_n(x)$ - эмпирическая функция распределения, построенная по второй выборке, супремум берется по всем действительным числам x. Для обсуждения проблемы соотношения точных и предельных результатов ограничимся случаем равных объемов выборок, т.е. $m = n$. Положим

$$H(n,t) = P(D^+(n,n) \geq \frac{t}{\sqrt{n}}).$$

В цитированной статье [10] Н.В. Смирнов показал, что при безграничном возрастании объема выборки n вероятность $H(n, t)$ стремится к $exp(- t^2)$.

В работе [11] 1951 г. Б.В. Гнеденко и В.С. Королюк показали, что при целом $c = t\sqrt{n}$ (именно при таких t вероятность $H(n, t)$ как функция t имеет скачки, поскольку статистика Смирнова $D^+(n,n)$ кратна $1/n$) рассматриваемая вероятность $H(n, t)$ выражается через биномиальные коэффициенты, а именно,

$$H(n,t) = \binom{2n}{n-c} \Big/ \binom{2n}{n}. \qquad (1)$$

К сожалению, непосредственные расчеты по формуле (1) возможны лишь при сравнительно небольших объемах выборок, поскольку величина n!. (n-факториал) уже при n=100 имеет более 200 цифр и не может быть без преобразований использована в вычислениях. Следовательно, наличие точной формулы для интересующей нас вероятности не снимает необходимости использования предельного распределения и изучения точности приближения с его помощью.

Широко известная формула Стирлинга для гамма-функции и, в частности, для факториалов позволяет преобразовать последнее выражение в асимптотическое разложение, т.е. построить бесконечный степенной ряд (по степеням n) такой что каждая следующая частичная сумма дает все более точное приближение для интересующей нас вероятности $H(x, t)$. Это и было сделано в работе А.А. Боровкова 1962 г. Большое количество подобных разложений для различных статистических задач приведено в работах В.М. Калинина и О.В. Шалаевского конца 1960-х - начала 1970-х годов. (Интересно отметить, что асимптотические разложения в ряде случаев расходятся, т.е. остаточные члены имеют нетривиальную природу.)

Затем в работах конца семидесятых годов была сделана попытка теоретически оценить остаточный член второго порядка. Итоги подведены в монографии [1, § 2.2, с.37-45]. Справедливо равенство

$$H(n, t) = \exp(-t^2).(1 + f(t)/n + g(n,t)/n^2),$$

где

$$f(t) = t^2 (1/2 - t^{2/} / 6).$$

Целью последних из названных работ было получение равномерных по n, t оценок остаточного члена второго порядка $g(n,t)$ сверху и снизу в области, задаваемой условиями

$$0 < \frac{t}{\sqrt{n}} < A, \quad 0 < t < t_{\max}, \quad , n \geq n_0. \qquad (2)$$

где A, $t_{\max}$, n_0 - некоторые параметры. С помощью длинных цепочек оценок остаточных членов в формулах, получаемых при преобразовании формулы (1) к предельному виду, сформулированная выше цель была достигнута, и для различных наборов параметров A, $t_{\max}$, n_0 получены равномерные по n, t оценки (сверху и снизу) остаточного члена второго порядка $g(n,t)$ в области (2). Так, например, при A = 0,5, t_{max} = 1,73, n_0 = 8 нижняя граница равна (- 0,71), а верхняя есть 2,65.

Основными недостатками такого подхода являются, во первых, зависимость оценок от параметров A, $t_{\max}$, n_0, задающих границы областей, во-вторых, завышение оценок, иногда в сотни раз, обусловленное желанием получить равномерные оценки по области (оценкой реальной погрешности в конкретной точке является значение следующего члена асимптотического разложения).

Поэтому при составлении рассчитанной на практическое использование методики [12] проверки однородности двух выборок с помощью статистики Смирнова было решено перейти на несколько другую методологию (назовем ее "методологией заданной точности"), которую кратко можно описать следующим образом.

1) выбирается достаточно малое положительное число p, например p = 0,05 или p = 0,20;

2) приводятся точные значения $H(n, t)$ для всех значений n таких, что

$$| H(n, t) - \exp(-t^2) | > p \exp(-t^2);$$

3) если же последнее неравенство не выполнено, то предлагается пользоваться вместо $H(n, t)$ предельным значением $\exp(-t^2)$.

Таким образом, принятая в методике [12] методология предполагает интенсивное использование вычислительной техники. Результатами расчетов являются *граничные значения* объемов выборок $n(p,t)$ такие, что при меньших значениях объемов выборок рекомендуется пользоваться точными значениями функции распределения статистики Смирнова, а при больших - предельными. Описывается этот результат таблицей, а не формулой. Отметим, что при построении реальных таблиц не обойтись без выбора того или иного конкретного значения p, задающего объемы таблиц.

Оценки скорости сходимости. Теоретические оценки скорости сходимости в различных задачах эконометрики и прикладной математической статистики иногда формулируются в весьма абстрактном виде. Так, в 1960-1970-х годах была популярна задача оценки скорости сходимости распределения классической статистики омега-квадрат (Крамера-Мизеса-Смирнова). Для максимума модуля разности допредельной и предельной функций распределения этой статистики различные авторы доказывали, что для любого e>0 существует константа $C(e)$ такая, что упомянутый максимум не превосходит $C(e)\ n^{-w+e}$. Прогресс состоял в увеличении константы w. Сформулированный выше результат был доказан последовательно для w = 1/10, 1/6, 1/5, 1/4, 1/3, 1/2 и 1 (подробнее история этих исследований рассказана в §2.3 монографии [1]).

Конечно, все эти исследования не могли дать конкретных практических рекомендаций. Однако необходимой исходной точкой является само существование предельного распределения. Представим себе, что некто, не зная, что у распределения Коши нет математического ожидания, моделирует выборочные средние арифметические результатов наблюдений из этого распределения. Ясно, что его попытки оценить скорость сходимости выборочных средних к пределу обречены на провал.

Последовательное улучшение теоретических оценок скорости сходимости дает надежду на быструю реальную сходимость. Действительно, численные расчеты показали, что предельным распределением для статистики омега-квадрат (Крамера-Мизеса-Смирнова) можно пользоваться уже при объеме выборки, равном 4.

Использование датчиков псевдослучайных чисел. Если же предельное распределение известно, то возникает возможность изучить скорость сходимости численно методом статистических испытаний (Монте-Карло). Однако при этом обычно возникают две проблемы.

Во-первых, откуда известно, что скорость сходимости монотонна? Если при данном объеме выборки различие мало, то будет ли оно мало и при дальнейших объемах? Иногда отклонения допредельного распределения от предельного объясняются довольно сложными причинами. Так, для распределения хи-квадрат они связаны с рядом до сих пор не решенных теоретико-числовых проблем о числе целых точек в эллипсоиде растущего диаметра.

Во-вторых, с помощью датчиков псевдослучайных чисел получаем допредельные распределения с погрешностью, которая может преуменьшать различие. Поясним мысль аналогией. Растущий сигнал измеряется с погрешностями. Когда можно гарантировать, что его величина наверняка превзошла заданную границу?

Напомним, что проблема качества датчиков псевдослучайных чисел продолжает оставаться открытой (см. главу 11). Для моделирования в пространствах фиксированной размерности датчики псевдослучайных чисел решают поставленные задачи. Но для рассматриваемых нами задач размерность не фиксирована - мы не знаем, при каком конкретно объеме выборки можно переходить к предельному распределению согласно "методологии заданной точности".

Нужны дальнейшие работы по изучению качества датчиков псевдослучайных чисел в задачах неопределенной размерности. Поскольку критиков датчиков обычно обвиняют в том, что они сами их не используют, отмечу, что мы применяли этот инструментарий при изучении помех, создаваемых электровозами (см. монографию [1]), при изучении статистических критериев проверки однородности двух выборок (см. работу [13]).

А нужна ли вообще асимптотика? В настоящее время развивается актуальное направление прикладной статистики, связанное с интенсивным использованием вычислительной техники для изучения свойств статистических процедур. Как уже отмечалось, математические методы в статистике обычно позволяют получать лишь асимптотические результаты, и для переноса выводов на конечные объемы выборок приходится применять вычислительные методы. В Новосибирском государственном техническом университете разработан и успешно применяется оригинальный подход, основанный на интенсивном использовании современной вычислительной техники. Основная идея такова: в качестве альтернативы асимптотическим методам математической статистики используется анализ результатов статистического моделирования (порядка 2000 испытаний) выборок конкретных объемов (200, 500, 1000). При этом анализ предельных распределений заменяется на анализ распределений соответствующих статистик при указанных объемах выборок.

К достоинствам подхода относится возможность замены теоретических исследований расчетами. Разработанная программная система дает в принципе возможность численно изучить свойства любого статистического алгоритма для любого конкретного распределения результатов наблюдений и любого конкретного объема выборки. К недостаткам рассматриваемого подхода относится зависимость от свойств датчиков псевдослучайных чисел, а также - что более важно - неизвестность предельного распределения (и даже самого факта его существования), а потому невозможность обоснованного переноса полученных выводов на объемы выборок, отличные от исследованных. Поэтому с точки зрения теории математической статистики полученные рассматриваемым способом результаты следует рассматривать как правдоподобные (а не доказательные, как в классической математической статистике).

Кроме того, они принципиально неточные. Даже в наиболее благоприятных условиях отклонение смоделированного распределения, построенного по 2000 испытаниям, от теоретического предельного распределения, по нашей оценке, может иметь порядок $(1/2000 + 1/1000)^{1/2} = 0{,}038$ (ср. главу 4). Это означает, в частности, что процентные точки, соответствующие уровням значимости 0,05 и особенно 0,01, могут сильно отличаться от соответствующих процентных точек предельных распределений. Очевидно, следующий этап работ - изучение точности полученных в рассматриваемом подходе выводов, прежде всего приближений и процентных точек.

Однако сразу все не сделаешь. Поэтому новосибирцы совершенно правы, развивая новые компьютерные подходы к давним задачам эконометрики и прикладной математической статистики. В частности, весьма полезными и интересными являются результаты, касающиеся непараметрических критериев согласия. Весьма интересным и полезным представляется также метод построения оптимального группирования, в частности, при использовании критериев типа хи-квадрат. Важен результат о неробастности (неустойчивости) оценок максимального правдоподобия по негруппированным данным. Надо поддержать идею использования одновременно двух оценок по группированным данным с использованием как оптимального, так и равновероятного группирования. Этот подход сибиряков соответствует современным идеям в области устойчивости (робастности) статистических выводов.

Однако стоит сделать два замечания. В работе [14] сравниваются два плана контроля надежности технических изделий. Оказывается, что при объемах выборки, меньших 150, лучше первый план, а при объемах, больших 150 - второй. Значит, если бы по новосибирскому методу сравнивались эти планы при достаточно большом объеме выборки n=100, то лучшим был бы признан первый план, что неверно - наступит момент (объем выборки), когда лучшим станет второй план.

Другая относящаяся к делу ассоциация - из весьма содержательной монографии о прикладной математике [15]. Будем суммировать бесконечный ряд с членами $z_n = 1/n$. Поскольку члены его убывают, то обычно используемые алгоритмы остановят вычисления на каком-то шагу. А сумма-то - бесконечна!

Кажется, что компьютер дал универсальную отмычку ко всем проблемам вообще и в области эконометрики в частности. Но это только кажется.

Итак, в Новосибирском государственном техническом университете предложен интересный эконометрический инструментарий и проделана полезная работа. Однако этот подход никоим образом не является панацеей.

10.4. Устойчивость по отношению к горизонту планирования

Продолжим начатое выше обсуждение влияние горизонта планирования на принимаемые решения. Отметим, что во многих реальных ситуациях продолжительность инвестиционного проекта не полностью определена либо горизонт планирования инвестора не охватывает всю продолжительность реализации проекта до этапа утилизации. В таких случаях важно изучить влияние горизонта планирования на принимаемые решения.

Рассмотрим условный пример. Предположим, я являюсь владельцем завода. Если горизонт моего планирования - 1 месяц, то наибольший денежный доход я получу, продав предприятие (включая здания, сырье, технологическое оборудование, землю, на котором

стоит предприятие - если, конечно, я имею право ее продать). Если же планирую на год, то я сначала понесу затраты, закупив сырье и оплатив труд рабочих, и только затем, продав продукцию, получу прибыль. Если я планирую на 10 лет, то пойду на крупные затраты, закупив лицензии и новое оборудование, с целью увеличения дохода в дальнейшие годы. При планировании на 30 лет имеет смысл вложить средства в создание и развитие собственного научно-исследовательского центра, и т.д.
Подчеркнем - реальные инвестиции (в основные фонды - в здания, оборудование, в конструкторские разработки и т.д.), которые окупятся в следующие годы, в текущем году ухудшат многие финансово-хозяйственные показатели работы предприятия, сократят его прибыль, уменьшат показатели рентабельности, в итоги акционеры получат - в данном году - меньше.

Таким образом, популярное утверждение "фирма работает ради максимизации прибыли" или "цель фирмы - максимизация прибыли" не имеет точного смысла. За какой период максимизировать прибыль - за месяц, год, 10 или 30 лет? От горизонта планирования зависят принимаемые решения. Понимая это, ряд западных экономистов отказываются рассматривать фирмы как инструменты для извлечения прибыли, предпочитают смотреть на них как на квазиживые существа, старающиеся обеспечить продолжение своего существования и дальнейшее развитие. Соответственно с этим стратегический менеджмент исходит из понятий "миссия фирмы", "стратегические цели" (например, стратегическая цель может иметь вид: "повысить долю рынка, контрлируемую фирмой"), которые невозможно непосредственно выразить в денежных единицах (подробнее об этом см., например, учебное пособие [2]).

Прежде чем обсуждать непосредственно влияние горизонта планирования на принимаемые менеджером решения, рассмотрим классификацию используемых при принятии решепний оптимизационных моделей.

Характеризация моделей с дисконтированием. Пусть для простоты изложения время принимает дискретные значения. Тогда развитие экономической ситуации описывается последовательностью $x_1, x_2, ..., x_m$, где переменные x_j лежат в некотором пространстве X, возможно, достаточно сложной природы. Надо отметитть также, что положение в следующий момент не может быть проризвольным, оно связано с положением в предыдущий момент. Проще всего принять, что существует некоторое множество K такое, что $(x_j, x_{j+1}) \in K, \quad j = 1,2,..., m-1$. Результат экономической деятельности за j-ый период описывается величиной $f_j(x_j, x_{j+1})$. Зависимость не только от начального и конечного положения, но и от номера периода объясняется тем, что через номер периода осуществляется связь с общей экономической ситуацией. Желая максимизировать суммарные результаты экономической деятельности, приходим к постановке стандартной задаче динамического программирования:

$$F_m(x_1, x_2, ..., x_m) = \sum_{1 \le j \le m-1} f_j(x_j, x_{j+1}) \to \max, \quad (1)$$

$$(x_j, x_{j+1}) \in K, \quad j = 1,2,..., m-1.$$

Таким образом, необходимо выбрать план $(x_1, x_2, ..., x_m)$, удовлетворяющий приведенным ограничениям, на котором достигает максимума функционал F_m . Естественно, предполагается, что множество возможных переходов K таково, что область определения функционала F_m не пуста. ПРи обычных математических предположениях максимум достигается.

Как известно, задача (1) часто возникает во многих прикладных экономических и эконометрических областях, в макроэкономике, в логистике (управлении запасами) (см., например, монографию [1]).

Широко предлагаются, исследуются и применяются модели, приводящи5 к следующему частному случаю задачи (1):

$$F_m(x_1, x_2, ..., x_m) = \sum_{1 \le j \le m-1} \alpha^{o-1} f(x_j, x_{j+1}) \to \max, \quad (2)$$
$$(x_j, x_{j+1}) \in K, \quad j = 1,2,..., m-1.$$

Это - модели с дисконтированием (как известно, α - дисконт-фактор). Естественно попытаться выяснить, какими "внутренними" свойствами выделяются задачи типа (2) из всех задач типа (1). В частности, почему такой большой популярностью пользуется характеристика инвестиционного проекта *NPV (Net Present Value* - чистая текущая стоимость), относящаяся к характеристикам дисконтированного типа.

Представляет интерес изучение и сравнение между собой планов возможного экономического поведения на k шагов $X_1 = (x_{11?} x_{21?} ..., x_{k1})$ и $X_2 = (x_{12?} x_{22}, ..., x_{k2})$. (Естественно, предполагаем, что все пары соседних элементов входят в множество *K*.) Естественно сравнение проводить с помощью описывающих результаты экономической деятельности функций, участвующих в задачах (1) и (2). Именно, будем говорить, что план X_1 лучше плана X_2 при реализации с момента *i*, если

$$f_i(x_{11}, x_{21}) + f_{i+1}(x_{21}, x_{31}) + ... + f_{i+k-1}(x_{(k-1)1}, x_{k1}) >$$
$$> f_i(x_{12}, x_{22}) + f_{i+1}(x_{22}, x_{32}) + ... + f_{i+k-1}(x_{(k-1)2}, x_{k2}). \quad (3)$$

Будем писать $X_1 R(i) X_2$, если выполнено неравенство (3), где *R(i)* - бинарное отношение на множестве планов, задающее упорядочение планов отношением "лучше".

Ясно, что упорядоченность планов на *k* шагов, определяемая с помощью бинарного отношения *R(i),* может зависеть от *i*, т.е. "хорошесть" плана зависит от того, с какого момента *i* он начинает осуществляться. С точки зрения реальной экономики это вполне понятно. Например, планы действий, вполне рациональные для периода стабильного развития, никуда не годятся в период гиперинфляции. И наоборот, приемлемые в период гиперинфляции операции не принесут эффекта в стабильной обстановке.

Однако, как легко видеть, в моделях с дисконтированием (2) все упорядочения *R(i)* совпадают*, i = 1,2, ...,m-k.* Оказывается - это и есть основной результат настоящего подпункта - верно и обратное: если упорядочения совпадают, то мы имеем дело с задачей (2) - с задачей с дисконтированием, причем достаточно совпадения только при *k*=1,2. Сформулируем более подробно предположения об устойчивости упорядочения планов.

(I). Пусть $(x, y) \in K, \quad (x', y') \in K$. Верно одно из двух: либо

$$f_i(x, y) > f_i(x', y')$$

для всех $i = 1,2,..., m-1;$ либо

$$f_i(x, y) \le f_i(x', y')$$

для всех $i = 1,2,..., m-1$.

(II). Пусть $(x, y) \in K, \quad (y, z) \in K, \quad (x', y') \in K, \quad (y', z') \in K$. Верно одно из двух: либо

$$f_i(x, y) + f_{i+1}(y, z) > f_i(x', y') + f_{i+1}(y', z')$$

для всех $i = 1,2,..., m-2;$ либо

$$f_i(x, y) + f_{i+1}(y, z) \le f_i(x', y') + f_{i+1}(y', z')$$

для всех $i = 1,2,..., m-2$.

Как впервые подробно показано в работе [16], при некоторых внутриматематических условиях регулярности из условий устойчивости упорядоченности планов (I) и (II) следует существование констант $\alpha > 0$ и $d_j, \quad j = 2,..., m-1$, таких, что

$$f_j(x, y) = \alpha^{j-1} f_1(x, y) + d_j, \quad j = 2,..., m-1.$$

Поскольку прибавление константы не меняет точки, в которой функция достигает максимума, то последнее соотношение означает, что условия устойчивости упорядоченности планов (I) и (II) характеризуют (другими словами, однозначно выделяют) модели с дисконтированием среди всех моделей динамического программирования.

Как и в случае проблем устойчивости результата сравнения средних в теории измерений (см. главу 3), проблему характеризации моделей с дисконтированием нетрудно выразить на языке общей схемы устойчивости. Речь идет о проблеме Б. Подробности оставляем читателю (см. также монографию [1]).

Математические условия, при которых доказывалась теорема о характеризации моделей с дисконтированием, постепенно ослаблялись на протяжении 1970-х годов (см. об этом в [1]), однако на экономическую сторону дела эти внутриматематические усовершенствования не влияли.

Асимптотически оптимальные планы. Рассмотрим модель (2) с $\alpha = 1$:

$$F_m(x_1, x_2, ..., x_m) = \sum_{1 \le j \le m-1} f(x_j, x_{j+1}) \to \max, \quad (4)$$

$$(x_j, x_{j+1}) \in K, \quad j = 1,2,..., m-1.$$

При естественных математических предположениях, на которых не будем останавливаться, при каждом *m* существует оптимальный план $(x_1(m), x_2(m),..., x_m(m))$, при котором достигает максимума оптимизируемая функция. Поскольку выбор горизонта планирования нельзя рационально обосновать, хотелось бы построить план действий, близкий к оптимальному при различных горизонтах планирования. Это значит, что целью является построение бесконечной последовательности $(y_1, y_2, ...)$ такой, что ее начальный отрезок длины *m,* т.е. $(y_1, y_2, ... y_m)$, дает примерно такое же значение оптимизируемого функционала, как и значение для оптимального плана $(x_1(m), x_2(m),..., x_m(m))$. Бесконечную последовательность $(y_1, y_2, ...)$ назовем асимптотически оптимальным планом.

Выясним, можно ли использовать для построения асимптотически оптимального плана непосредственно оптимальный план. Зафиксируем *k* и рассмотрим последовательность $x_k(m)$. Нетрудно построить примеры, показывающие, что, во-первых, элементы в этой последовательности будут меняться; во-вторых, они могут не иметь пределов. Следовательно, оптимальные планы могут вести себя крайне нерегулярно, а потому в таких случаях их нельзя использовать для построения асимптотически оптимальных планов.

Тем не менее можно доказать (соответствующая экономико-математическая теория развита в главе 5 монографии [1]), что асимптотически оптимальные планы существуют, т.е. можно указать такие бесконечные последовательности $(y_1, y_2, ...)$, что

$$\lim_{m \to \infty} \frac{F_m(x_1(m), x_2(m),..., x_m(m))}{F_m(y_1, y_2, ... y_m)} = 1.$$

Таким образом решается проблема горизонта планирования - надо использовать асимптотически оптимальные планы, не зависящие от горизонта планирования. Интересно, что оптимальная траектория движения состоит из трех участков - начального, конечного и основного, а основной - это движение по магистрали. Полная аналогия с движением автотранспорта: чтобы попасть куда-либо, нужно сначала выехать на магистраль (шоссе), подъехать по хорошей дороге возможно ближе к цели, потом преодолеть заключительный участок.

Обсуждение проблем устойчивости будет продолжено в дальнейших главах книги.

Цитированная литература

1. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979. -296 с.
2. Менеджмент. Учебное пособие / Под ред. Ж.В. Прокофьевой. - М.: Знание, 2000. -288 с.
3. Гнеденко Б.В. Курс теории вероятностей: Учебник. - Изд. 6-е, перераб. и доп. - М.: Наука, Гл. ред. физ.-мат. лит., 1988. - 448 с.
4. Смоляк С.А., Титаренко Б.П. Устойчивые методы оценивания: Статистическая обработка неоднородных совокупностей. - М;: Статистика, 1980. - 208 с.
5. Хьюбер П. Робастность в статистике. - М.: Мир, 1984. - 304 с.
6. Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В. Робастность в статистике. Подход на основе функций влияния. - М.: Мир, 1989. - 512 с.
7. Эльясберг П.Е. Измерительная информация. Сколько ее нужно, как ее обрабатывать? - М.: Наука, 1983. - 208 с.
8. Гнеденко Б.В., Колмогоров А.Н. Предельные распределения для сумм независимых случайных величин. - М.-Л.: ГИТТЛ, 1949. 264 с.
9. Ибрагимов И.А., Хасьминский Р.З. Асимптотическая теория оценивания. -М.: Наука, 1979. 528 с.
10. Смирнов Н.В. Оценка расхождения между эмпирическими кривыми распределения в двух независимых выборках. // Бюллетень. МГУ им. М.В. Ломоносова. Сер. А. 1939. Т.2. № 2. С.3-14.
11. Гнеденко Б.В., Королюк В.С.О максимальном расхождении двух эмпирических распределений. // Доклады АН СССР. 1951. Т.80. № 4. С.525-528.
12. Методика. Проверка однородности двух выборок параметров продукции при оценке ее технического уровня и качества. - М.: Всесоюзный научно-исследовательский институт стандартизации Госстандарта СССР, 1987. - 116 с.
13. Камень Ю.Э., Камень Я.Э., Орлов А.И. . Реальные и номинальные уровни значимости в задачах проверки статистических гипотез // Заводская лаборатория. 1986. Т. 52. №. 12. С. 55-57.
14. Левин Б.Р., Демидович Н.О. Использование непараметрических методов при обработке результатов испытаний на надежность. // Надежность средств связи. - Киев: Техніка, 1976. - С.59-72.
15. Блехман И.И., Мышкис А.Д., Пановко Я.Г. Механика и прикладная математика: Логика и особенности приложений математики. - М.: Наука, 1983. - 328 с.
16. Orlov A. Sur la stabilite' dans les modeles economiques discrets et les modeles de gestion des stocks. // Publications Econometriques. 1977. Vol.X. F. 2. Pp.63-81.

Глава 11. Эконометрические информационные технологии

В этой главе рассматриваем эконометрические технологии, связанные с использованием нескольких эконометрических методов и с применением компьютеров. И то, и другое дает качественный скачок - переход от ручного анализа данных к современной информационной технологии обработки экономической информации. Рассматриваются проблемы последовательной проверки статистических гипотез, разработки и обоснования статистических технологий, методов статистических испытаний (Монте-Карло) и датчиков псевдослучайных чисел, методов размножения выборок (бутстреп-методов), эконометрической поддержки принятия решений в контроллинге.

11.1. Проблема множественных проверок статистических гипотез

Практика применения эконометрических методов часто выходит за границы классической математико-статистической теории. В качестве примера рассмотрим проверку статистических гипотез.

Базовая теоретическая модель касается проверки одной-единственной статистической гипотезы. На практике же при выполнении того или иного прикладного исследования гипотезы зачастую проверяют неоднократно. При этом, как правило, остается неясным, как влияют результаты предыдущих проверок на характеристики (уровень значимости, мощность) последующих проверок. Есть ли вообще влияние? Как его оценить? Как его учесть при формулировке окончательных выводов?

Изучены лишь некоторые схемы множественных проверок, например, схема последовательного анализа А. Вальда или схема оценивания степени полинома в регрессии путем последовательной проверки адекватности модели (см. главу 5 выше). В таких исключительных постановках удается рассчитать характеристики статистических процедур, включающих множественные проверки статистических гипотез.

Однако в большинстве важных для практики случаев статистические свойства процедур анализа данных, основанных на множественных проверках, остаются пока неизвестными. Примерами являются процедуры нахождения информативных подмножеств признаков (коэффициенты для таких и только таких признаков отличны от 0) в регрессионном анализе или выявления отклонений параметров в автоматизированных системах управления.

В таких системах происходит слежение за большим числом параметров. Резкое изменение значения параметра свидетельствует об изменении режима работы системы, что, как правило, требует управляющего воздействия. Существует теория для определения границ допустимых колебаний одного или фиксированного числа параметров. Например, можно использовать контрольные карты Шухарта или кумулятивных сумм, а также их многомерные аналоги (см. главу 13). В подавляющем большинстве постановок, согласно обычно используемым вероятностным моделям, для каждого параметра, находящемся в стабильном ("налаженном") состоянии, существует хотя и малая, но положительная вероятность того, что его значение выйдет за заданные границы. Тогда система зафиксирует резкое изменение значения параметра ("ложная разладка"). При достаточно большом числе параметров с вероятностью, близкой к 1, будет обнаружено несколько "случайных сбоев", среди которых могут "затеряться" и реальные отказы подсистем. Можно доказать, что при большом числе параметров имеется два крайних случая - независимых (в совокупности) параметров и функционально связанных параметров, а для всех остальных систем вероятность обнаружения резкого отклонения хотя бы у одного параметра лежит между соответствующими вероятностями для этих двух крайних случаев.

Почему трудно изучать статистические процедуры, использующие множественные проверки гипотез? Причина состоит в том, что результаты последовательно проводящихся проверок, как правило, не являются независимыми (в смысле независимости случайных величин). Более того, последовательность проверок зачастую задается исследователем произвольно.

Проблема множественных проверок статистических гипотез - часть более общей проблемы "стыковки" (сопряжения) статистических процедур. Дело в том, что каждая процедура может применяться лишь при некоторых условиях, а в результате применения предыдущих процедур эти условия могут нарушаться. Например, часто рекомендуют перед восстановлением зависимости (регрессионным анализом) разбить данные на однородные группы с помощью какого-либо алгоритма классификации, а затем строить зависимости для каждой из выделенных групп отдельно. Здесь идет речь о "стыковке" алгоритмов классификации и регрессии. Как вытекает из рассмотрений главы 5 выше, попадающие в одну однородную группу результаты наблюдений зависимы и их распределение не является нормальным (гауссовым), поскольку они лежат в ограниченной по некоторым направлениям области, а границы зависят от всей совокупности результатов наблюдений. При этом при росте объема выборки зависимость уменьшается, но ненормальность остается Распределение результатов наблюдений, попавших в одну группу, приближается не к нормальному, а к усеченному нормальному. Следовательно, алгоритмами регрессионного анализа, основанными на "нормальной теории", пользоваться некорректно. Согласно рекомендациям главы 10 целесообразно применять робастную регрессию.

Проблема "стыковки" статистических процедур обсуждается давно. По проблеме "стыковки" был проведен ряд исследований, результаты некоторые из которых упомянуты выше, но сколько-нибудь окончательных результатов получено не было. По нашему мнению, на скорое решение проблемы "стыковки" рассчитывать нельзя. Возможно, она является столь же "вечной", как и проблема выбора между средним арифметическим и медианой как характеристиками "центра" выборки.

В качестве примера обсудим одно интересное исследование по проблеме повторных проверок статистических гипотез - работу С.Г.Корнилова [1].

Как уже отмечалось, теоретическое исследование является весьма сложным, сколько-нибудь интересные результаты удается получить лишь для отдельных постановок. Поэтому вполне естественно, что С.Г. Корнилов применил метод статистического моделирования на ЭВМ. Однако нельзя забывать о проблеме качества псевдослучайных чисел. Достоинства и недостатки различных алгоритмов получения псевдослучайных чисел много лет обсуждаются в различных изданиях (см. ниже).

В работе С.Г.Корнилова хорошо моделируется *мышление* статистика-прикладника. Видно, насколько мешает устаревшее представление о том, что для проверки гипотез необходимо задавать определенный уровень значимости. Особенно оно мешает, если в дальнейшем понадобятся дальнейшие проверки. Гораздо удобнее использовать "достигаемый уровень значимости", т.е. вероятность того, что статистика критерия покажет большее отклонение от нулевой гипотезы, чем то, что соответствует имеющимся экспериментальным данным (см. терминологическое приложение 1 в конце книги). Если есть желание, можно сравнивать "достигаемый уровень значимости" с заданными значениями 0,05 или 0,01. Так, если "достигаемый уровень значимости" меньше 0,01, то нулевая гипотеза отвергается на уровне значимости 0,01, в противном случае - принимается. Следует рассчитывать "достигаемый уровень значимости" всегда, когда для этого есть вычислительные возможности.

Переход к "достигаемому уровню значимости" может избавить прикладника от еще одной трудности, связанной с использованием непараметрических критериев. Дело в том, что их распределения, как правило, дискретны, поскольку эти критерии используют только ранги наблюдений. Поэтому невозможно построить критерий с заданным номинальным уровнем значимости, реальный уровень значимости может принимать лишь конечное число значений, среди которых, как правило, нет ни 0,05, ни 0,01, ни других популярных номинальных значений.

Невозможность построения критических областей критериев с заданными уровнями значимости затрудняет сравнение критериев по мощности, как это продемонстрировано в работе [2]. Есть формальный способ достичь заданного номинального уровня значимости - провести рандомизацию, т.е. при определенном (граничном) значении статистики критерия провести независимый случайный эксперимент, в котором одни исходы (с заданной суммарной вероятностью) приводят к принятию гипотезы, а остальные - к ее отклонению. Однако подобную процедуру рандомизации прикладнику трудно принять - как оправдать то, что одни и те же экспериментальные данные могут быть основанием как для принятия гипотезы, так и для ее отклонения? Вспоминается обложка журнала "Крокодил", на которой один хозяйственник говорит другому: "Бросим монетку. Упадет гербом - будем строить завод, а упадет решкой - нет". Описанная процедура рандомизации имеет практический смысл лишь при массовой рутинной проверке гипотез, например, при статистическом контроле больших выборок изделий или деталей (см. главу 13, посвященную эконометрике качества).

У все еще распространенных критерия Стьюдента и других параметрических статистических критериев - свои проблемы. Они исходят из предположения о том, что функции распределения результатов наблюдений входят в определенные параметрические семейства небольшой размерности. Наиболее распространена гипотеза нормальности распределения. Однако давно известно, что подавляющее большинство реальных распределений результатов измерений не являются нормальными. Об этом говорится, например, в классической для инженеров и организаторов производства монографии проф. В.В. Налимова [3]. Ряд недавно полученных конкретных экспериментальных фактов и теоретических соображений рассмотрен в главе 4.

Как же быть? Проверять нормальность распределения своих данных? Но это дело непростое, можно допустить те или иные ошибки, в частности, применяя критерии типа Колмогорова или омега-квадрат (одна из наиболее распространенных ошибок состоит в том, что в статистики вместо неизвестных параметров подставляют их оценки, но при этом пользуются критическими значениями, рассчитанными для случая, когда параметры полностью известны [4]). Кроме того, для сколько-нибудь надежной проверки нормальности нужны тысячи наблюдений (см. главу 4). Поэтому в подавляющем большинстве реальных задач нет оснований принимать гипотезу нормальности. В лучшем случае можно говорить о том, что распределение результатов наблюдений мало отличается от нормального.

Как влияют отклонения от нормальности на свойства статистических процедур? Для разных процедур - разный ответ. Если речь идет об отбраковке выбросов - влияние отклонений от нормальности настолько велико, что делает процедуру отбраковки с практической точки зрения эвристической, а не научно обоснованной (см. главу 4). Если же речь идет о проверке однородности двух выборок с помощью критерия Стьюдента (при априорном предположении о равенстве дисперсий) или Крамера-Уэлча (при отсутствии такого предположения), то при росте объемов выборок влияние отклонений от нормальности убывает, как это подробно показано в главе 4). Это вытекает из Центральной Предельной Теоремы. Правда, при этом оказывается, что процентные точки распределения Стьюдента не приносят реальной пользы, достаточно использовать процентные точки предельного нормального распределения.

Весьма важна обсуждаемая, в частности, в работе [1] постоянно встающая перед эконометриком проблема выбора того или иного статистического критерия для решения конкретной прикладной задачи. Например, как проверять однородность двух независимых выборок числовых результатов наблюдений? Известны параметрические критерии:

Стьюдента, Лорда; непараметрические: Крамера-Уэлча, Вилкоксона, Ван-дер-Вардена, Сэвиджа, Мартынова, Смирнова, типа омега-квадрат (Лемана-Розенблатта) и многие другие (см., например, главу 4 и справочник [5]). Какой из них выбрать для конкретных расчетов?

Некоторые авторы предлагают формировать технологию принятия статистического решения, согласно которой решающее правило формируется на основе комбинации нескольких критериев. Например, технология может предусматривать проведение "голосования": если из 5 критериев большинство "высказывается" за отклонение гипотезы, то итоговое решение - отвергнуть ее, в противном случае - принять. Эти авторы не всегда понимают, что в их подходе нет ничего принципиально нового, просто к уже имеющимся критериям они добавляют их комбинации - очередные варианты, тем или иным образом выделяющие критические области в пространствах возможных значений результатов измерений, т.е. увеличивают число рассматриваемых критериев.

Итак, имеется некоторая совокупность критериев. У каждого - свой набор значений уровней значимости и мощностей на возможных альтернативах. Математическая статистика демонстрирует в этой ситуации виртуозную математическую технику для анализа частных случаев и полную беспомощность при выдаче практических рекомендаций. Так, оказывается, что практически каждый из известных критериев является оптимальным в том или ином смысле для какого-то набора нулевых гипотез и альтернатив. Математики изучают асимптотическую эффективность в разных смыслах - по Питмену, по Бахадуру и т.д., но - для узкого класса альтернативных гипотез, обычно для альтернативы сдвига. При попытке переноса асимптотических результатов на конечные объемы выборок возникают новые нерешенные проблемы, связанные, в частности, с численным оцениванием скорости сходимости (см. главу 10). В целом эта область математической статистики может активно развиваться еще многие десятилетия, выдавая "на гора" превосходные теоремы (которые могут послужить основанием для защит кандидатских и докторских диссертаций, выборов в академики РАН и т.д.), но не давая ничего практике. Хорошо бы, чтобы этот пессимистический прогноз не вполне оправдался!

С точки зрения эконометрики и прикладной статистики необходимо изучать проблему выбора критерия проверки однородности двух независимых выборок. Такое изучение было проведено, в том числе методом статистических испытаний, и в результате был получен вывод о том, что наиболее целесообразно применять критерий Лемана-Розенблатта типа омега-квадрат (см. главу 4).

В литературе по прикладным статистическим методам, как справедливо замечает С.Г. Корнилов в работе [1], имеется масса ошибочных рекомендаций. Чего стоят хотя бы принципиально неверные государственные стандарты СССР по статистическим методам, а также соответствующие им стандарты СЭВ и ИСО, т.е. Международной организации по стандартизации (о них см. главу 13, а также работу [6]). Особо выделяются своим количеством ошибочные рекомендации по применению критерия Колмогорова для проверки нормальности (см. ссылки в работе [4]). Ошибки есть и в научных статьях, и в нормативных документах (государственных стандартах), и в методических разработках, и даже в вузовских учебниках. К сожалению, нет способа оградить инженера и научного работника, экономиста и менеджера, нуждающихся в применении эконометрических и статистических методов, от литературных источников и нормативно-технических и инструктивно-методических документов с ошибками, неточностями и погрешностями. Единственный способ - либо постоянно поддерживать профессиональные контакты с квалифицированными специалистами в эконометрике, либо самому стать таким специалистом.

Как оценить достигаемый уровень значимости конкретного критерия, предусматривающего повторные проверки? Сразу ясно, что в большинстве случаев никакая современная теория математической статистики не поможет. Остается использовать современные компьютеры. Методика статистического моделирования, описанная в работе [1], может стать ежедневным рабочим инструментом специалиста, занимающегося применением эконометрических методов. Для этого она должна быть реализована в виде соответствующей диалоговой программной системы. Современные персональные компьютеры позволяют проводить статистическое моделирование весьма быстро (за доли секунд). Можно использовать различные модификации бутстрепа - одного из вариантов применения статистического моделирования (см. ниже).

Проведенное обсуждение показывает, как много нерешенных проблем стоит перед специалистом, занимающимся, казалось бы, рутинным применением стандартных статистических процедур. Эконометрика - молодая наука, ее основные проблемы, по нашему мнению, еще не до конца решены. Много работы как в сравнительно новых областях, например, в анализе нечисловых и интервальных данных (см. главы 8 и 9 выше), так и в классических.

11.2. Проблемы разработки и обоснования статистических технологий

В настоящем пункте рассматриваются проблемы практического использования эконометрических методов для системного анализа конкретных экономических данных. При этом применяются не отдельные методы описания данных, оценивания,проверки гипотез, а развернутые цельные процедуры - так называемые "статистические технологии". Понятия "статистические технологии" или "эконометрические технологии" аналогичны понятию "технологический процесс" в теории организации производства.

Статистические технологии. Поскольку термин "технология" сравнительно редко используется применительно к эконометрике и статистике, поясним суть рассматриваемой проблемы. Статистический анализ конкретных экономических данных, как правило, включает в себя целый ряд процедур и алгоритмов, выполняемых последовательно, параллельно или по более сложной схеме. В частности, с точки зрения менеджера эконометрического проекта можно выделить следующие этапы:

- планирование статистического исследования (включая разработку форм учета, их апробацию; подготовку сценариев интервью и анализа данных и т.п.);

- организация сбора необходимых статистических данных по оптимальной или рациональной программе (планирование выборки, создание организационной структуры и подбор команды статистиков, подготовка кадров, которые будут заниматься сбором данных, а также контролеров данных и т.п.);

- непосредственный сбор данных и их фиксация на тех или иных носителях (с контролем качества сбора и отбраковкой ошибочных данных по соображениям предметной области);

- первичное описание данных (расчет различных выборочных характеристик, функций распределения, непараметрических оценок плотности, построение гистограмм, корреляционных полей, различных таблиц и диаграмм и т.д.),

- оценивание тех или иных числовых или нечисловых характеристик и параметров распределений (например, непараметрическое интервальное оценивание коэффициента вариации или восстановление зависимости между откликом и факторами, т.е. оценивание функции),

- проверка статистических гипотез (иногда их цепочек - после проверки предыдущей гипотезы принимается решение о проверке той или иной последующей гипотезы; например, после проверки адекватности линейной регрессионной модели и отклонения этой гипотезы может проверяться адекватность квадратичной модели),

- более углубленное изучение, т.е. одновременное применение различных алгоритмов многомерного статистического анализа, алгоритмов диагностики и построения классификации, статистики нечисловых и интервальных данных, анализа временных рядов и др.;

- проверка устойчивости полученных оценок и выводов относительно допустимых отклонений исходных данных и предпосылок используемых вероятностно-статистических моделей, в частности, изучение свойств оценок методом размножения выборок и другими численными методами;

- применение полученных статистических результатов в прикладных целях, т.е. для формулировки выводов в терминах содержательной области (например, для диагностики конкретных материалов, построения прогнозов, выбора инвестиционного проекта из предложенных вариантов, нахождения оптимальных режима осуществления технологического процесса, подведения итогов испытаний образцов технических устройств и др.),

- составление итоговых отчетов, в частности, предназначенных для тех, кто не является специалистами в статистических методах анализа данных, в том числе для руководства - "лиц, принимающих решения".

Возможны и многие иные структуризации различных статистических технологий (см., например, аналогичную схему для процедур экспертных оценок в главе 12). Важно подчеркнуть, что квалифицированное и результативное применение статистических методов - это отнюдь не проверка одной отдельно взятой статистической гипотезы или оценка характеристик или параметров одного заданного распределения из фиксированного семейства. Подобного рода операции - только отдельные кирпичики, из которых складывается статистическая технология. Между тем учебники и монографии по статистике обычно рассказывают только об отдельных кирпичиках, но не обсуждают проблемы их организации в технологию, предназначенную для прикладного использования.

Итак, процедура статистического анализа данных – это информационный технологический процесс, другими словами, та или иная информационная технология. Статистическая информация подвергается разнообразным операциям (последовательно, параллельно или по более сложным схемам). В настоящее время об автоматизации всего процесса статистического анализа данных говорить было бы несерьезно, поскольку имеется слишком много нерешенных проблем, вызывающих дискуссии среди эконометриков и статистиков. Так называемые "экспертные системы" в области статистического анализа данных пока не стали рабочим инструментом статистиков. Ясно, что и не могли стать. Можно сказать и жестче - это пока научная фантастика или даже вредная утопия.

Проблема "стыковки" алгоритмов. В литературе статистические технологии рассматриваются явно недостаточно. В частности, обычно все внимание сосредотачивается на том или ином элементе технологической цепочки, а переход от одного элемента к другому остается в тени. Между тем проблема "стыковки" статистических алгоритмов, как известно, требует специального рассмотрения (см. предыдущий пункт настоящей главы), поскольку в результате использования предыдущего алгоритма зачастую нарушаются условия применимости последующего. В частности, результаты наблюдений могут перестать быть независимыми, может измениться их распределение и т.п.

Так, вполне резонной выглядит рекомендация: сначала разбейте данные на однородные группы, а потом в каждой из групп проводите статистическую обработку, например, регрессионный анализ. Однако эта рекомендация под кажущейся прозрачностью содержит подводные камни. Действительно, как поставить задачу в вероятностно-статистических терминах? Если, как обычно, примем, что исходные данные - это выборка, т.е. совокупность независимых одинаково распределенных случайных элементов, то классификация приведет к разбиению этих элементов на группы. В каждой группе элементы будут зависимы между собой, а их распределение будет зависеть от группы, куда они попали. Отметим, что в типовых ситуациях границы классов стабилизируются, а это значит, что асимптотически элементы кластеров статновятся независимыми. Однако их распределение не может быть нормальным. Например, если исходное распределение было нормальным, то распределения в классах будет усеченным нормальным. Это означает, что необходимо пользоваться непараметрическими методами, о чем уже не раз говорилось в главах 4 и 5 (подробнее этот пример разобран в работе [7]).

Разберем другой пример При проверке статистических гипотез большое значение имеют такие хорошо известные характеристики статистических критериев, как уровень значимости и мощность. Методы их расчета и использования при проверке одной гипотезы обычно хорошо известны. Если же сначала проверяется одна гипотеза, а потом с учетом результатов ее проверки (конкретнее, если первая гипотеза принята) - вторая, то итоговая процедура, которую также можно рассматривать как проверку некоторой (более сложной) статистической гипотезы, имеет характеристики (уровень значимости и мощность), которые, как правило, нельзя простыми формулами выразить через характеристики двух составляющих гипотез, а потому они обычно неизвестны. Лишь в некоторых простых случаях характеристики итоговой процедуры можно рассчитать (см. примеры в главе 13). В результате итоговую процедуру нельзя рассматривать как научно обоснованную, она относится к эвристическим алгоритмам. Конечно, после соответствующего изучения, например, методом Монте-Карло, она может войти в число научно обоснованных процедур эконометрики или прикладной статистики.

О термине "высокие статистические технологии". Как понятно, технологии бывают разные. Бывают адекватные и неадекватные, современные и устаревшие. Обратим внимание на термин "высокие технологии". Он популярен в современной научно-технической литературе и используется для обозначения наиболее передовых технологий, опирающихся на последние достижения научно-технического прогресса. Есть такие технологии и среди технологий эконометрического и статистического анализа данных - как в любой интенсивно развивающейся научно-практической области.

Примеры высоких эконометрических и статистических технологий и входящих в них алгоритмов анализа экономических данных постоянно обсуждаются на страницах настоящей книги. Подробный анализ современного состояния и перспектив развития эконометрики дан в главе 15 при обсуждении “точек роста” нашей научно-практической дисциплины. В этой главе в качестве примеров "высоких статистических технологий" выделены технологии непараметрического анализа данных (см. главы 4, 5 и 6); устойчивые (робастные) технологии (см. главу 10); технологии, основанные на размножении выборок (см. ниже в настоящей главе), на использовании достижений статистики нечисловых данных (см. главы 8 и 12) и статистики интервальных данных (см. главу 9).

Подробнее обсудим здесь пока не вполне привычный термин "высокие статистические технологии". Каждое из трех слов несет свою смысловую нагрузку.

"Высокие", как и в других научно-технических областях, означает, что статистическая технология опирается на современные научные достижения и передовой опыт реальной деятельности, а именно, достижения эконометрической и статистической теории и практики, в частности, на современные результаты теории вероятностей и прикладной математической статистики. При этом формулировка "опирается на современные научные достижения" означает, во-первых, что математическая основа технологии получена сравнительно недавно в рамках соответствующей научной дисциплины, во-вторых, что алгоритмы расчетов разработаны и обоснованы в соответствии с нею (а не являются т.н. "эвристическими"). Со временем, если новые подходы и результаты не заставляют пересмотреть оценку применимости и возможностей технологии, заменить ее на более современную, "высокие статистические технологии" переходят в "классические статистические технологии", такие, как метод наименьших квадратов. Как известно, несмотря на солидный возраст (более 200 лет), метод наименьших квадратов остается одним из наиболее часто используемых эконометрических методов. Итак, высокие статистические технологии - плоды недавних серьезных научных исследований. Здесь два ключевых понятия - "молодость" технологии (во всяком случае, не старше 50 лет, а лучше - не старше 10 или 30 лет), и опора на "высокую науку".

Термин "статистические" привычен, но разъяснить его нелегко. Во всяком случае, к деятельности Государственного комитета РФ по статистике высокие статистические технологии непосредственного отношения не имеют. В главе 1 уже шла речь о том разрыве между различными группами лиц, употребляющих термин "статистика", который имеется в нашей стране. Впрочем, сам термин "статистика" прошел долгий путь. Как известно, сотрудники проф. В.В. Налимова, одного из наиболее известных отечественных статистиков XX в., собрали более 200 определений термина "статистика" [8]. Полемика вокруг терминологии иногда принимает весьма острые формы (см., например, редакционные замечания к статье [9], написанные в стиле известных высказываний о генетике и кибернетике 1940-х годов - впрочем, каких-либо организационных выводов не последовало). Современное представление о терминологии в области теории вероятностей и прикладной математической статистики отражено в приложении 1 к настоящей книге, подготовленной в противовес распространенным ошибкам и неточностям в этой области. В частности, с точки зрения эконометрики статистические данные – это результаты измерений, наблюдений, испытаний, анализов, опытов, а "статистические технологии" - это технологии анализа статистических данных.

Всегда ли нужны "высокие статистические технологии"? "Высоким статистическим технологиям" противостоят, естественно, "низкие статистические технологии" (а между ними расположены "классические статистические технологии"). Это те технологии, которые не соответствуют современному уровню науки и практики. Обычно они одновременно и устарели, и не вполне адекватны сути решаемых эконометрических и статистических задач.

Примеры таких технологий неоднократно критически рассматривались, в том числе и на страницах этой книги. Достаточно вспомнить критику использования критерия Стьюдента для проверки однородности при отсутствии нормальности и равенства дисперсии или критику применения классических процентных точек критериев Колмогорова и омега-квадрат в ситуациях, когда параметры оцениваются по выборке и эти оценки подставляются в "теоретическую" функцию распределения (подробный разбор проведен, например, в работе [4]). Приходилось констатировать широкое распространение таких порочных технологий и конкретных алгоритмов, в том числе в государственных и международных стандартах (перечень ошибочных стандартов дан в работе [10]), учебниках и распространенных пособиях. Тиражирование ошибок происходит обычно в процессе обучения в вузах или путем самообразования при использовании недоброкачественной литературы.

На первый взгляд вызывает удивление устойчивость «низких статистических технологий», их постоянное возрождение во все новых статьях, монографиях, учебниках. Поэтому, как ни странно, наиболее "долгоживущими" оказываются не работы, посвященные новым научным результатам, а публикации, разоблачающие ошибки, типа статьи [4]. Прошло больше 15 лет с момента ее публикации, но она по-прежнему актуальна, поскольку ошибочное применение критериев Колмогорова и омега-квадрат по-прежнему распространено.

Целесообразно рассмотреть здесь по крайней мере четыре обстоятельства, которые определяют эту устойчивость ошибок.

Во-первых, прочно закрепившаяся традиция. Учебники по т.н. "Общей теории статистики", написанные "чистыми" экономистами (поскольку учебная дисциплина "Статистика" официально относится к экономике), если беспристрастно проанализировать их содержание, состоят в основном из введения в прикладную статистику, изложенного в стиле "низких статистических технологий", т.е. на уровне 1950-х годов, а во многом и на уровне начала XX в.. К "низкой" прикладной статистике добавлена некоторая информация о деятельности органов Госкомстата РФ. Некорректно обвинять только экономистов - примерно таково же положение со статистическими методами в медицине: одни и те же "низкие статистические технологии" переписываются из книги в книгу. Новое поколение, обучившись ошибочным подходам, идеям, алгоритмам, их использует, а с течением времени и достижением должностей, ученых званий и степеней – пишет новые учебники со старыми ошибками.

Второе обстоятельство связано с большими трудностями при оценке экономической эффективности применения статистических методов вообще и при оценке вреда от применения ошибочных методов в частности. (А без такой оценки как докажешь, что "высокие статистические технологии" лучше "низких"?) Некоторые соображения по первому из этих вопросов приведены в статье [9], содержащей оценки экономической эффективности ряда работ по применению статистических методов (см. также главу 13, посвященную эконометрике качества). При оценке вреда от применения ошибочных методов приходится учитывать, что общий успех в конкретной инженерной или научной работе вполне мог быть достигнут вопреки применению ошибочных методов, за счет "запаса прочности" других составляющих общей работы. Например, преимущество одного технологического приема над другим можно продемонстрировать как с помощью критерия Крамера-Уэлча проверки равенства математических ожиданий (что правильно), так и с помощью двухвыборочного критерия Стьюдента (что, вообще говоря, неверно, т.к. обычно не выполняются условия применимости этого критерия - нет ни нормальности распределения, ни равенства дисперсий). Кроме того, приходится выдерживать натиск невежд, защищающих свои ошибочные работы, например, государственные стандарты. Вместо исправления ошибок применяются самые разные приемы бюрократической борьбы с теми, кто разоблачает ошибки.

Третье существенное обстоятельство – трудности со знакомством с высокими статистическими технологиями. В нашей стране в силу ряда исторических обстоятельств развития статистических методов и эконометрики (см. главу 1) в течение последних 10 лет только журнал "Заводская лаборатория" предоставлял такие возможности. К сожалению, поток современных отечественных и переводных статистических книг, выпускавшихся ранее, в частности, издательствами "Наука", "Мир", “Финансы и статистика”, практически превратился в узкий ручеек… Возможно, более существенным является влияние естественной задержки во времени между созданием "новых статистических технологий" и написанием полноценной и объемной учебной и методической литературы. Она должна позволять знакомиться с новой методологией, новыми методами, теоремами, алгоритмами, методами расчетов и интерпретации ихъ результатов, статистическими технологиями в целом не по кратким оригинальным статьям, а при обычном вузовском и последипломном обучении.

И, наконец, наиболее важное. Всегда ли нужны высокие статистические технологии? Приведем аналогию - нужна ли современная сельскохозяйственная техника для обработки приусадебногоучастка? Нужны ли трактора и комбайны? Может быть, достаточно технологий, основанных на использовании лопаты? Вернемся к данным государственной статистики. Применяются статистические технологии первичной обработки (описания) данных, основанные на построении разнообразных таблиц, диаграмм, графиков. Большинство потребителей статистической информации это представление данных удовлетворяет. Итак, чтобы высокие статистические технологии успешно использовались, необходимы два условия: чтобы они были *объективно* нужны для решения практической задачи и чтобы потенциальный пользователь технологий *субъективно* понимал это.

Таким образом, весь арсенал реально используемых в настоящее время эконометрических и статистических технологий можно распределить по трем потокам:

- высокие статистические технологии;
- классические статистические технологии,

- низкие статистические технологии.

Под классическими статистическими технологиями, как уже отмечалось, понимаем технологии почтенного возраста, сохранившие свое значение для современной статистической практики. Таковы технололгии на основе метода наименьших квадратов (включая методы точечного оценивания параметров прогностической функции, непараметрические методы доверительного оценивания параметров, прогностической функции,проверок различных гипотез о них - см.главу 5), статистик типа Колмогорова, Смирнова, омега-квадрат, непараметрических коэффициентов корреляции Спирмена и Кендалла (относить их только к методам анализа ранжировок - значит делать уступку "низким статистическим технологиям", см. главу 5) и многих других статистических процедур.

Основная современная проблема в области эконометрических и статистических технологий состоит в том, чтобы в конкретных эконометрических исследованиях использовались только технологии первых двух типов.

Каковы возможные пути решения основной современной проблемы в области статистических технологий?

Бороться с конкретными невеждами - дело почти безнадежное. Отстаивая свое положение и должности, они либо нагло игнорируют информацию о своих ошибках, как это обычно делают авторы учебников по "Общей теории статистики" и их издатели, либо с помощью различных бюрократических приемов уходят и от ответственности, и от исправления ошибок по существу (как это было со стандартами по статистическим методам - см. статью [10]). Третий вариант - признание и исправление ошибок - встречается, увы, редко. Но встречается.

Конечно, необходима демонстрация квалифицированного применения высоких статистических технологий. В 1960-70-х годах этим занималась Лаборатория статистических методов акад. А.Н. Колмогорова в МГУ им. М.В. Ломоносова. Секция "Математические методы исследования" журнала "Заводская лаборатория" опубликовала за последние 40 лет более 1000 статей в стиле "высоких статистических технологий". В настоящее время действует Институт высоких статистических технологий и эконометрики МГТУ им. Н.Э.Баумана. Есть, конечно, целый ряд других научных коллективов, работающих на уровне "высоких статистических технологий".

Очевидно, самое основное - это обучение. Какие бы новые научные результаты ни были получены, если они остаются неизвестными студентам, то новое поколение экономистов ии менеджеров, исследователей и инженеров вынуждено осваивать их поодиночке, в порядке самообразования, а то и переоткрывать. Т.е. практически новые научные результаты почти исчезают, едва появившись. Как ни странно это может показаться, избыток научных публикаций превратился в тормоз развития науки. По нашим данным, к настоящему времени по эконометрическим и статистическим технологиям опубликовано не менее миллиона статей и книг, в основном во второй половине XX в., из них не менее 100 тысяч являются актуальными для современного специалиста. При этом реальное число публикаций, которые способен освоить исследователь за свою проофессиональную жизнь, по нашей оценке, не превышает 2-3 тысяч. Во всяком случае, в наиболее "толстом" на русском языке трехтомнике по статистике М. Дж. Кендалла и А. Стьюарта приведено около 2 тысяч литературных ссылок. Итак, каждый исследователь-эконометрик знаком не более чем с 2-3% актуальных для него литературных источников. Поскольку существенная часть публикаций заражена "низкими статистическими технологиями", то исследователь-самоучка, увы, имеет мало шансов выйти на уровень "высоких статистических технологий". С подтверждениями этого печального вывода постоянно приходится сталкиваться. Одновременно приходится констатировать, что масса полезных результатов погребена в изданиях прошлых десятилетий и имеет мало шансов пробиться в ряды используемых в настоящее время "высоких статистических технологий" без специально организованных усилий современных специалистов.

Итак, основное - обучение. Несколько огрубляя, можно сказать так: что попало в учебные курсы и соответствующие учебные пособия - то сохраняется, что не попало - то пропадает.

Необходимость высоких статистических технологий. Может возникнуть естественный вопрос: зачем нужны высокие статистические технологии, разве недостаточно обычных статистических методов? Специалисты по эконометрике справедливо считают и доказывают своими теоретическими и прикладными работами, что совершенно недостаточно. Так, совершенно очевидно, что многие данные в информационных системах имеют нечисловой характер, например, являются словами или принимают значения из конечных множеств. Нечисловой характер имеют и упорядочения, которые дают эксперты или менеджеры, например, выбирая главную цель, следующую по важности и т.д. Значит, нужна статистика нечисловых данных. Мы ее построили. Далее, многие величины известны не абсолютно точно, а с некоторой погрешностью - от и до. Другими словами, исходные данные - не числа, а интервалы. Нужна статистика интервальных данных. Мы ее развиваем. В широко известной монографии по контроллингу [11] на с.138 хорошо сказано: "Нечеткая логика - мощный элегантный инструмент современной науки, который на Западе (и на Востоке - в Японии, Китае - А.О.) можно встретить в десятках изделий - от бытовых видеокамер до систем управления вооружениями, - у нас до самого последнего времени был практически неизвестен". Напомним, первая монография российского автора по теории нечеткости содержит основы высоких статистических технологий анализа выборок нечетких множеств (см. книгу [12]). Ни статистики нечисловых данных, ни статистики интервальных данных, ни статистики нечетких данных нет и не могло быть в классической статистике. Все это - высокие статистические технологии. Они разработаны за последние 10-30-50 лет. А обычные вузовские курсы по общей теории статистики и по математической статистике разбирают научные результаты, полученные в первой половине XX века.

Важная и весьма перспективная часть эконометрики - применение высоких статистических технологий к анализу конкретных экономических данных, что зачастую требует дополнительной теоретической работы по доработке статистических технологий применительно к конкретной ситуации. Большое значение имеют конкретные эконометрические модели, например, модели экспертных оценок или эконометрики качества. И конечно, такие конкретные применения, как расчет и прогнозирование индекса инфляции. Сейчас уже многих ясно, что годовой бухгалтерский баланс предприятия может быть использован для оценки его финансово-хозяйственной деятельности только с привлечением данных об инфляции (см. главу 7).

О подготовке специалистов по высоким статистическим технологиям. Приходится с сожалением констатировать, что в России практически отсутствует подготовка специалистов по высоким статистическим технологиям. В курсах по теории вероятностей и математической статистике обычно даются лишь классические основы этих дисциплин, разработанные в первой половине XX в., а преподаватели-математики свою научную деятельность предпочитают посвящать доказательству теорем, имеющих лишь внутриматематическое значение, а не развитию высоких статистических технологий. В настоящее время появилась надежда на эконометрику. В России начинают развертываться эконометрические исследования и преподавание эконометрики. Экономисты, менеджеры и инженеры, прежде всего специалисты по контроллингу, должны быть вооружены современными средствами информационной поддержки, в том числе высокими статистическими технологиями и эконометрикой. Очевидно, преподавание должно идти впереди практического применения. Ведь как применять то, чего не знаешь?

Приведем два примера - отрицательный и положительный, - показывающие связь преподавания с внедрением передовых технологий.

Один раз - в 1990-1992 гг. мы уже обожглись на недооценке необходимости предварительной подготовки тех, для кого предназначены современные программные продукты. Наш коллектив (Всесоюзный центр статистических методов и информатики Центрального Правления Всесоюзного экономического общества) разработал систему диалоговых программных систем обеспечения качества продукции. Их созданием руководили ведущие специалисты страны. Но распространение программных продуктов шло на 1-2 порядка медленнее, чем мы ожидали. Причина стала ясна не сразу. Как оказалось, работники предприятий просто не понимали возможностей разработанных систем, не знали, какие задачи можно решать с их помощью, какой экономический эффект они дадут. А не понимали и не знали потому, что в вузах никто их не учил статистическим методам управления качеством. Без такого систематического обучения нельзя обойтись - сложные концепции "на пальцах" за 5 минут не объяснишь.

Есть и противоположный пример - положительный. В середине 1980-х годов в советской средней школе ввели новый предмет "Информатика". И сейчас молодое поколение превосходно владеет компьютерами, мгновенно осваивая быстро появляющиеся новинки, и этим заметно отличается от тех, кому за 30-40 лет.

Если бы удалось ввести в средней школе курс теории вероятностей и статистики - а такой курс есть в Японии и США, Швейцарии, Кении и Ботсване, почти во всех странах (и ЮНЕСКО проводит всемирные конференции по преподаванию математической статистики в средней школе - см сборник докладов [13]) - то ситуация с внедрением высоких статистических технологий могла бы быть резко улучшена. Надо, конечно, добиться того, чтобы такой курс был построен на высоких статистических технологиях, а не на низких. Другими словами, он должен отражать современные достижения, а не концепции пятидесятилетней или столетней давности.

11.3. Методы статистических испытаний (Монте-Карло) и датчики псевдослучайных чисел

Многие эконометрические информационные технологии опираются на использование методов статистических испытаний. Этот термин применяется для обозначения компьютерных технологий, в которых в эконометрическую модель искусственно вводится большое число случайных элементов. Обычно моделируется последовательность независимых одинаково распределенных случайных величин или же последовательность, построенная на основе такой, например, последовательность накапливающихся (кумулятивных) сумм.

Необходимость в методе статистических испытаний возникает потому, что чисто теоретические методы дают точное решение, как правило, лишь в исключительных случаях. Либо тогда, когда исходные случайные величины имеют вполне определенные функции распределения, например, нормальные, чего, как правило, не бывает. Либо когда объемы выборок очень велики (с практической точки зрения - бесконечны). Эта проблема уже обсуждалась в главе 10.

Не только в чисто эконометрических задачах обработки статистических данных возникает необходимость в методе статистических испытаний. Она не менее актуальна и при экономико-математическом моделировании технико-экономических и торговых процессов. Представим себе всем знакомый объект - торговый зал самообслуживания по продаже продовольственных товаров. Сколько нужно работников в зале, сколько касс? Необходимо просчитать загрузку в разное время суток, в разные сезоны года, с учетом замены товаров и смены сотрудников. Нетрудно увидеть, что теоретическому анализу подобная система не поддается, а компьютерному - вполне.

Методы статистических испытаний стали развиваться после второй мировой войны с появлением компьютеров. Второе название - методы Монте-Карло - они получили по наиболее известному игорному дому, а точнее, по его рулетке, поскольку исходный материал для получения случайных чисел с произвольным распределением - это случайные натуральные числа.

В методах статистических испытаний можно выделить две составляющие. Базой являются датчики псевдослучайных чисел. Результатом работы таких датчиков являются последовательности чисел, которые обладают некоторыми свойствами последовательностей случайных величин (в смысле теории вероятностей). Надстройкой являются различные алгоритмы, использующие последовательности псевдослучайных чисел.

Что же это могут быть за алгоритмы? Приведем примеры. Пусть мы изучаем распределение некоторой статистики при заданном объеме выборки. Тогда естественно много раз (например, 100000 раз) смоделировать выборку заданного объема (т.е. набор независимых одинаково распределенных случайных величин) и рассчитать значение статистики. Затем по 100000 значениям статистики можно достаточно точно построить функцию распределения изучаемой статистики, оценить ее характеристики. Однако эта схема годится лишь для так называемой "свободной от распределения" статистики, распределение которой не зависит от распределения элементов выборки. Если же такая зависимость есть, то одной точкой моделирования не обойдешься, придется много раз моделировать выборку, беря различные распределения, меняя параметры. Чтобы общее время моделирования было приемлемым, возможно, придется сократить число моделирований в одной точке, зато увеличив общее число точек. Точность моделирования может быть оценена по общим правилам выборочных обследований (см. главу 2).

Второй пример - частично описанное выше моделирование работы торгового зала самообслуживания по продаже продовольственных товаров. Здесь одна последовательность псевдослучайных чисел описывает интервалы между появлениями покупателей, вторая, третья и т.д. связаны с выбором ими первого, второго и т.д. товаров в зале (например, число - номер в перечне товаров). Короче, все действия покупателей, продавцов, работников предприятия разбиты на операции, каждая операция, в продолжительности или иной характеристике которой имеется случайность, моделируется с помощью соответствующей последовательности псевдослучайных чисел. Затем итоги работы сотрудников торговой организации и зала в целом выражаются через характеристики случайных величин. Формулируется критерий оптимальности, решается задача оптимизации и находятся оптимальные значения параметров.

Оптимальные планы статистического контроля, построенные на основе вероятностно-статистических моделей, строятся в главе 13.

Теперь обсудим свойства датчиков псевдослучайных чисел. Здесь стоит слово "псевдослучайные", а не "случайные". Это весьма важно.

Дело в том, что за последние 50 лет обсуждались в основном три принципиально разных варианта получения последовательностей чисел, которые в дальнейшем использовались в методах статистических испытаний.

Первый - таблица случайных чисел. К сожалению, объем любой таблицы конечен, и сколько-нибудь сложные расчеты с ее помощью невозможны. Через некоторое время приходится повторяться. Кроме того, обычно обнаруживались те или иные отклонения от случайности (см. об этом в работе [9]).

Второй - физические датчики случайных чисел. Основной недостаток - нестабильность, непредсказуемые отклонения от заданного распределения (обычно - равномерного).

Третий - расчетный. В простейшем случае каждый следующий член последовательности рассчитывается по предыдущему. Например, так:

$$z_{n+1} \equiv M z_n \pmod P,$$

где z_0 - начальное значение (заданное целое положительное число) M - параметр алгоритма (заданное целое положительное число), $P=2^m$, где m - число двоичных разрядов представления чисел, с которыми манипулирует компьютер. Знак $\equiv$здесь означает теоретико-числовую операцию сравнения, т.е. взятие дробной части от $\frac{Mz_n}{P}$ и отбрасывание целой.

В настоящее время применяется именно третий вариант. Совершенно ясно, что он не соответствует интуитивному представлению о случайности. Например, интуитивно очевидно, что по предыдущему элементу случайной последовательности с независимыми элементами нельзя предсказать значение следующего элемента. Расчетный путь получения последовательности псевдослучайных чисел противоречит не только интуиции, но и подходу к определению случайности на основе теории алгоритмов, развитому акад. А.Н. Колмогоровым и его учениками в 1960-х годах. Однако во многих прикладных задачах он работает, и это основное.

Методу статистических испытаний посвящена обширная литература (см., например, монографии [14-16]). Время от времени обнаруживаются недостатки у популярных датчиков псевдослучайных чисел. Так, например, в середине 1980-х годов выяснилось, что для одного из наиболее известных датчиков

$$Z_{n+2} = aZ_{n+1} + bZ_n, \quad n = 1,2,...$$

После этого в 1985 г. в журнале "Заводская лаборатория" началась дискуссия о качестве датчиков псевдослучайных чисел, которая продолжалась до 1993 г. и закончилась статьей проф. С.М.Ермакова [17] и нашим комментарием.

Итоги можно подвести так. Во многих случаях решаемая методом статистических испытаний задача сводится к оценке вероятности попадания в некоторую область в многомерном пространстве *фиксированной* размерности. Тогда из чисто математических соображений теории чисел следует, что с помощью датчиков псевдослучайных чисел поставленная задача решается корректно. Сводка соответствующих математических обоснований приведена, например, в работе С.М. Ермакова [17].

В других случаях приходится рассматривать вероятности попадания в области в пространствах *переменной* размерности. Типичным примером является ситуация, когда на каждом шагу проводится проверка, и по ее результатам либо остаемся в данном пространстве, либо переходим в пространство большей размерности. Например, в главе 5 при оценивании степени многочлена либо останавливались на данной степени, либо увеличивали степень, переходя в параметрическое пространство большей размерности. Так вот, вопрос об обоснованности применения метода статистических испытаний (а точнее, о свойствах датчиков псевдослучайных чисел) в случае пространств переменной размерности остается в настоящее время открытым. О важности этой проблемы говорил академик РАН Ю.В. Прохоров на Первом Всемирном Конгрессе Общества математической статистики и теории вероятностей им. Бернулли (Ташкент, 1986 г.).

Имитационное моделирование. Поскольку постоянно говорим о моделировании, приведем несколько общих формулировок.

Модель в общем смысле (обобщенная модель) - это создаваемый с целью получения и (или) хранения информации специфический объект (в форме мысленного образа, описания знаковыми средствами либо материальной системы), отражающей свойства, характеристики и связи объекта-оригинала произвольной природы, существенные для задачи, решаемой субъектом (это определение взято из монографии [18, с.44]).

Например, в менеджменте производственных систем используют:

- модели технологических процессов (контроль и управление по технико-экономическим критериям, АСУ ТП - автоматизированные системы управления технологическими процессами);

- модели управления качеством продукции (в частности, модели оценки и контроля надежности);

- модели массового обслуживания (теории очередей);

- модели управления запасами (в современной терминологии - модели логистики, т.е. теории и практики управления материальными, финансовыми и информационными потоками);

- имитационные и эконометрические модели деятельности предприятия (как единого целого) и управления им (АСУ предприятием) и др.

Согласно академику РАН Н.Н. Моисееву [19, с.213], имитационная система - это совокупность моделей, имитирующих протекание изучаемого процесса, объединенная со специальной системой вспомогательных программ и информационной базой, позволяющих достаточно просто и оперативно реализовать вариантные расчеты. Другими словами, имитационная система - это совокупность имитационных моделей. А имитационная модель предназначена для ответов на вопросы типа: "Что будет, если…" Что будет, если параметры примут те или иные значения? Что будет с ценой на продукцию, если спрос будет падать, а число конкурентов расти? Что будет, если государство резко усилит вмешательство в экономику? Что будет, если остановку общественного транспорта перенесут на 100 м дальше от входа в торговый зал, о котором шла речь выше, и поток покупателей резко упадет? Кроме компьютерных моделей, на вопросы подобного типа часто отвечают эксперты при использовании метода сценариев (см. главу 12).

При имитационном моделировании часто используется метод статистических испытаний (Монте-Карло). Теорию и практику машинных имитационных экспериментов с моделями экономических систем еще 30 лет назад подробно разобрал Т. Нейлор в обширной классической монографии [20]. Вернемся к внутриэконометрическому применению датчиков псевдослучайных чисел.

11.4. Методы размножения выборок (бутстреп-методы)

Эконометрика и прикладная статистика бурно развиваются последние десятилетия. Серьезным (хотя, разумеется, не единственным и не главным) стимулом является стремительно растущая производительность вычислительных средств. Поэтому понятен острый интерес к статистическим методам, интенсивно использующим компьютеры. Одним из таких методов является так называемый "бутстреп", предложенный в 1977 г. Б.Эфроном из Станфордского университета (США).

Сам термин "бутстреп" - это "bootstrap" русскими буквами и буквально означает что-то вроде: "вытягивание себя (из болота) за шнурки от ботинок". Термин специально придуман и заставляет вспомнить о подвигах барона Мюнхгаузена.

В истории эконометрики было несколько более или менее успешно осуществленных рекламных кампаний. В каждой из них "раскручивался" тот или иной метод, который, как правило, отвечал нескольким условиям:

- по мнению его пропагандистов, полностью решал актуальную научную задачу;

- был понятен (при постановке задачи, при ее решении и при интерпретации результатов) широким массам потенциальных пользователей;

- использовал современные возможности вычислительной техники.

Пропагандисты метода, как правило, избегали беспристрастного сравнения его возможностей с возможностями иных эконометрических методов. Если сравнения и проводились, то с заведомо слабым "противником".

В нашей стране в условиях отсутствия систематического эконометрического образования подобные рекламные кампании находили особо благоприятную почву, поскольку у большинства затронутых ими специалистов не было достаточных знаний в области методологии построения эконометрических моделей для того, чтобы составить самостоятельное квалифицированное мнение.

Речь идет о таких методах как бутстреп, нейронные сети, метод группового учета аргументов, робастные оценки по Тьюки-Хуберу (см. главу 10), асимптотика пропорционального роста числа параметров и объема данных и др. Бывают локальные всплески энтузиазма, например, московские социологи в 1980-х годах пропагандировали так называемый "детерминационный анализ" - простой эвристический метод анализа таблиц сопряженности, хотя в Новосибирске в это время давно уже было разработано продвинутое программное обеспечение анализа векторов разнотипных признаков (см. главу 8).

Однако даже на фоне всех остальных рекламных кампаний судьба бутстрепа исключительна. Во-первых, признанный его автор Б. Эфрон с самого начала признавался, что он ничего принципиально нового не сделал. Его исходная статья (первая в сборнике [21]) называлась: "Бутстреп-методы: новый взгляд на методы складного ножа". Во вторых, сразу появились статьи и дискуссии в научных изданиях, публикации рекламного характера, и даже в научно-популярных журналах. Бурные обсуждения на конференциях, спешный выпуск книг. В 1980-е годы финансовая подоплека всей этой активности, связанная с выбиванием грантов на научную деятельность, содержание учебных заведений и т.п. была мало понятна отечественным специалистам.

В чем основная идея группы методов "размножения выборок", наиболее известным представителем которых является бутстреп?

Пусть дана выборка $x_1, x_2, x_3, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-1}, x_n$. В вероятностно-статистической теории предполагаем, что это - набор независимых одинаково распределенных случайных величин. Пусть эконометрика интересует некоторая статистика $f_n(x_1, x_2, \ldots, x_n)$. Как изучить ее свойства? Подобными проблемами мы занимались на протяжении всей книги и знаем, насколько это непросто. Идея, которую предложил в 1949 г. М. Кенуй (это и есть "метод складного ножа") состоит в том, чтобы из одной выборки сделать много, исключая по одному наблюдению (и возвращая ранее исключенные). Перечислим выборки, которые получаются из исходной:

$$x_2, x_3, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-1}, x_n.;$$
$$x_1, x_3, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-1}, x_n.;$$
$$x_1, x_2, x_4, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-1}, x_n.;$$
$$\ldots$$
$$x_1, x_2, x_3, \ldots x_{k-1}, x_{k+1}, \ldots, x_{n-1}, x_n.;$$
$$\ldots$$
$$x_1, x_2, x_3, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-2}, x_n.;$$
$$x_1, x_2, x_3, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-1}.$$

Всего *n* новых (размноженных) выборок объемом *(n-1)* каждая. По каждой из них можно рассчитать значение интересующей эконометрика статистики (с уменьшенным на 1 объемом выборки):

$$f_{n-1,1}(\omega) = f_{n-1}(x_2, x_3, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-1}, x_n);$$
$$f_{n-1,2}(\omega) = f_{n-1}(x_1, x_3, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-1}, x_n);$$
$$f_{n-1,3}(\omega) = f_{n-1}(x_1, x_2, x_4, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-1}, x_n);$$
$$\ldots$$
$$f_{n-1,k}(\omega) = f_{n-1}(x_1, x_2, x_3, \ldots x_{k-1}, x_{k+1}, \ldots, x_{n-1}, x_n);$$
$$\ldots$$
$$f_{n-1,n-1}(\omega) = f_{n-1}(x_1, x_2, x_3, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-2}, x_n);$$
$$f_{n-1,n}(\omega) = f_{n-1}(x_1, x_2, x_3, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-1}).$$

Полученные значения статистики позволяют судить о ее распределении и о характеристиках распределения - о математическом ожидании, медиане, квантилях, разбросе, среднем квадратическом отклонении. Значения статистики, построенные по размноженным подвыборкам, не являются независимыми, однако, как мы видели в главе 5 на примере ряда статистик, возникающих в методе наименьших квадратов и в кластер-анализе (при обсуждении возможности объединения двух кластеров), при росте объема выборки влияние зависимости может ослабевать и со значениями статистик типа $f_{n-1,k}(\omega), \quad k = 1,2,...,n,$ можно обращаться как с независимыми случайными величинами.

Однако и без всякой вероятностно-статистической теории разброс величин $f_{n-1,k}(\omega), \quad k = 1,2,...,n,$ дает наглядное представление о том, какую точность может дать рассматриваемая статистическая оценка.

Сам М. Кенуй и его последователи использовали размножение выборок в основном для построения оценок с уменьшенным смещением. А вот Б. Эфрон преложил новый способ размножения выборок, существенно использующий датчики псевдослучайных чисел. А именно, он предложил строить новые выборки, *моделируя выборки из эмпирического распределения* (см. определения в терминологическом Приложении 1 в конце книги). Другими словами, Б. Эфрон предложил взять конечную совокупность из n элементов исходной выборки $x_1, x_2, x_3, ... x_{k-1}, x_k, x_{k+1}, ..., x_{n-1}, x_n$. и с помощью датчика случайных чисел сформировать из нее любое число размноженных выборок. Процедура, хотя и нереальна без ЭВМ, проста с точки зрения программирования. По сравнению с описанной выше процедурой появляются новые недостатки - неизбежные совпадения элементов размноженных выборок и зависимость от качества датчиков псевдослучайных чисел (см. выше). Однако существует математическая теория, позволяющая (при некоторых предположениях и безграничном росте объема выборки) обосновать процедуры бутстрепа (см. сборник статей [21]).

Есть много способов развития идеи размножения выборок (см., например, статью [22]). Можно по исходной выборке построить эмпирическую функцию распределения, а затем каким-либо образом от кусочно-постоянной функции перейти к непрерывной функции распределения, например, соединив точки $(x(i); \frac{i}{n}), i = 1,2,...n,$ отрезками прямых. Другой вариант - перейти к непрерывному распределению, построив непараметрическую оценку плотности. После этого рекомендуется брать размноженные выборки из этого непрерывного распределения (являющегося состоятельной оценкой исходного), непрерывность защитит от совпадений элементов в этих выборках.

Другой вариант построения размноженных выборок - более прямой. Исходные данные не могут быть определены совершенно точно и однозначно. Поэтому предлагается к исходным данным добавлять малые независимые одинаково распределенные погрешности. При таком подходе одновременно соединяем вместе идеи устойчивости (см. главу 10) и бутстрепа. При внимательном анализе многие идеи эконометрики тесно друг с другом связаны (см. статью [22]).

В каких случаях целесообразно применять бутстреп, а в каких - другие эконометрические методы? В период рекламной кампании встречались, в том числе в научно-популярных журналах, утверждения о том, что и для оценивания математического ожидания полезен бутстреп. Как показано в статье [22], это совершенно не так. При росте числа испытаний методом Монте-Карло бутстреп-оценка приближается к классической оценке - среднему арифметическому результатов наблюдений. Другими словами, бутстреп-оценка отличается от классической только шумом псевдослучайных чисел.

Аналогичной является ситуация и в ряде других случаев. Там, где эконометрическая теория хорошо развита, где найдены методы анализа данных, в том или иной смысле близкие к оптимальным, бутстрепу делать нечего. А вот в новых областях со сложными алгоритмами, свойства которых недостаточно ясны, он представляет собой ценный инструмент для изучения ситуации.

11.5.Эконометрика в контроллинге

Контроллеру и сотрудничающему с ним эконометрику нужна разнообразная экономическая и управленческая информация, не менее нужны удобные инструменты ее анализа. Следовательно, информационная поддержка контроллинга необходима для успешной работы контроллера. Без современных компьютерных инструментов анализа и управления, основанных на продвинутых эконометрических и экономико-математических методах и моделях, невозможно эффективно принимать управленческие решения. Недаром специалисты по контроллингу большое внимание уделяют проблемам создания, развития и применения компьютерных систем поддержки принятия решений. Высокие статистические технологии и эконометрика - неотъемлемые части любой современной системы поддержки принятия экономических и управленческих решений.

Важная часть эконометрики - применение высоких статистических технологий к анализу конкретных экономических данных. Такие исследования зачастую требуют дополнительной теоретической работы по "доводке" статистических технологий применительно к конкретной ситуации. Большое значение для контроллинга имеют не только общие методы, но и конкретные эконометрические модели, например, вероятностно-статистические модели тех или иных процедур экспертных оценок (глава 12) или эконометрики качества (глава 13), имитационные модели деятельности организации, прогнозирования в условиях риска (глава 14). И конечно, такие конкретные применения, как расчет и прогнозирование индекса инфляции. Сейчас уже многим специалистам ясно, что годовой бухгалтерский баланс предприятия может быть использован для оценки его финансово-хозяйственной деятельности только с привлечением данных об инфляции. Различные области экономической теории и практики в настоящее время еще далеко не согласованы. При оценке и сравнении инвестиционных проектов принято использовать такие характеристики, как чистая текущая стоимость, внутренняя норма доходности, основанные на введении в рассмотрение изменения стоимости денежной единицы во времени (это осуществляется с помощью дисконтирования). А при анализе финансово-хозяйственной деятельности организации на основе данных бухгалтерской отчетности изменение стоимости денежной единицы во времени по традиции не учитывают.

Специалисты по контроллингу должны быть вооружены современными средствами информационной поддержки, в том числе средствами на основе высоких статистических технологий и эконометрики. Очевидно, преподавание должно идти впереди практического применения. Ведь как применять то, чего не знаешь?

Статистические технологии применяют для анализа данных двух принципиально различных типов. Один из них - это результаты измерений (наблюдений, испытаний, анализов, опытов и др.) различных видов, например, результаты управленческого или бухгалтерского учета, данные Госкомстата и др. Короче, речь идет об объективной информации. Другой - это оценки экспертов, на основе своего опыта и интуиции делающих заключения относительно экономических явлений и процессов. Очевидно, это - субъективная информация. В стабильной экономической ситуации, позволяющей рассматривать длинные временные ряды тех или иных экономических величин, полученных в сопоставимых условиях, данные первого типа вполне адекватны. В быстро меняющихся условиях приходятся опираться на экспертные оценки. Такая новейшая часть эконометрики, как статистика нечисловых данных, была создана как ответ на запросы теории и практики экспертных оценок (см. главы 8 и 12).

Для решения каких экономических задач может быть полезна эконометрика? Практически для всех, использующих конкретную информацию о реальном мире. Только чисто абстрактные, отвлеченные от реальности исследования могут обойтись без нее. В частности, эконометрика необходима для прогнозирования, в том числе поведения потребителей, а потому и для планирования. Выборочные исследования, в том числе выборочный контроль, основаны на эконометрике. Но планирование и контроль - основа контроллинга. Поэтому эконометрика - важная составляющая инструментария контроллера, воплощенного в компьютерной системе поддержки принятия решений. Прежде всего оптимальных решений, которые предполагают опору на адекватные эконометрические модели. В производственном менеджменте это может означать, например, использование моделей экстремального планирования эксперимента (судя по накопленному опыту их практического использования, такие модели позволяют повысить выход полезного продукта на 30-300%).

Высокие статистические технологии в эконометрике предполагают адаптацию применяемых методов к меняющейся ситуации. Например, параметры прогностического индекса меняются вслед за изменением характеристик используемых для прогнозирования величин. Таков метод экспоненциального сглаживания. В соответствующем алгоритме расчетов значения временного ряда используются с весами. Веса уменьшаются по мере удаления в прошлое. Многие методы дискриминантного анализа основаны на применении обучающих выборок. Например, для построения рейтинга надежности банков можно с помощью экспертов составить две обучающие выборки - надежных и ненадежных банков. А затем с их помощью решать для вновь рассматриваемого банка, каков он - надежный или ненадежный, а также оценивать его надежность численно, т.е. вычислять значение рейтинга.

Один из способов построения адаптивных эконометрических моделей - нейронные сети (см., например, монографию [23]). При этом упор делается не на формулировку адаптивных алгоритмов анализа данных, а - в большинстве случаев - на построение виртуальной адаптивной структуры. Термин "виртуальная" означает, что "нейронная сеть" - это специализированная компьютерная программа, "нейроны" используются лишь при общении человека с компьютером. Методология нейронных сетей идет от идей кибернетики 1940-х годов. В компьютере создается модель мозга человека (весьма примитивная с точки зрения физиолога). Основа модели - весьма простые базовые элементы, называемые нейронами. Они соединены между собой, так что нейронные сети можно сравнить с хорошо знакомыми экономистам и инженерам блок-схемами. Каждый нейрон находится в одном из заданного множества состояний. Он получает импульсы от соседей по сети, изменяет свое состояние и сам рассылает импульсы. В результате состояние множества нейтронов изменяется, что соответствует проведению эконометрических вычислений.

Нейроны обычно объединяются в слои (как правило, два-три). Среди них выделяются входной и выходной слои. Перед началом решения той или иной задачи производится настройка. Во-первых, устанавливаются связи между нейронами, соответствующие решаемой задаче. Во-вторых, проводится обучение, т.е. через нейронную сеть пропускаются обучающие выборки, для элементов которых требуемые результаты расчетов известны. Затем параметры сети модифицируются так, чтобы получить максимальное соответствие выходных значений заданным величинам.

С точки зрения точности расчетов (и оптимальности в том или ином эконометрическом смысле) нейронные сети не имеют преимуществ перед другими адаптивными эконометрическими системами. Однако они более просты для восприятия. Надо отметить, что в эконометрике используются и модели, промежуточные между нейронными сетями и "обычными" системами регрессионных уравнений (одновременных и с лагами). Они тоже используют блок-схемы, как, например, универсальный метод моделирования связей экономических факторов ЖОК (этот метод описан в работе [24]).

Заметное место в математико-компьютерном обеспечении принятия решений в контроллинге занимают методы теории нечеткости (по-английски - *fuzzy theory*, причем термин *fuzzy* переводят на русский язык по-разному: нечеткий, размытый, расплывчатый, туманный, пушистый и др.). Начало современной теории нечеткости положено работой Л.А. Заде 1965г., хотя истоки прослеживаются со времен Древней Греции (об истории теории нечеткости см., например, книгу [12]). Это направление прикладной математики в последней трети XX в. получило бурное развитие. К настоящему времени по теории нечеткости опубликованы тысячи книг и статей, издается несколько международных журналов (половина - в Китае и Японии), постоянно проводятся международные конференции, выполнено достаточно много как теоретических, так и прикладных научных работ, практические приложения дали ощутимый технико-экономический эффект.

Основоположник рассматриваемого научного направления Лотфи А. Заде рассматривал теорию нечетких множеств как аппарат анализа и моделирования гуманистических систем, т.е. систем, в которых участвует человек. Его подход опирается на предпосылку о том, что элементами мышления человека являются не числа, а элементы некоторых нечетких множеств или классов объектов, для которых переход от "принадлежности" к "непринадлежности" не скачкообразен, а непрерывен. В настоящее время методы теории нечеткости используются почти во всех прикладных областях, в том числе при управлении качеством продукции и технологическими процессами.

Нечеткая математика и логика - мощный элегантный инструмент современной науки, который на Западе и на Востоке (в Японии, Китае, Корее) можно встретить в программном обеспечении сотен видов изделий - от игрушек и бытовых видеокамер до систем управления предприятиями. В России он был достаточно хорошо известен с начала 1970-х годов. Однако первая монография российского автора по теории нечеткости [12] была опубликована лишь в 1980 г. В дальнейшем проводившиеся раз в год всесоюзные конференции собирали около 100 участников - по мировым меркам немного. В настоящее время интерес к теории нечеткости среди экономистов и менеджеров растет.

При изложении теории нечетких множеств обычно не подчеркивается связь с вероятностными моделями. Между тем еще в середине 1970-х годов установлено (цикл соответствующих теорем приведен, в частности, в монографии [12], но это отнюдь не первая публикация), что теория нечеткости в определенном смысле сводится к теории случайных множеств, хотя эта связь и имеет, возможно, лишь теоретическое значение. В США подобные работы появились лет на пять позже.

Профессионалу в области контроллинга полезны многочисленные интеллектуальные инструменты анализа данных, относящиеся к высоким статистическим технологиям и эконометрике.

Цитированная литература

1. Корнилов С.Г. Накопление ошибки первого рода при повторной проверке статистических гипотез. Регламент повторных проверок. // Заводская лаборатория. 1996. Т.62. No.5. С. 45-51.
2. Камень Ю.Э., Камень Я.Э., Орлов А.И. Реальные и номинальные уровни значимости в задачах проверки статистических гипотез. // Заводская лаборатория.1986. Т.52. No.12. С.55-57.
3. Налимов В.В. Применение математической статистики при анализе вещества. - М.:Физматгиз,1960. - 430 с.
4. Орлов А.И. Распространенная ошибка при использовании критериев Колмогорова и омега-квадрат. // Заводская лаборатория. 1985. Т.51. No.1. С.60-62.
5. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. Изд.3-е.- М.: Наука, 1983. - 416 с.
6. Орлов А.И.О современных проблемах внедрения прикладной статистики и других статистических методов. // Заводская лаборатория. 1992. Т.58. No.1. С.67-74.

7. Орлов А.И. Некоторые вероятностные вопросы теории классификации. – В сб.: Прикладная статистика. Ученые записки по статистике, т.45. - М.: Наука, 1983. С.166-179.
8. Никитина Е.П., Фрейдлина В.Д., Ярхо А.В. Коллекция определений термина "статистика" / Межфакультетская лаборатория статистических методов. Вып.37. - М.: Изд-во Московского государственного университета им. М.В. Ломоносова, 1972. - 46 с.
9. Орлов А.И. Что дает прикладная статистика народному хозяйству? // Вестник статистики. - 1986. - No.8. - С.52-56.
10. Орлов А.И. Сертификация и статистические методы (обобщающая статья). // Заводская лаборатория. - 1997. - Т.63. - No.3. - С.55-62.
11. Контроллинг в бизнесе. Методологические и практические основы построения контроллинга в организациях / А.М. Карминский, Н.И. Оленев, А.Г. Примак, С.Г.Фалько. - М.: Финансы и статистика, 1998. - 256 с.
12. Орлов А. И. Задачи оптимизации и нечеткие переменные. - М.: Знание, 1980.- 64 с.
13. The teaching of statistics / Studies in mathematics education. Vol.7. - Paris, UNESCO, 1989. - 258 pp.
14. Ермаков С.М. Метод Монте-Карло и смежные вопросы. - М.: Наука, 1975. - 471 с.
15. Ермаков С.М., Михайлов Г.А. Статистическое моделирование. - М.: Наука, 1982. - 296 с.
16. Иванова И.М. Случайные числа и их применения. - М.: Финансы и статистика, 1984. - 111 с.
17. Ермаков С.М. О датчиках случайных чисел. // Заводская лаборатория. 1993. Т.59. No.7. С.48-50.
18. Неуймин Я.Г. Модели в науке и технике. История, теория, практика. - Л.: Наука, 1984. - 190 с.
19. Моисеев Н.Н. Математические задачи системного анализа. - М.: Наука, 1981. - 488 с.
20. Нейлор Т. Машинные имитационные эксперименты с моделями экономических систем. - М.: Мир, 1975. - 500 с.
21. Эфрон Б. Нетрадиционные методы многомерного статистического анализа. - М.: Финансы и статистика, 1988. - 263 с.
22. Орлов А.И. О реальных возможностях бутстрепа как статистического метода. // Заводская лаборатория. 1987. Т.53. No.10. С.82-85.
23. Бэстенс Д.Э., Берт В.М. ван дер, Вуд Д. Нейронные сети и финансовые рынки: принятие решений в торговых операциях. - М.: ТВП, 1998.
24. Орлов А.И., Жихарев В.Н., Кольцов В.Г. Новый эконометрический метод "ЖОК" оценки результатов взаимовлияний факторов в инженерном менеджменте // Проблемы технологии, управления и экономики / Под общей редакцией к. э. н. Панкова В.А. Ч.1. Краматорск: Донбасская государственная машиностроительная академия, 1999. С.87-89.

Глава 12. Эконометрические методы проведения экспертных исследований и анализа оценок экспертов

Что будет через десять лет? Как изменится социально-экономическая обстановка? Будет ли обеспечена экологическая безопасность промышленных производств или же вокруг раскинется рукотворная пустыня? Достаточно вдуматься в эти постановки естественных вопросов, проанализировать, как десять лет назад мы представляли себе сегодняшний день, чтобы понять, что стопроцентно надежных прогнозов просто не может быть. Вместо утверждений с наборами конкретных чисел можно ожидать лишь качественных оценок. Тем не менее мы должны принимать решения, например, об экономических, социальных, экологических и иных проектах, в том числе требующих крупных инвестиций. Их последствия скажутся через десять, двадцать и т.д. лет. А будущего мы не можем знать. Как быть? Остается обратиться к методам экспертных оценок. Что это за методы?

12.1. Примеры процедур экспертных оценок

Бесспорно совершенно, что для принятия обоснованных решений необходимо опираться на опыт, знания и интуицию специалистов. После второй мировой войны в рамках научного движения, включающего в себя эконометрику, кибернетику, теорию управления, менеджмент и исследование операций стала развиваться самостоятельная дисциплина - теория и практика экспертных оценок.

Методы экспертных оценок - это методы организации работы со специалистами-экспертами и анализа мнений экспертов. Эти мнения обычно выражены частично в количественной, частично в качественной форме. Экспертные исследования проводят с целью подготовки информации для принятия решений ЛПР (лицом, принимающим решения). Для проведения работы по методу экспертных оценок создают Рабочую группу (сокращенно РГ), которая и организует по поручению ЛПР деятельность экспертов, объединенных (формально или по существу) в экспертную комиссию (ЭК).

Экспертные оценки бывают *индивидуальные* и *коллективные*. *Индивидуальные оценки* - это оценки одного специалиста. Например, преподаватель единолично ставит отметку студенту, а врач - диагноз больному. Но в сложных случаях заболевания или угрозе отчисления студента за плохую учебу обращаются к *коллективному* мнению - симпозиуму врачей или комиссии преподавателей. Аналогичная ситуация - в армии. Обычно командующий принимает решение единолично. Но в сложных и ответственных ситуациях проводят военный совет. Один из наиболее известных примеров такого рода - военный совет 1812 г. в Филях, на котором под председательством М.И. Кутузова решался вопрос: "Давать или не давать французам сражение под Москвой?"

Другой простейший пример экспертных оценок - оценка исполненных командами номеров в КВН. Каждый из членов жюри поднимают фанерку со своей оценкой, а технический работник вычисляет среднюю арифметическую оценку, которая и объявляется как коллективное мнение жюри (отметим, что такой подход некорректен с точки зрения теории измерений).

В фигурном катании процедура усложняется - перед усреднением *отбрасываются самая большая и самая маленькая оценки.* Это делается для того, чтобы у судьи не было соблазна завысить оценку одной спортсменке (например, соотечественнице) или занизить другой. Такие резко выделяющиеся из общего ряда оценки будут сразу отброшены.

Экспертные оценки часто используются при выборе - одного варианта технических устройств из нескольких, группы космонавтов из многих претендентов, набора проектов научно-исследовательских работ для финансирования из массы заявок, получателей кредитов из многих желающих, при выборе инвестиционных проектов для реализации среди представленных, при проведении тендера (выбора исполнителей заказа из многих желающих) и т.д.

Существует масса методов получения экспертных оценок. В одних с каждым экспертом работают отдельно, он даже не знает, кто еще является экспертом, а потому высказывает свое мнение независимо от авторитетов. В других экспертов собирают вместе для подготовки материалов для ЛПР, при этом эксперты обсуждают проблему друг с другом, учатся друг у друга, и неверные мнения отбрасываются. В одних методах число экспертов фиксировано и таково, чтобы статистические методы проверки согласованности мнений и затем их усреднения позволяли принимать обоснованные решения. В других - число экспертов растет в процессе проведения экспертизы, например, при использовании метода "снежного кома" (о нем - дальше).

Не меньше существует и методов обработки ответов экспертов, в том числе весьма насыщенных математикой и компьютеризированных. Многие из них основаны на достижениях статистики объектов нечисловой природы и других современных методах эконометрики и прикладной статистики.

Один из наиболее известных методов экспертных оценок - это *метод "Дельфи"*. Название дано по ассоциации с древним обычаем для получения советов и поддержки при принятии решений обращаться в Дельфийский храм. Он был расположен у выхода ядовитых вулканических газов. Жрицы храма, надышавшись отравы, начинали пророчествовать, произнося непонятные слова. Специальные "переводчики" - жрецы храма толковали эти слова и отмечали на вопросы пришедших со своими проблемами паломников. По традиции говорят, что Дельфийский храм находился в Греции. Но там нет вулканов. Видимо, он был в Италии - у Везувия или Этны, а сами описанные предсказания происходили в XII-XIV вв. Эти места и даты вытекают из высшего достижения современной исторической науки - новой статистической хронологии.

В США в 1960-х годах методом Дельфи назвали экспертную процедуру прогнозирования научно-технического развития. В первом туре эксперты называли вероятные даты тех или иных будущих научно-технических свершений. Во втором туре каждый эксперт знакомился с прогнозами всех остальных (без указания фамилий авторов прогнозов). Если его прогноз сильно отличался от прогнозов основной массы, его просили пояснить свою позицию, и эксперт довольно часто изменял свои оценки, приближаясь к средним значениям. Эти средние значения и выдавались заказчику как групповое мнение. Надо сказать, что реальные результаты прогностического исследования оказались довольно скромными - хотя дата высадки американцев на Луну была предсказана с точностью до месяца, все остальные прогнозы провалились - холодного термоядерного синтеза и средства от рака в XX в. человечество не дождалось. Однако сама методика оказалась популярной - за последующие годы она использовалась не менее 40 тыс. раз. Средняя стоимость экспертного исследования по методу Дельфи - 5 тыс. долларов США, но в ряде случаев приходилось расходовать и более крупные суммы - до 130 тыс. долларов.

Несколько в стороне от основного русла экспертных оценок лежит *метод сценариев*, применяемый прежде всего для экспертного прогнозирования. Рассмотрим основные идеи технологии сценарных экспертных прогнозов. Социально-экономическое или экологическое прогнозирование, как и любое прогнозирование вообще, может быть успешным лишь при некоторой стабильности условий. Однако решения органов власти, отдельных лиц, иные события, например, землетрясения, меняют условия, в которых живет население и текут экономические процессы, и события развиваются по-иному, чем ранее предполагалось. Например, вполне очевидно, что после первого тура президентских выборов 1996 г. о дальнейшем развитии событий можно было говорить лишь в терминах сценариев: если победит Б.Н. Ельцин, то будет то-то и то-то, а если победит Г.А. Зюганов, то события пойдут так-то и так-то.

Метод сценариев необходим не только в экологической или социально-экономической области. Например, при разработке методологического, программного и информационного обеспечения *анализа риска* химико-технологических проектов необходимо составить детальный каталог сценариев аварий, связанных с утечками токсических химических веществ. Каждый из таких сценариев описывает аварию своего типа, со своим индивидуальным происхождением, развитием, последствиями, возможностями предупреждения.

Таким образом, метод сценариев - это метод декомпозиции задачи прогнозирования, предусматривающий выделение набора отдельных вариантов развития событий (сценариев), в совокупности охватывающих все возможные варианты развития. При этом каждый отдельный сценарий должен допускать возможность достаточно точного прогнозирования, а общее число сценариев должно быть обозримо.

Возможность подобной декомпозиции не очевидна. При применении метода сценариев необходимо осуществить два этапа исследования:

- построение исчерпывающего, но обозримого набора сценариев;
- прогнозирование в рамках каждого конкретного сценария с целью получения ответов на интересующие исследователя вопросы.

Каждый из этих этапов лишь частично формализуем. Существенная часть рассуждений проводится на качественном уровне, как это принято в общественно-экономических и гуманитарных науках. Одна из причин заключается в том, что стремление к излишней формализации и математизации приводит к *искусственному* внесению определенности там, где ее нет по существу, либо к использованию громоздкого математического аппарата. Так, рассуждения на словесном уровне считаются доказательными в большинстве ситуаций, в то время как попытка уточнить смысл используемых слов с помощью, например, теории нечетких множеств приводит к весьма громоздким математическим моделям.

Набор сценариев должен быть обозрим. Приходится исключать различные маловероятные события - прилет инопланетян, падение астероида, массовые эпидемии ранее неизвестных болезней, и т.д. Само по себе создание набора сценариев - предмет экспертного исследования. Кроме того, эксперты могут оценить вероятности реализации того или иного сценария.

Прогнозирование в рамках каждого конкретного сценария с целью получения ответов на интересующие исследователя вопросы также осуществляется в соответствии с описанной выше методологией прогнозирования. При стабильных условиях могут быть применены статистические методы прогнозирования временных рядов. Однако этому предшествует анализ с помощью экспертов, причем зачастую прогнозирование на словесном уровне является достаточным (для получения интересующих исследователя и ЛПР выводов) и не требующим количественного уточнения.

Как известно, при принятии решений на основе *анализа ситуации*, в том числе результатов прогнозных исследований, можно исходить из различных критериев. Так, можно ориентироваться на то, что ситуация сложится наихудшим, или наилучшим, или средним (в каком-либо смысле) образом. Можно попытаться наметить мероприятия, обеспечивающие минимально допустимые полезные результаты при любом варианте развития ситуации, и т.д.

Еще один вариант экспертного оценивания - *мозговой штурм*. Организуется он как собрание экспертов, на выступления которых наложено одно, но очень существенное ограничение - нельзя критиковать предложения других. Можно их развивать, можно высказывать свои идеи, но нельзя критиковать! В ходе заседания эксперты, "заражаясь" друг от друга, высказывают все более экстравагантные соображения. Часа через два записанное на магнитофон или видеокамеру заседание заканчивается, и начинается второй этап мозгового штурма - анализ высказанных идей. Обычно из 100 идей 30 заслуживают дальнейшей проработки, из 5-6 дают возможность сформулировать прикладные проекта, а 2-3 оказываются в итоге приносящими полезный эффект - прибыль, повышение технологической или экологической безопасности и т.п. При этом интерпретация идей - творческий процесс. Например, при обсуждении возможностей защиты кораблей от торпедной атаки была высказана идея: "Выстроить матросов вдоль борта и дуть на торпеду, чтобы изменить ее курс". После проработки эта идея привела к созданию специальных устройств, создающих волны, сбивающиеся торпеду с курса.

Более подробно рассмотрим отдельные этапы экспертного исследования.

12.2. Основные стадии экспертного опроса

Как показывает опыт проведения экспертных исследований, с точки зрения менеджера - организатора такого исследования целесообразно выделять следующие стадии проведения экспертного опроса.

1) *Принятие решения* о необходимости проведения экспертного опроса и формулировка Лицом, Принимающим Решения (ЛПР) его цели. Таким образом, инициатива должна исходить от руководства, что в дальнейшем обеспечит успешное решение организационных и финансовых проблем. Очевидно, что исходный толчок может быть дан докладной запиской одного из сотрудников или дискуссией на совещании, но реальное начало работы - решение ЛПР.

2) *Подбор и назначение ЛПР основного состава Рабочей группы,* сокращенно РГ (обычно - научного руководителя и секретаря). При этом научный руководитель отвечает за организацию и проведение экспертного исследования в целом, а также за анализ собранных материалов и формулировку заключения экспертной комиссии. Он участвует в формировании коллектива экспертов и выдаче задания каждому (вместе с ЛПР или его представителем). Дело секретаря - ведение документации экспертного опроса, решение организационных задач.

3) *Разработка* РГ (точнее, ее основным составом, прежде всего научным руководителем и секретарем) и утверждение у ЛПР *технического задания* на проведение экспертного опроса. На этой стадия решение о проведении экспертного опроса приобретает четкость во времени, финансовом, кадровом, материальном и организационном обеспечении. В частности, в РГ выделяются различные группы специалистов - аналитическая, эконометрическая (специалисты по методам), компьютерная, по работе с экспертами (например, интервьюеров), организационная. Очень важно для успеха, чтобы все эти позиции были утверждены ЛПР.

4) *Разработка* аналитической группой РГ подробного сценария (т.е. *регламента*) проведения сбора и анализа экспертных мнений (оценок). Сценарий включает в себя прежде всего конкретный вид информации, которая будет получена от экспертов (например, слова, условные градации, числа, ранжировки, разбиения или иные виды объектов нечисловой природы). Например, довольно часто экспертов просят высказаться в свободной форме, ответив при этом на некоторые количество заранее сформулированных вопросов. Кроме того, их просят заполнить формальную карту, в каждом пункте выбрав одну из нескольких градаций. Сценарий должен содержать и конкретные методы анализа собранной информации. Например, вычисление медианы Кемени, статистический анализ люсианов, применение иных методов статистики объектов нечисловой природы и других разделов прикладной статистики (о некоторых из названных методов речь пойдет ниже). Эта работа ложится на эконометрическую и компьютерную группу РГ. Традиционная ошибка - сначала собрать информацию, а потом думать, что с ней делать. В результате информация используется на 1-2%.

5) *Подбор экспертов* в соответствии с их компетентностью. На этой стадии РГ составляет список возможных экспертов.

6) *Формирование экспертной комиссии*. На этой стадии РГ проводит переговоры с экспертами, получает их согласие на работу в экспертной комиссии (сокращенно ЭК),

возможно, часть намеченных РГ экспертов отказывается по тем или иным причинам. ЛПР утверждает состав экспертной комиссии, возможно, вычеркнув или добавив часть экспертов к предложениям РГ. Проводится заключение договоров с экспертами об условиях их работы и ее оплаты.

7) Проведение *сбора* экспертной информации. Часто перед этим проводится набор и обучение интервьюеров - одной из групп, входящих в РГ.

8) Компьютерный *анализ экспертной информации* с помощью включенных в сценарий методов. Ему обычно предшествует введение информации в компьютеры.

9) При применении согласно сценарию экспертной процедуры из нескольких туров - *повторение* двух предыдущих этапов.

10) Итоговый анализ экспертных мнений, интерпретация полученных результатов аналитической группой РГ и *подготовка заключительного документа* ЭК для ЛПР.

11) Официальное *окончание* деятельности РГ, в том числе утверждение ЛПР заключительного документа ЭК, подготовка и утверждение научного и финансового отчетов РГ о проведении экспертного исследования, оплата труда экспертов и сотрудников РГ, официальное прекращение деятельности (роспуск) ЭК и РГ.

Разберем подробнее отдельные стадии экспертного исследования. Начнем с подбора экспертов: кадры решают все! Каковы эксперты - таково и качество заключения экспертной комиссии.

12.3. Подбор экспертов

Проблема подбора экспертов является одной из наиболее сложных в теории и практике экспертных исследований. Очевидно, в качестве экспертов необходимо использовать тех людей, чьи суждения наиболее помогут принятию адекватного решения. Но как выделить, найти, подобрать таких людей? Надо прямо сказать, что в настоящее время и в обозримом будущем не будет методов подбора экспертов, наверняка обеспечивающих успех экспертизы. Сейчас мы не будем возвращаться к обсуждению проблемы существования различных "партий" среди экспертов и обратим внимание на иные стороны процедур подбора экспертов.

В проблеме подбора экспертов можно выделить две составляющие - 1)*составление списка возможных экспертов и 2) выбор из них экспертной комиссии в соответствии с компетентностью кандидатов.*

Составление списка возможных экспертов облегчается тогда, когда рассматриваемый вид экспертизы проводится многократно. В таких ситуациях обычно ведется *реестр* возможных экспертов, например, в области государственной экологической экспертизы или судейства фигурного катания. Из обширного реестра можно выбирать по различным критериям или с помощью *датчика псевдослучайных чисел*.

Как быть, если экспертиза проводится впервые, а потому устоявшиеся списки возможных экспертов отсутствуют? Однако и в этом случае у каждого конкретного специалиста есть некоторое представление о том, что требуется от эксперта в подобной ситуации. Для формирования списка есть полезный метод "*снежного кома*", при котором от каждого специалиста, привлекаемого в качестве эксперта, получают несколько (например, пять) фамилий тех, кто может быть экспертом по рассматриваемой тематике. Очевидно, некоторые из этих фамилий встречались ранее и зафиксированы в списках РГ, а некоторые - новые. Каждого вновь появившегося опрашивают по той же схеме. Процесс расширения списка останавливается, когда список экспертов расширяется до нужных размеров или когда новые фамилии практически перестают встречаться. В результате получается достаточно обширный список возможных экспертов. Метод "снежного кома" имеет и недостатки. Число туров до остановки процесса наращивания "*снежного кома*" нельзя заранее предсказать, а потому нельзя предварительно установить продолжительность и стоимость этой работы. Кроме того, ясно, что если на первом этапе все эксперты были из одного "клана", придерживались в чем-то близких взглядов или занимались сходной деятельностью, то и метод "снежного кома" даст, скорее всего, прежде всего лиц из этого "клана". Мнения и аргументы других "кланов" будут упущены.

Вопрос об оценке компетентности экспертов не менее сложен. Успешность участия в предыдущих экспертизах - хороший критерий для деятельности дегустатора, врача, судьи в спортивных соревнованиях, т.е. таких экспертов, которые участвуют в длинных сериях однотипных экспертиз. Однако, увы, наиболее интересны и важны уникальные экспертизы больших проектов, не имеющих аналогов. Использование формальных показателей экспертов (должность, ученые степень и звание, стаж, число публикаций, награды, ...), очевидно, в современных условиях может носить лишь вспомогательный характер, хотя подобные показатели проще всего применять при формировании экспертной комиссии.

Часто предлагают использовать методы *самооценки* и *взаимооценки* компетентности экспертов. Обсудим их, начав с метода самооценки, при котором эксперт сам дает информацию о том, в каких областях он компетентен, а в каких - нет. С одной стороны, кто лучше может знать возможности эксперта, чем он сам? С другой стороны, при самооценке компетентности скорее оценивается степень самоуверенности эксперта, чем его реальная компетентность. Тем более, что само понятие *"компетентность"* строго не определено. Можно его уточнять, выделяя составляющие, но при этом усложняется предварительная часть деятельности экспертной комиссии. Достаточно часто эксперт преувеличивает свою реальную компетентность. Например, большинство людей считают, что они хорошо разбираются в политике, экономике, проблемах образования и воспитания, семьи и медицины. На самом деле экспертов (и даже знающих людей) в этих областях не так уж много, особено в сравнении с претензиями профанов. Бывают уклонения и в другую сторону, излишне критичное отношение к своим возможностям. При этом специалист вполне сознательно искусственно сужает зону своей компетенции. Так, научный работник может заявить, что он компетентен только в том, чему посвящены его последние публикации.

При использовании метода взаимооценки, помимо возможности проявления личностных и групповых симпатий и антипатий, играет роль малая осведомленность экспертов о возможностях друг друга. В современных условиях достаточно хорошее знакомство с работами и возможностями друг друга может быть лишь у специалистов, много лет (не менее 3-4) работающих совместно, в одной комнате, над одной темой. Именно про такие пары можно сказать, что они "*пуд соли вместе съели*". Однако привлечение таких пар специалистов не очень-то целесообразно, поскольку их взгляды из-за схожести жизненного пути слишком похожи друг на друга. Малая осведомленность экспертов о возможностях друг друга приводит к взаимооценкам на основе недостаточных, а иногда и не вполне достоверных сведений, или на базе ранее описанных формальных показателей.

Если процедура экспертного опроса предполагает непосредственное общение экспертов, необходимо учитывать еще ряд обстоятельств. Большое значение имеют их личностные (социально-психологические) качества. Так, один-единственный "*говорун*" может парализовать деятельность всей комиссии на совместном заседании. К срыву могут привести и неприязненные отношения членов комиссии, и сильно различающийся научный и должностной статус членов комиссии. В подобных случаях важно соблюдение регламента работы, разработанного РГ.

Необходимо подчеркнуть, что подбор экспертов в конечном счете - функция Рабочей группы, и никакие методики подбора не снимают с нее ответственности. Другими словами, именно на Рабочей группе лежит ответственность за компетентность экспертов, за их принципиальную способность решить поставленную задачу. Важным является требование к ЛПР об утверждении списка экспертов. При этом ЛПР может как добавить в комиссию отдельных экспертов, так и вычеркнуть некоторых из них - по собственным соображениям, с которыми членам РГ и ЭК знакомиться нет необходимости.

Существует ряд нормативных документов, регулирующих деятельность экспертных комиссий в тех или иных областях. Примером является Закон Российской Федерации "Об экологической экспертизе" от 23 ноября 1995 г., в котором регламентируется процедура экспертизы "намечаемой хозяйственной или иной деятельности" с целью выявления возможного вреда, который может нанести рассматриваемая деятельность окружающей природной среде.

12.4. О разработке регламента проведения сбора и анализа экспертных мнений

Как уже отмечалось, существует масса методов получения экспертных оценок. В одних с каждым экспертом работают отдельно, он даже не знает, кто еще является экспертом, а потому высказывает свое мнение независимо от авторитетов, "кланов" и отдельных коллег. В других экспертов собирают вместе для подготовки материалов для ЛПР, при этом эксперты обсуждают проблему друг с другом, принимают или отвергают аргументы друг друга, учатся друг у друга, и неверные или недостаточно обоснованные мнения отбрасываются. В одних методах число экспертов фиксировано и таково, чтобы статистические методы проверки согласованности мнений и затем (в случае достаточно хорошей согласованности мнений) их усреднения позволяли принимать обоснованные решения с точки зрения эконометрики. В других - число экспертов растет в процессе проведения экспертизы, например, при использовании метода "снежного кома" для формирования команды экспертов.

В настоящее время *не существует* общепринятой научно обоснованной классификации методов экспертных оценок и тем более - однозначных рекомендаций по их применению. Попытка силой (приняв какой-либо нормативно-правовой документ) утвердить одну из возможных точек зрения может принести лишь вред.

Однако для рассказа о многообразии экспертных оценок необходима какая-либо рабочая классификация методов. Одну из таких возможных классификаций мы даем ниже, перечисляя основания, по которым мы делим экспертные оценки.

Какова цель работы комиссии? Один из основных вопросов - что именно должна представить экспертная комиссия в результате своей работы - информацию для принятия решения ЛПР или проект самого решения? От ответа на этот методологический вопрос зависит организация работы экспертной комиссии, и он служит первым основанием для разбиения методов на группы с целью их классификации.

ЦЕЛЬ - СБОР ИНФОРМАЦИИ ДЛЯ ЛПР. Тогда Рабочая группа должна собрать возможно больше относящейся к делу информации, аргументов "за" и "против" определенных вариантов решений. Полезен следующий метод постепенного увеличения числа экспертов. Сначала первый эксперт приводит свои соображения по рассматриваемому вопросу. Составленный им материал передается второму эксперту, который добавляет свои аргументы. Накопленный материал поступает к следующему - третьему - эксперту... Процедура заканчивается, когда иссякает поток новых соображений.

Отметим, что эксперты в рассматриваемом методе только поставляют информацию, аргументы "за" и "против", но не вырабатывают согласованного проекта решения. Нет никакой необходимости стремиться к тому, чтобы экспертные мнения были согласованы между собой. Более того, наибольшую пользу приносят эксперты с мышлением, отклоняющимся от массового. Таких обычно называют диссидентами (т.е. инакомыслящими). Именно от них следует ожидать наиболее оригинальных аргументов.

ЦЕЛЬ - ПОДГОТОВКА ПРОЕКТА РЕШЕНИЯ ДЛЯ ЛПР. Эконометрические методы в экспертных оценках применяются обычно именно для решения задач, связанных с подготовкой проекта решения, основанного на итоговом коллективном мнении комиссии экспертов. При этом зачастую некритически принимают догмы согласованности и одномерности. Эти догмы "кочуют" из одной публикации в другую, поэтому целесообразно их обсудить.

ДОГМА СОГЛАСОВАННОСТИ. Часто без всяких оснований считается, что решение может быть принято лишь на основе согласованных мнений экспертов. Поэтому исключают из экспертной группы тех, чье мнение отличается от мнения большинства. (В лучшем случае им разрешает составить документ под названием: "Особое мнение".) При этом отсеиваются как неквалифицированные лица, попавшие в состав экспертной комиссии по недоразумению или по соображениям, не имеющим отношения к их компетентности и профессиональному уровню, так и наиболее оригинальные мыслители, глубже проникшие в проблему, чем большинство. Следовало бы выяснить их аргументы, предоставить им возможность для обоснования их точек зрения. Вместо этого их мнением пренебрегают.

Бывает и так, что эксперты делятся на две или более групп, имеющих единые *групповые* точки зрения. Так, хорошо известны примеры деления специалистов при оценке результатов научно-исследовательских работ (НИР) на две группы: "теоретиков", явно предпочитающих НИР, в которых получены теоретические результаты, и "практиков", выбирающих те НИР, которые позволяют получать непосредственные прикладные результаты (речь идет, например, об истории конкурсов НИР в академическом Институте проблем управления (автоматики и телемеханики)).

Иногда заявляют, что в случае обнаружения двух или нескольких групп экспертов (вместо одной согласованной во мнениях) опрос не достиг цели. Это не так! *Цель достигнута - установлено, что единого мнения нет.* Это весьма важно. И ЛПР при принятии решений должен это учитывать. Стремление обеспечить согласованность мнений экспертов любой целой может приводить к сознательному одностороннему подбору экспертов, игнорированию всех точек зрения, кроме одной, наиболее полюбившейся Рабочей группе (или даже "подсказанной" ЛПР).

Часто не учитывают еще одного чисто эконометрического обстоятельства. Поскольку число экспертов обычно не превышает 20-30, то формальная статистическая согласованность мнений экспертов (установленная с помощью тех или иных критериев проверки статистических гипотез) может сочетаться с реально имеющимся разделением экспертов на группы, что делает дальнейшие расчеты не имеющими отношения к действительности, а потому не имеющими практического смысла. Для примера обратимся к конкретным методам расчетов с помощью коэффициентов конкордации на основе коэффициентов ранговой корреляции Кендалла или Спирмена. Необходимо напомнить, что согласно эконометрической теории положительный результат проверки согласованности таким способом означает ни больше, ни меньше, как отклонение конкретной статистической гипотезы, а именно, гипотезы о независимости и равномерной распределенности мнений экспертов на множестве всех ранжировок. Таким образом, проверяется нулевая гипотеза, согласно которой ранжировки, описывающие мнения экспертов, являются независимыми случайными бинарными отношениями, равномерно распределенными на множестве всех ранжировок [1]. Отклонение этой нулевой гипотезы толкуется как принятие альтернативной гипотезы согласованности ответов экспертов. Другими словами, мы падаем жертвой заблуждений, вытекающих из различного толкования одних и тех же слов в не вполне связанных друг с другом научных дисциплинах: принятие статистической гипотезы согласованности в указанном математико-статистическом смысле вовсе не является обоснованием согласованности мнений экспертов в смысле практики экспертных оценок. (Именно ущербность рассматриваемых математико-статистических методов анализа ранжировок привела группу специалистов к разработке нового эконометрического аппарата для проверки согласованности - непараметрических методов, основанных на т.н. *люсианах* и входящих в современный раздел эконометрики - *статистику нечисловых данных*). Группы экспертов с близкими взглядами можно выделить эконометрическими методами кластер-анализа.

МНЕНИЯ ДИССИДЕНТОВ. С целью искусственно добиться согласованности стараются уменьшить влияние мнений экспертов-*диссидентов*, т.е. инакомыслящих по сравнению с большинством. *Жесткий* способ борьбы с диссидентами состоит в игнорировании их мнений, т.е. фактически в их исключении из состава экспертной комиссии. Отбраковка экспертов, как и отбраковка резко выделяющихся результатов наблюдений (выбросов), приводит к процедурам, имеющим плохие или неизвестные статистические свойства. Так, известна *крайняя неустойчивость* классических методов отбраковки выбросов по отношению к отклонениям от предпосылок модели.

Мягкий способ борьбы с диссидентами состоит в применении *робастных (устойчивых) статистических процедур*. Простейший пример: если ответ эксперта - действительное число, то резко выделяющееся мнение диссидента сильно влияет на среднее арифметическое ответов экспертов и не влияет на их медиану. Поэтому разумно в качестве согласованного мнения рассматривать медиану. Однако при этом игнорируются (не достигают ЛПР) аргументы диссидентов.

В любом из двух способов борьбы с диссидентами ЛПР лишается информации, идущей от диссидентов, а потому может принять необоснованное решение, которое впоследствии приведет к отрицательным последствиям. С другой стороны, представление ЛПР всего набора мнений снимает часть ответственности и труда по подготовке окончательного решения с комиссии экспертов и рабочей группы по проведению экспертного опроса и перекладывает эти ответственность и труд на плечи ЛПР.

ДОГМА ОДНОМЕРНОСТИ. Среди менеджеров и инженеров распространен довольно примитивный подход так называемой "квалиметрии", согласно которому объект экспертизы всегда можно оценить *одним числом*. Странная идея! *Оценивать человека одним числом приходило в голову лишь на невольничьих рынках*. Вряд ли даже самые рьяные квалиметристы рассматривают книгу или картину как эквивалент числа - ее "рыночной стоимости".

Вместе с тем нельзя полностью отрицать саму идею поиска обобщенных показателей качества, технического уровня и аналогичных им. Так, каждый объект можно оценивать по многим показателям качества. Например, легковой автомобиль можно оценивать по таким показателям:

расход бензина на 100 км пути (в среднем);

надежность (средняя стоимость ремонта за год);

экологическая безопасность, оцениваемая по содержанию вредных веществ в выхлопных газах;

маневренность;

быстрота набора скорости 100 км/час после начала движения; максимальная достигаемая скорость;

длительность сохранения в салоне положительной температуры при низкой наружной температуре (-50 градусов по Цельсию) и выключенном двигателе;

дизайн (привлекательность и "модность" внешнего вида и отделки салона);

представительность;

вес;

срок службы;

эксплуатационные расходы (за год);

цена;

приведенная к сопоставимым ценам стоимость 1 км пробега, и т.д.

Можно ли свести оценки по этим показателям вместе? Определяющей является конкретная ситуация, для которой выбирается автомашина. Максимально достигаемая скорость важна для гонщика, но, как нам представляется, не имеет большого практического значения для водителя рядовой частной машины, особенно в большом городе с суровым ограничением на максимальную скорость. Для такого водителя важнее расход бензина, маневренность и надежность. Для машин различных служб государственного управления, видимо, надежность важнее, чем для частника, а расход бензина - наоборот. Представительность важна для высших менеджеров и чиновников, занимающих высокие посты. Для районов Крайнего Севера важна теплоизоляция салона, а для южных районов - нет. И т.д.

Таким образом, важна конкретная (узкая) постановка задачи перед экспертами. Но такой постановки зачастую нет. А тогда "игры" по разработке обобщенного показателя качества - например, в виде линейной функции от перечисленных переменных - могут не дать объективных выводов. Альтернативой единственному обобщенному показателю является математический аппарат типа *многокритериальной оптимизации* - множества Парето и т.д.

В некоторых случаях все-таки можно глобально сравнить объекты - например, с помощью тех же экспертов получить упорядочение рассматриваемых объектов - изделий или проектов. Тогда можно ПОДОБРАТЬ коэффициенты при отдельных показателях так, чтобы *упорядочение с помощью линейной функции возможно точнее соответствовало глобальному упорядочению* (например, найти эти коэффициенты методом наименьших квадратов). Наоборот, в подобных случаях НЕ СЛЕДУЕТ оценивать указанные коэффициенты непосредственно с помощью экспертов. Эта простая идея до сих пор не стала очевидной для отдельных составителей методик по проведению экспертных опросов и анализу их результатов. Они упорно стараются заставить экспертов делать то, что они выполнить *не в состоянии* - указывать веса, с которыми отдельные показатели качества должны входить в итоговый обобщенный показатель.

Эксперты обычно могут сравнить объекты или проекты в целом, но не могут вычленить вклад отдельных факторов. *Раз организаторы опроса спрашивают, эксперты отвечают,* но эти ответы не несут в себе надежной информации о реальности...

Отметим, что есть экспертные процедуры, в которых веса отдельных факторов вычисляются в результате тщательного анализа иерархической системы показателей. Для таких процедур приведенные выше критические замечания по поводу экспертного определения весов факторов не имеют силы.

ВТОРОЕ ОСНОВАНИЕ КЛАССИФИКАЦИИ ЭКСПЕРТНЫХ ПРОЦЕДУР - ЧИСЛО ТУРОВ. Экспертизы могут включать один тур, некоторое фиксированное число туров (два, три,…) или неопределенное число туров. Чем больше туров, тем более тщательным является анализ ситуации, поскольку эксперты при этом обычно много раз возвращаются к рассмотрению предмета экспертизы. Но одновременно увеличивается общее время на экспертизу и возрастает ее стоимость. Можно уменьшить расходы, вводя в экспертизу не всех экспертов сразу, а постепенно. Так, например, если цель состоит в сборе аргументов "за" и "против", то первоначальный перечень аргументов может быть составлен одним экспертом. Второй добавит к нему свои аргументы. Суммарный материал поступит к первому и третьему, которые внесут свои аргументы и контраргументы. И так далее - добавляется по одному эксперту на каждый новый тур.

Наибольшие сложности вызывают процедуры с заранее неопределенным числом туров, например, "снежный ком". Часто задают максимально возможное число туров, и тогда неопределенность сводится к тому, придется ли проводить это максимальное число туров или удастся ограничиться меньшим числом.

ТРЕТЬЕ ОСНОВАНИЕ КЛАССИФИКАЦИИ ЭКСПЕРТНЫХ ПРОЦЕДУР - ОРГАНИЗАЦИЯ ОБЩЕНИЯ ЭКСПЕРТОВ. Рассмотрим достоинства и недостатки каждого из элементов следующей шкалы: отсутствие общения - заочное анонимное общение - заочное общение без анонимности - очное общение с ограничениями - очное общение без ограничений.

При отсутствии общения эксперт высказывает свое мнение, ничего не зная о других экспертах и об их мнениях. Он полностью независим, что и хорошо, и плохо.

Хорошо - потому что он полностью независим, защищен от любого давления. Плохо - он не знает соображений других экспертов, а потому опирается лишь на собственную информационную базу. Обычно такая ситуация соответствует однотуровой экспертизе.

Заочное анонимное общение, например, как в методе Дельфи, означает, что эксперт знакомится с мнениями и аргументами других экспертов, но не знает, кто именно высказал то или иное положение. Следовательно, в экспертизе должно быть предусмотрено хотя бы два тура, чтобы эксперт смог скорректировать свое мнение, познакомившись с мнениями других.

Заочное общение без анонимности соответствует, например, общению по Интернету. Экспертные опросы на основе информационных технологий -весьма перспективное направление развития в области организации экспертиз.

Все варианты заочной экспертизы хороши тем, что нет необходимости собирать экспертов вместе, следовательно, находить для этого удобное для всех экспертов время и место.

При очных экспертизах эксперты говорят, а не пишут, как при заочных, и потому успевают за то же время сказать существенно больше. *Очная экспертиза с ограничениями* весьма распространена. Это - собрание, идущее по фиксированному регламенту. Примером является военный совет в императорской русской армии, когда эксперты (офицеры и генералы) высказывались в порядке от младшего (по чину и должности) к старшему. Подробно такой совет описан в повести А.С. Пушкина "Капитанская дочка". Другой пример - "мозговой штурм", при котором запрещается критиковать чужие высказывания.

Наконец, *очная экспертиза без ограничений* - это свободная дискуссия. Все очные экспертизы имеют недостатки, связанные с возможностями отрицательного влияния на их проведение социально-психологических свойств и клановых (партийных) пристрастий участников, а также неравенства их профессионального, должностного, научного статусов. Представьте себе, что соберутся вместе 5 лейтенантов и 3 генерала. Независимо от того, какая информация имеется у того или иного участника встречи, ход ее предсказать нетрудно: генералы будут говорить, а лейтенанты - помалкивать. Хотя опыт генералов несравним с лейтенантским, но в процессе недавнего обучения лейтенанты познакомились с последними научными достижениями, которые почти наверняка прошли мимо внимания генералов. Разумеется, этот пример можно обсудить не для совещания военных, а для собрания управленцев (менеджеров), врачей или преподавателей.

Необходимо обратить внимание на психологические особенности экспертов. Один начнет громко говорить на разные темы, как только ему это удастся, и словесный понос может продолжаться часами. От попыток прервать его и сказать свое такой деятель легко отбивается, в частности, за счет высокой скорости произнесения слов. Поведение другого эксперта противоположно. Он предпочитает помалкивать, пока ему не предоставлено слово, не тратить сил на попытки вклиниться в словесный понос "говоруна". Из сказанного ясно, что *очная экспертиза без ограничений,* т.е. свободная дискуссия - это недостижимый идеал, реально роль и возможности председателя заседания (зафиксированные в регламенте работы ЭК) должны быть достаточны велики.

КОМБИНАЦИЯ РАЗЛИЧНЫХ ВИДОВ ЭКСПЕРТИЗЫ. Реальные экспертизы часто представляют собой комбинации различных описанных выше типов экспертиз. В качестве примера рассмотрим систему экспертиз при подготовке и защите студентом дипломного проекта. Сначала идет многотуровая очная экспертиза, проводимая научным руководителем и консультантами, в результате учета результатов этой экспертизы студент подготавливает проект к защите. Затем два эксперта работают заочно - это автор отзыва сторонней организации и заведующий кафедрой, допускающий работу к защите. Обратите внимание на различие задач этих экспертов и объемов выполняемой ими работы - один пишет подробный отзыв и дает оценку проекту, второй росписью на титульном листе проекта разрешает его защиту. Наконец - очная экспертиза без ограничений (для членов государственной аттестационной комиссии). Дипломный проект оценивается коллегиально, по большинству голосов, при этом лишь один из экспертов (научный руководитель) знает работу подробно, а остальные - в основном лишь по докладу студента. Таким образом, имеем сочетание многотуровой и однотуровой, заочных и очных экспертиз. Подобные сочетания характерны для многих реально проводящихся экспертиз.

12.5. Методы средних баллов

В настоящее время распространены экспертные, маркетинговые, квалиметрические, социологические и иные опросы, в которых опрашиваемых просят выставить баллы объектам, изделиям, технологическим процессам, предприятиям, проектам, заявкам на выполнение научно-исследовательских работ, идеям, проблемам, программам, политикам и т.п., а затем рассчитывают средние баллы и рассматривают их как интегральные оценки, выставленные коллективом опрошенных.

Какими формулами пользоваться для вычисления средних величин? Ведь разных видов средних величин, как мы знаем, очень много (см. главу 3). Обычно в старых или устаревших литературных источниках рекомендуют применять среднее арифметическое. Однако эта устоявшаяся рекомендация противоречит теории измерений. Однако уже более 25 лет известно, что *такой способ некорректен*, поскольку баллы обычно измерены в порядковой шкале. Обоснованным является использование медиан в качестве средних баллов. Это вытекает из теорем эконометрики, приведенных в главе 3.

Однако полностью *игнорировать средние арифметические нерационально из-за их привычности и распространенности*. Поэтому **целесообразно использовать одновременно оба метода - и метод средних арифметических рангов (баллов), и методов медианных рангов.** Эта рекомендация находится в согласии с концепцией устойчивости, согласно которой следует использовать различные методы для обработки одних и тех же данных с целью выделить выводы, получаемые одновременно при всех методах. Такие выводы, видимо, соответствуют реальной действительности, в то время как заключения, меняющиеся от метода к методу, зависят от субъективизма исследователя, выбирающего метод обработки исходных данных, например, экспертных оценок.

Пример сравнения восьми проектов. Рассмотрим конкретный пример применения только что сформулированного подхода.

По заданию руководства фирмы анализировались восемь проектов, предлагаемых для включения в план стратегического развития фирмы. Они были обозначены следующим образом: Д, Л, М-К, Б, Г-Б, Сол, Стеф, К (по фамилиям менеджеров, предложивших их для рассмотрения). Все проекты были направлены 12 экспертам, назначенным Советом директоров фирмы. В приведенной ниже табл.1 приведены ранги восьми проектов, присвоенные им каждым из 12 экспертов в соответствии с представлением экспертов о целесообразности включения проекта в стратегический план фирмы. При этом эксперт присваивает ранг 1 самому лучшему проекту, который обязательно надо реализовать. Ранг 2 получает от эксперта второй по привлекательности проект,..., наконец, ранг 8 - наиболее сомнительный проект, который реализовывать стоит лишь в последнюю очередь).

Анализируя результаты работы экспертов (т.е. упомянутую таблицу), члены Правления фирмы были вынуждены констатировать, что полного согласия между экспертами нет, а потому данные, приведенные в табл.1, следует подвергнуть более тщательному математическому анализу.

Метод средних арифметических рангов. Сначала был применен метод средних арифметических рангов. Для этого прежде всего была подсчитана сумма рангов, присвоенных проектам (см. табл.1). Затем эта сумма была разделена на число экспертов, в результате рассчитан средний арифметический ранг (именно эта операция дала название методу). По средним рангам строится итоговая ранжировка (в другой терминологии - упорядочение), исходя из принципа - чем меньше средний ранг, чем лучше проект.

Табл. 1. Ранги 8 проектов по степени привлекательности для включения в план стратегического развития фирмы

№ эксперта	Д	Л	М-К	Б	Г-Б	Сол	Стеф	К
1	5	3	1	2	8	4	6	7
2	5	4	3	1	8	2	6	7
3	1	7	5	4	8	2	3	6
4	6	4	2,5	2,5	8	1	7	5
5	8	2	4	6	3	5	1	7
6	5	6	4	3	2	1	7	8
7	6	1	2	3	5	4	8	7
8	5	1	3	2	7	4	6	8
9	6	1	3	2	5	4	7	8
10	5	3	2	1	8	4	6	7
11	7	1	3	2	6	4	5	8
12	1	6	5	3	8	4	2	7

Примечание. Эксперт № 4 считает, что проекты М-К и Б равноценны, но уступают лишь одному проекту - проекту Сол. Поэтому проекты М-К и Б должны были бы стоять на втором и третьем местах и получить баллы 2 и 3. Поскольку они равноценны, то получают средний балл (2+3)/ 2 = 5/ 2 = 2,5.

Наименьший средний ранг, равный 2,625, у проекта Б, - следовательно, в итоговой ранжировке он получает ранг 1. Следующая по величине сумма, равная 3,125, у проекта М-К. И он получает итоговый ранг 2. Проекты Л и Сол имеют одинаковые суммы (равные 3,25), значит, с точки зрения экспертов они равноценны (при рассматриваемом способе сведения вместе мнений экспертов с целью получения итоговой ранжировки), а потому они должны бы стоять на 3 и 4 местах и получают средний балл (3+4) /2 = 3,5. Дальнейшие результаты приведены в табл. 2 ниже.

Итак, ранжировка по суммам рангов (или, что то же самое, по средним арифметическим рангам) имеет вид:

$$Б < М\text{-}К < \{Л, Сол\} < Д < Стеф < Г\text{-}Б < К . \qquad (1)$$

Здесь запись типа "А<Б" означает, что проект А предшествует проекту Б (т.е. проект А лучше проекта Б). Поскольку проекты Л и Сол получили одинаковую сумму баллов, то по рассматриваемому методу они эквивалентны, а потому объединены в группу - класс эквивалентности (в фигурных скобках). В терминологии математической статистики ранжировка (1) имеет одну связь.

Метод медиан рангов. Значит, наука сказала свое слово, итог расчетов - ранжировка (1), и на ее основе предстоит принимать решение? Но тут наиболее знакомый с современной эконометрикой член Совета директоров вспомнил то, о чем говорилось в главе 3, посвященной теории измерений. Он вспомнил, что ответы экспертов измерены в порядковой шкале, а потому для них неправомерно проводить усреднение методом средних арифметических. Надо использовать метод медиан.

Что это значит? Напомним, что надо взять ответы экспертов, соответствующие одному из проектов, например, проекту Д. Это ранги 5, 5, 1, 6, 8, 5, 6, 5, 6, 5, 7, 1. Затем их надо расположить в порядке неубывания (проще было бы сказать - "в порядке возрастания", но поскольку некоторые ответы совпадают, то приходится использовать непривычный термин "неубывание"). Получим последовательность: 1, 1, 5, 5, 5, 5, 5, 6, 6, 6, 7, 8. На центральных местах - шестом и седьмом - стоят 5 и 5. Следовательно, медиана равна 5.

Табл. 2. Результаты расчетов по методу средних арифметических и методу медиан для данных, приведенных в табл. 1.

	Д	Л	М-К	Б	Г-Б	Сол	Стеф	К
Сумма рангов	60	39	37,5	31.5	76	39	64	85
Среднее арифметическое рангов	5	3,25	3,125	2,625	6,333	3,25	5,333	7,083
Итоговый ранг по среднему арифметическому	5	3,5	2	1	7	3,5	6	8
Медианы рангов	5	3	3	2,25	7,5	4	6	7
Итоговый ранг по медианам	5	2,5	2,5	1	8	4	6	7

Медианы совокупностей из 12 рангов, соответствующих определенным проектам, приведены в предпоследней строке табл.2. (При этом медианы вычислены по обычным правилам статистики - как среднее арифметическое центральных членов вариационного ряда.) Итоговое упорядочение по методу медиан приведено в последней строке табл.2. Ранжировка (т.е. упорядочение - итоговое мнение комиссии экспертов) по медианам имеет вид:

$$Б < \{М\text{-}К, Л\} < Сол < Д < Стеф < К < Г\text{-}Б \ . \qquad (2)$$

Поскольку проекты Л и М-К имеют одинаковые медианы баллов, то по рассматриваемому методу ранжирования они эквивалентны, а потому объединены в группу (кластер), т.е. с точки зрения математической статистики ранжировка (2) имеет одну связь.

Сравнение ранжировок по методу средних арифметических и методу медиан. Сравнение ранжировок (1) и (2) показывает их близость (похожесть). Можно принять, что проекты М-К, Л, Сол упорядочены как М-К < Л < Сол, но из-за погрешностей экспертных оценок в одном методе признаны равноценными проекты Л и Сол (ранжировка (1)), а в другом - проекты М-К и Л (ранжировка (2)). Существенным является только расхождение, касающееся упорядочения проектов К и Г-Б: в ранжировке (1) Г-Б<К, а в ранжировке (2), наоборот, К<Г-Б. Однако эти проекты - наименее привлекательные из восьми рассматриваемых, и при выборе наиболее привлекательных проектов для дальнейшего обсуждения и использования на указанное расхождение можно не обращать внимания.

Рассмотренный пример демонстрирует сходство и различие ранжировок, полученных по методу средних арифметических рангов и по методу медиан, а также пользу от их совместного применения. Однако нельзя не отметить, что только что проведенное сравнение ранжировок (1) и (2) осуществлено не вполне строго. Ясно, что в эконометрическом инструментарии специалиста по проведению экспертных исследований должен быть алгоритм согласования ранжировок, полученных различными методами.

12.6. Метод согласования кластеризованных ранжировок

Рассматриваемая здесь проблема состоит в выделении общего нестрогого порядка из набора кластеризованных ранжировок (на статистическом языке - ранжировок со связями). Этот набор может отражать мнения нескольких экспертов или быть получен при обработке мнений экспертов различными методами. *Рассматривается метод согласования кластеризованных ранжировок, позволяющий «загнать» противоречия внутрь специальным образом построенных кластеров (групп), в то время как упорядочение кластеров соответствует всем исходным упорядочениям.*

В различных прикладных областях возникает необходимость анализа нескольких кластеризованных ранжировок объектов. К таким областям относятся прежде всего инженерный бизнес, менеджмент, экономика, социология, прогнозирование, экология, научные и технические исследования и т.д., особенно те их разделы, что связаны с экспертными оценками. В качестве объектов могут выступать образцы продукции, технологии, математические модели, проекты, кандидаты на должность и др. Кластеризованные ранжировки могут быть получены как с помощью экспертов, так и объективным путем, например, при сопоставлении математических моделей с экспериментальными данными с помощью того или иного критерия качества. Первоначально описанный ниже метод был разработан в связи с проблемами химической безопасности биосферы и экологического страхования.

Повторим более подробно постановку проблемы. В настоящем пункте учебного пособия рассматривается метод построения кластеризованной ранжировки, согласованной (в раскрытом ниже смысле) со всеми рассматриваемыми кластеризованными ранжировками. При этом противоречия между отдельными исходными ранжировками оказываются заключенными внутри кластеров согласованной ранжировки. В результате упорядоченность кластеров отражает общее мнение экспертов, точнее, то общее, что содержится в исходных ранжировках. В кластеры заключены объекты, по поводу которых некоторые из исходных ранжировок *противоречат* друг другу. Для их упорядочения необходимо провести новые исследования. Эти исследования могут быть как формально-математическими (например, вычисление медианы Кемени, упорядочения по средним рангам или по медианам и т.п.), так и требовать привлечения новой информации из соответствующей прикладной области, возможно, проведения дополнительных научных или прикладных работ.

Введем необходимые понятия, затем сформулируем алгоритм согласования кластеризованных ранжировок в общем виде и рассмотрим его свойства.

Пусть имеется конечное число объектов, которые мы для простоты изложения будем изображать натуральными числами 1,2,3,...,k и называть «носителем». *Под кластеризованной ранжировкой, определенной на заданном носителе, понимаем следующую математическую конструкцию*. Пусть объекты разбиты на группы, которые будем называть кластерами. В кластере может быть и один элемент. Входящие в один кластер объекты будем заключать в фигурные скобки. Например, объекты 1,2,3,...,10 могут быть разбиты на 7 кластеров: {1}, {2,3}, {4}, {5,6,7}, {8}, {9}, {10}. В этом разбиении один кластер {5,6,7} содержит три элемента, другой - {2,3} - два, остальные пять - по одному элементу. Кластеры не имеют общих элементов, а объединение их (как множеств) есть все рассматриваемое множество объектов.

Вторая составляющая кластеризованной ранжировки - это строгий линейный порядок между кластерами. Задано, какой из них первый, какой второй, и т.д. Будем изображать упорядоченность с помощью знака < . При этом кластеры, состоящие из одного элемента, будем для простоты изображать без фигурных скобок. Тогда кластеризованную ранжировку (одну из возможных) на основе введенных выше кластеров можно изобразить так:

$$A = [\ 1 < \{2,3\} < 4 < \{5,6,7\} < 8 < 9 < 10\ ]\ .$$

Конкретные кластеризованные ранжировки будем заключать в квадратные скобки. Если для простоты речи термин "кластер" применять только к кластеру не менее чем из 2-х элементов, то можно сказать, что в кластеризованную ранжировку А входят два кластера {2,3} и {5,6,7} и 5 отдельных элементов.

Введенная описанным образом кластеризованная ранжировка является **бинарным отношением** на множестве {1,2,3,...,10}. Его структура такова. Задано *отношение эквивалентности* с 7-ю классами эквивалентности, а именно, {2,3}, {5,6,7}, а остальные состоят из оставшихся 5 отдельных элементов. Затем введен *строгий линейный порядок* между классами эквивалентности.

Введенный математический объект известен в литературе как *"ранжировка со связями"* (М. Холлендер, Д.Вулф), *"упорядочение"* (Дж. Кемени, Дж. Снелл), *"квазисерия"* (Б.Г.Миркин), *"совершенный квазипорядок"* (Ю.А.Шрейдер [2, с.127, 130]). Учитывая разнобой в терминологии, мы сочли полезным ввести собственный термин *"кластеризованная ранжировка",* поскольку в нем явным образом названы основные элементы изучаемого математического объекта - кластеры, рассматриваемые на этапе согласования ранжировок как классы эквивалентности, и ранжировка - строгий совершенный порядок между ними (в терминологии Ю.А. Шрейдера [2, гл.IV]).

Следующее важное понятие - *противоречивость*. Оно определяется для четверки - две кластеризованные ранжировки на одном и том же носителе и два различных объекта - элементы того же носителя. При этом два элемента из одного кластера будем связывать символом равенства = , как эквивалентные .

Определение 1. *Пусть A и B - две кластеризованные ранжировки. Пару объектов (a,b) назовем «противоречивой» относительно A и B, если эти два элемента по-разному упорядочены в A и B, т.е. a < b в A и a > b в B (первый вариант противоречивости) либо a >b в A и a < b в B (второй вариант противоречивости).*

Отметим, что в соответствии с этим определением пара объектов (a,b), эквивалентная хотя бы в одной кластеризованной ранжировке, не может быть противоречивой: эквивалентность a = b не образует "противоречия" ни с a < b, ни с a > b.

Определение 2. *Совокупность противоречивых пар объектов для двух кластеризованных ранжировок A и B назовем «ядром противоречий» и обозначим S(A,B).*

В качестве примера рассмотрим две кластеризованные ранжировки

$$B = [\{1,2\} < \{3,4,5\} < 6 < 7 < 9 < \{8, 10\}],$$
$$C = [3 < \{1, 4\} < 2 < 6 < \{5, 7, 8\} < \{9, 10\}].$$

Для трех кластеризованных ранжировок A, B и C, определенных на одном и том же носителе {1, 2, 3,..., 10}, имеем

$$S(A,B) = [\ (8, 9)],\ S(A,C) = [\ (1, 3), (2,4)\ ],$$
$$S(B,C) = [\ (1, 3), (2, 3), (2, 4), (5, 6), (8,9)\ ].$$

Как при ручном, так и при программном нахождении ядра можно в поисках противоречивых пар просматривать пары (1,2), (1,3), (1.,4), ..., (1, k), затем (2,3), (2,4), ..., (2, k), потом (3,4), ..., (3, k), и т.д., вплоть до (k-1, k).

Пользуясь понятиями дискретной математики, «ядро противоречий» можно изобразить *графом* с вершинами в точках носителя. При этом *противоречивые пары задают ребра этого графа.* Граф для S(A,B) имеет только одно ребро (следовательно, у него одна связная компонента более чем из одной точки), для S(A,C) - 2 ребра (у этого графа две связные компоненты более чем из одной точки), для S(B,C) - 5 ребер (здесь три связные компоненты более чем из одной точки, а именно, {1, 2 , 3, 4}, {5, 6} и {8, 9}).

Каждую кластеризованную ранжировку, как и любое бинарное отношение, можно задать матрицей $\| x(a, b) \|$ из 0 и 1 порядка k x k. При этом x(a, b) = 1 тогда и только тогда, когда a < b либо a = b. В первом случае x(b, a) = 0, а во втором x(b, a) = 1. При этом хотя бы одно из чисел x(a, b) и x(b, a) равно 1. Из определения противоречивости пары (a, b) вытекает, что для нахождения всех таких пар достаточно поэлементно перемножить две матрицы $\|x(a,b)\|$ и $\|y(a, b)\|$, соответствующие двум кластеризованным ранжировкам, и отобрать те и только те пары, для которых $x(a,b)y(a,b)=x(b,a)y(b,a)=0$.

Кроме ядер противоречий, представляют интерес пары объектов, эквивалентных во всех исходных кластеризованных ранжировках.

Определение 3. *Ядром всеобщей эквивалентности называется совокупность пар объектов, в которых оба объекта эквивалентны во всех исходных кластеризованных ранжировках.*

Рассматриваемый алгоритм согласования некоторого числа кластеризованных ранжировок состоят из трех этапов.

На первом *выделяются противоречивые пары* объектов во всех парах кластеризованных ранжировок и формируются (попарные) ядра противоречий.

На втором формируются кластеры итоговой кластеризованной ранжировки (т.е. классы эквивалентности - *связные компоненты графа*, соответствующего объединению попарных ядер противоречий и ядра всеобщей эквивалентности).

На третьем этапе эти *кластеры (классы эквивалентности) упорядочиваются*. Для установления порядка между кластерами произвольно выбирается один объект из первого кластера и второй - из второго, порядок между кластерами устанавливается такой же, какой имеет быть между выбранными объектами в любой из рассматриваемых кластеризованных ранжировок. Отметим, что в некоторых из исходных кластеризованных ранжировок выбранные объекты могут быть эквивалентны (т.е. находиться в одном кластере).). В таком случае надо рассмотреть упорядоченность этих объектов в какой-либо другой из исходных кластеризованных ранжировок.   Если же они эквивалентны во всех исходных ранжировках, то входят в ядро всеобщей эквивалентности и будут эквивалентны и в итоговой кластеризованной ранжировке, что обеспечивается выполнением процедур второго этапа.

Корректность подобного упорядочивания, т.е. его независимость от выбора той или иной пары объектов, вытекает из соответствующих теорем, доказанных в статье [3].

Определение 4. *Результат применения алгоритма согласования к совокупности исходных кластеризованных ранжировок называется кластеризованной ранжировкой, согласованной с исходными (в другой формулировке - согласующей исходные ранжировки).*

Результат согласования кластеризованных ранжировок A, B, C,... обозначим f(A, B, C,...). Ядра противоречий выписаны выше. Ядро всеобщей эквивалентности возникает лишь при рассмотрении ранжировок A и C. Оно состоит из пары (5,7), поскольку объекты 5 и 7 (и только они) эквивалентны и в A, и в C. Тогда

$$f(A, B) = [1<2<3<4<5<6<7<\{8, 9\}<10],$$
$$f(A, C) = [\{1,3\}<\{2, 4\}<6<\{5,7\}<8<9<10],$$
$$f(B, C) = [\{1,2,3,4\}<\{5,6\}<7<\{8,9\}<10],$$
$$f(A, B, C) = f(B, C) = [\{1,2,3,4\} <\{5,6\}<7<\{8, 9\}<10].$$

В случае f(A, B) дополнительного изучения с целью упорядочения требуют только объекты 8 и 9. В случае f(B, C) объекты 1,2,3,4 объединились в один кластер, т.е. кластеризованные ранжировки оказались настолько противоречивыми, что процедура согласования не позволила провести достаточно полную декомпозицию задачи нахождения итогового мнения экспертов.

Рассмотрим некоторые свойства алгоритмов согласования.

1. Пусть D = f(A, B, C,...). Если $a<b$ в согласующей кластеризованной ранжировке D, то $a<b$ или $a=b$ в каждой из исходных ранжировок A, B, C, ...

2. Построение согласующих кластеризованных ранжировок может осуществляться поэтапно. В частности, f(A, B, C) = f(f(A, B), f(A, C), f(B, C)) . Ясно, что *ядро противоречий для набора кластеризованных ранжировок является объединением таких ядер для всех пар рассматриваемых ранжировок, а ядро всеобщей эквивалентности - пересечением таких ядер для всех пар рассматриваемых ранжировок*.

3. Построение согласующих кластеризованных ранжировок нацелено на выделение общего упорядочения в исходных кластеризованных ранжировках. Однако при этом некоторые общие свойства исходных кластеризованных ранжировок могут теряться. Так, при согласовании ранжировок B и C, рассмотренных выше, противоречия в упорядочении элементов 1 и 2 не было - в ранжировке B эти объекты входили в один кластер, т.е. 1 = 2, в то время как $1<2$ в кластеризованной ранжировке C. Значит, при их отдельном рассмотрении можно принять упорядочение $1 < 2$. Однако в f(B,C) они попали в один кластер, т.е. возможность их упорядочения исчезла. Это связано с поведением объекта 3, который "перескочил" в C на первое место и "увлек с собой в противоречие" пару (1, 2),

образовав противоречивые пары и с 1, и с 2. Другими словами, связная компонента графа, соответствующего ядру противоречий, сама по себе не всегда является полным графом. Недостающие ребра при этом соответствуют парам типа (1, 2), которые сами по себе не являются противоречивыми, но "увлекаются в противоречие" другими парами.

4. Необходимость согласования кластеризованных ранжировок возникает, в частности, при разработке методики применения экспертных оценок в задачах экологического страхования и химической безопасности биосферы. Поясним, как возникает эта необходимость. Как уже говорилось, популярным является метод упорядочения по средним рангам, в котором итоговая ранжировка строится на основе средних арифметических рангов, выставленных отдельными экспертами. Однако из теории измерений известно (см. выше главу 3), что более обоснованным является использование не средних арифметических, а медиан. Вместе с тем метод средних рангов весьма известен и широко применяется, так что просто отбросить его нецелесообразно. Поэтому было принято решение об одновременном применении обеих методов. Реализация этого решения потребовала разработки методики согласования двух указанных кластеризованных ранжировок.

5. Область применения рассматриваемого метода не ограничивается экспертными оценками. Он может быть использован, например, для построения банка знаний с целью использования в задачах экологического страхования требовалось разработать методику сравнения эконометрических, экономико-математических моделей и математических моделей в смежных областях. В частности, для расчета экономического ущерба от аварий использовались математические модели процесса испарения жидкости. Как сравнивать качество таких моделей? Имелись данные экспериментов и результаты расчетов по 8 математическим моделям. Сравнивать модели можно по различным критериям качества. Например, по сумме модулей относительных отклонений расчетных и экспериментальных значений. Можно и по другому - в каждой экспериментальной точке упорядочить модели по качеству, а потом получать единую оценку методами средних рангов и медиан. Использовались и иные методы. Затем применялись методы согласования полученных кластеризованных ранжировок. В результате оказалось возможным упорядочить модели по качеству и использовать это упорядочение при разработке банка математических моделей, используемого в задачах химической безопасности биосферы.

6. Рассматриваемый метод согласования кластеризованных ранжировок построен в соответствии с *методологией теории устойчивости*, согласно которой результат обработки данных, инвариантный относительно метода обработки, соответствует реальности, а результат расчетов, зависящий от метода обработки, отражает субъективизм исследователя, а не объективные соотношения.

12.7. Математические методы анализа экспертных оценок

При анализе мнений экспертов можно применять самые разнообразные статистические методы, описывать их - значит описывать всю прикладную статистику. Тем не менее можно выделить основные широко используемые в настоящее время методы математической обработки экспертных оценок - это проверка согласованности мнений экспертов (или классификация экспертов, если нет согласованности) и усреднение мнений экспертов внутри согласованной группы.

Поскольку ответы экспертов во многих процедурах экспертного опроса - не числа, а такие объекты нечисловой природы, как градации качественных признаков, ранжировки, разбиения, результаты парных сравнений, нечеткие предпочтения и т.д., то для их анализа оказываются полезными методы статистики объектов нечисловой природы.

Почему ответы экспертов часто носят нечисловой характер? Наиболее общий ответ состоит в том, что люди не мыслят числами. В мышлении человека используются образы, слова, но не числа. Поэтому требовать от эксперта ответ в форме чисел - значит насиловать его разум. Даже в экономике предприниматели, принимая решения, лишь частично опираются на численные расчеты. Это видно из условного (т.е. определяемого произвольно принятыми соглашениями, обычно оформленными в виде инструкций) характера балансовой прибыли, амортизационных отчислений и других экономических показателей. Поэтому фраза типа "фирма стремится к максимизации прибыли" не может иметь строго определенного смысла. Достаточно спросить: "Максимизация прибыли - за какой период?" И сразу станет ясно, что степень оптимальности принимаемых решений зависит от горизонта планирования (на экономико-математическом уровне этот сюжет рассмотрен в монографии [4]).

Эксперт может сравнить два объекта, сказать, какой из двух лучше (метод парных сравнений), дать им оценки типа "хороший", "приемлемый", "плохой", упорядочить несколько объектов по привлекательности, но обычно не может ответить, во сколько раз или на сколько один объект лучше другого. Другими словами, ответы эксперта обычно измерены в порядковой шкале, или являются ранжировками, результатами парных сравнений и другими объектами нечисловой природы, но не числами. *Распространенное заблуждение состоит в том, что ответы экспертов стараются рассматривать как числа, занимаются "оцифровкой" их мнений, приписывая этим мнениям численные значения - баллы, которые потом обрабатывают с помощью методов прикладной статистики как результаты обычных физико-технических измерений.* В случае произвольности "оцифровки" выводы, полученные в результате обработки данных, могут не иметь отношения к реальности. В связи с "оцифровкой" уместно вспомнить классическую притчу о человеке, который ищет потерянные ключи под фонарем, хотя потерял их в кустах. На вопрос, почему он так делает, отвечает: "Под фонарем светлее". Это, конечно, верно. Но, к сожалению, весьма малы шансы найти потерянные ключи под фонарем. Так и с "оцифровкой" нечисловых данных. Она дает возможность имитации научной деятельности, но не возможность найти истину.

Проверка согласованности мнений экспертов и классификация экспертных мнений. Ясно, что мнения разных экспертов различаются. Важно понять, насколько велико это различие. Если мало - усреднение мнений экспертов позволит выделить то общее, что есть у всех экспертов, отбросив случайные отклонения в ту или иную сторону. Если велико - усреднение является чисто формальной процедурой. Так, если представить себе, что ответы экспертов равномерно покрывают поверхность бублика, то формальное усреднение укажет на центр дырки от бублика, а такого мнения не придерживается ни один эксперт. Из сказанного ясна важность проблемы проверки согласованности мнений экспертов.

Разработан ряд методов такой проверки. Статистические методы проверки согласованности зависят от математической природы ответов экспертов. Соответствующие статистические теории весьма трудны, если эти ответы - ранжировки или разбиения, и достаточно просты, если ответы - результаты независимых парных сравнений. Отсюда вытекает рекомендация по организации экспертного опроса: не старайтесь сразу получить от эксперта ранжировку или разбиение, ему трудно это сделать, да и имеющиеся математические методы не позволяют далеко продвинуться в анализе подобных данных. Например, рекомендуют проверять согласованность ранжировок с помощью коэффициента ранговой конкордации Кендалла-Смита. Но давайте вспомним, какая статистическая модель при этом используется. Проверяется нулевая гипотеза, согласно которой ранжировки независимы и равномерно распределены на множестве всех ранжировок. Если эта гипотеза принимается, то конечно, ни о какой согласованности мнений экспертов говорить нельзя. А если отклоняется? Тоже нельзя. Например, может быть два (или больше) центра, около которых группируются ответы экспертов. Нулевая гипотеза отклоняется. Но разве можно говорить о согласованности?

Эксперту гораздо легче на каждом шагу сравнивать только два объекта. Пусть он занимается парными сравнениями. *Непараметрическая теория парных сравнений (теория* *люсианов) позволяет решать более сложные задачи, чем статистика ранжировок или разбиений.* В частности, вместо гипотезы равномерного распределения можно рассматривать гипотезу однородности, т.е. вместо совпадения всех распределений с одним фиксированным (равномерным) можно проверять лишь совпадение распределений мнений экспертов между собой, что естественно трактовать как согласованность их мнений. Таким образом, удается избавиться от неестественного предположения равномерности.

При отсутствии согласованности экспертов естественно разбить их на группы сходных по мнению. Это можно сделать различными методами статистики объектов нечисловой природы, относящимися к кластер-анализу, предварительно введя метрику в пространство мнений экспертов. Идея американского математика Джона Кемени об аксиоматическом введении метрик (см. ниже) нашла многочисленных продолжателей. Однако методы кластер-анализа обычно являются эвристическими. В частности, невозможно с позиций статистической теории обосновать "законность" объединения двух кластеров в один. Имеется важное исключение - *для независимых парных сравнений (люсианов) разработаны методы, позволяющие проверять возможность объединения кластеров как статистическую гипотезу.* Это - еще один аргумент за то, чтобы рассматривать теорию люсианов как ядро математических методов экспертных оценок.

Нахождение итогового мнения комиссии экспертов. Пусть мнения комиссии экспертов или какой-то ее части признаны согласованными. Каково же итоговое (среднее, общее) мнение комиссии? Согласно идее Джона Кемени следует найти среднее мнение как решение *оптимизационной задачи*. А именно, надо минимизировать суммарное расстояние от кандидата в средние до мнений экспертов. Найденное таким способом среднее мнение называют "медианой Кемени".

Математическая сложность состоит в том, что мнения экспертов лежат в некотором пространстве объектов нечисловой природы. Общая теория подобного усреднения рассмотрена выше. В частности, показано, что в силу обобщения закона больших чисел среднее мнение при увеличении числа экспертов (чьи мнения независимы и одинаково распределены) приближается к некоторому пределу, который естественно назвать *математическим ожиданием* (случайного элемента, имеющего то же распределение, что и ответы экспертов).

В конкретных пространствах нечисловых мнений экспертов вычисление медианы Кемени может быть достаточно сложным делом. Кроме свойств пространства, велика роль конкретных метрик. Так, в пространстве ранжировок при использовании метрики, связанной с коэффициентом ранговой корреляции Кендалла, необходимо проводить достаточно сложные расчеты, в то время как применение показателя различия на основе коэффициента ранговой корреляции Спирмена приводит к упорядочению по средним рангам.

Бинарные отношения и расстояние Кемени. Как известно, бинарное отношение A на конечном множестве $Q = \{q_1, q_2, ..., q_k\}$ - это подмножество *декартова квадрата* $Q^2 = \{ (q_m, q_n), m,n = 1,2,...,k \}$. При этом пара (q_m, q_n) входит в A тогда и только тогда, когда между q_m и q_n имеется рассматриваемое отношение.

Каждую кластеризованную ранжировку, как и любое бинарное отношение, можно задать матрицей $\| x(a, b) \|$ из 0 и 1 порядка k x k. При этом $x(a, b) = 1$ тогда и только тогда, когда $a < b$ либо $a = b$. В первом случае $x(b, a) = 0$, а во втором $x(b, a) = 1$. При этом хотя бы одно из чисел $x(a, b)$ и $x(b, a)$ равно 1.

Как использовать связь между ранжировками и матрицами? Например, из определения противоречивости пары (a, b) вытекает, что для нахождения всех таких пар можно воспользоваться матрицами, соответствующими ранжировкам. Достаточно поэлементно перемножить две матрицы $\|x(a,b)\|$ и $\|y(a, b)\|$, соответствующие двум кластеризованным ранжировкам, и отобрать те и только те пары, для которых $x(a,b)y(a,b)=x(b,a)y(b,a)=0$.

В экспертных методах используют, в частности, такие бинарные отношения, как ранжировки (упорядочения, или разбиения на группы, между которыми имеется строгий порядок), отношения эквивалентности, толерантности (отношения сходства). Как следует из сказанного выше, каждое бинарное отношение A можно описать матрицей $\| a(i,j) \|$ из 0 и 1, причем $a(i,j) = 1$ тогда и только тогда, когда q_i и q_j находятся в отношении A, и $a(i,j) = 0$ в противном случае.

Определение. *Расстоянием Кемени между бинарными отношениями A и B, описываемыми матрицами* $\| a(i,j) \|$ и $\| b(i,j) \|$ *соответственно, называется число*

$$D(A, B) = \sum_{i,j=1}^{k} \left| a(i,j) - b(i,j) \right| ,$$

где суммирование производится по всем i,j от 1 до k, т.е. расстояние Кемени между бинарными отношениями равно сумме модулей разностей элементов, стоящих на одних и тех же местах в соответствующих им матрицах.

Легко видеть, что расстояние Кемени - это число несовпадающих элементов в матрицах $\| a(i,j) \|$ и $\| b(i,j) \|$.

Расстояние Кемени основано на некоторой системе аксиом. Эта система аксиом и вывод из нее формулы для расстояния Кемени между упорядочениями содержится в книге [5], которая сыграла большую роль в развитии в нашей стране такого научного направления, как анализ нечисловой информации. В дальнейшем под влиянием Дж. Кемени были предложены различные системы аксиом для получения расстояний в тех или иных используемых в социально-экономических исследованиях пространствах, например, в пространствах множеств [4].

Медиана Кемени и законы больших чисел. С помощью расстояния Кемени находят итоговое мнение комиссии экспертов. Пусть A_1 , A_2 , A_3 ,…, A_p - ответы p экспертов, представленные в виде бинарных отношений. Для их усреднения используют т.н. **медиану Кемени**

$$\mathrm{Arg\,min} \sum_{i=1}^{p} D(A_i, A) ,$$

где Arg min - то или те значения A, при которых достигает минимума указанная сумма расстояний Кемени от ответов экспертов до текущей переменной A, по которой и проводится минимизация. Таким образом,

$$\sum_{i=1}^{p} D(A_i, A) = D(A_1, A) + D(A_2, A) + D(A_3, A) + \ldots + D(A_p, A) .$$

Кроме медианы Кемени, используют **среднее по Кемени,** в котором вместо $D(A_i, A)$ стоит $D^2(A_i, A)$.

Медиана Кемени - частный случай определения эмпирического среднего в пространствах нечисловой природы. Для нее справедлив закон больших чисел, т.е. эмпирическое среднее приближается при росте числа составляющих (т.е. p - числа слагаемых в сумме), к теоретическому среднему:

$$\mathrm{Arg\,min} \sum_{i=1}^{p} D(A_i, A) \rightarrow \mathrm{Arg\,min}\, M\, D(A_1, A) .$$

Здесь M - символ математического ожидания. Предполагается, что ответы p экспертов A_1 , A_2 , A_3 ,…, A_p есть основания рассматривать как независимые одинаково распределенные случайные элементы (т.е. как случайную выборку) в соответствующем пространстве произвольной природы, например, в пространстве упорядочений или отношений эквивалентности. Систематически эмпирические и теоретические средние и соответствующие законы больших чисел рассмотрены в соответствующей главе настоящей книги.

Законы больших чисел показывают, во-первых, что медиана Кемени обладает *устойчивостью* по отношению к незначительному изменению состава экспертной комиссии; во-вторых, при увеличении числа экспертов она *приближается к некоторому пределу.* Его естественно рассматривать как *истинное мнение* экспертов, от которого каждый из них несколько отклонялся по случайным причинам.

Рассматриваемый здесь закон больших чисел является обобщением известного в статистике "классического" закона больших чисел. Он основан на иной математической базе - теории оптимизации, в то время как "классический" закон больших чисел использует суммирование. Упорядочения и другие бинарные отношения нельзя складывать, поэтому приходится применять иную математику. Рассмотрим пример вычисления медианы Кемени.

Пример. Пусть дана квадратная матрица (порядка 9) попарных расстояний для множества бинарных отношений из 9 элементов $A_1, A_2, A_3,..., A_9$ (см. табл.3). Необходимо найти в этом множестве *медиану* для множества из 5 элементов $\{A_2, A_4, A_5, A_8, A_9\}$.

Табл. 3. Матрица попарных расстояний

0	2	13	1	7	4	10	3	11
2	0	5	6	1	3	2	5	1
13	5	0	2	2	7	6	5	7
1	6	2	0	5	4	3	8	8
7	1	2	5	0	10	1	3	7
4	3	7	4	10	0	2	1	5
10	2	6	3	1	2	0	6	3
3	5	5	8	3	1	6	0	9
11	1	7	8	7	5	3	9	0

В соответствии с определением медианы Кемени следует ввести в рассмотрение функцию

$$C(A) = \sum D(A_i,A) = D(A_2,A)+D(A_4,A)+D(A_5,A)+D(A_8,A)+D(A_9,A),$$

рассчитать ее значения для всех A_1, A_2, A_3,..., A_9 и выбрать наименьшее. Проведем расчеты:

$$C(A_1) = D(A_2,A_1) + D(A_4,A_1) + D(A_5,A_1) + D(A_8,A_1) + D(A_9,A_1) = $$
$$= 2 + 1 + 7 + 3 + 11 = 24,$$
$$C(A_2) = D(A_2,A_2) + D(A_4,A_2) + D(A_5,A_2) + D(A_8,A_2) + D(A_9,A_2) = $$
$$= 0 + 6 + 1 + 5 + 1 = 13,$$
$$C(A_3) = D(A_2,A_3) + D(A_4,A_3) + D(A_5,A_3) + D(A_8,A_3) + D(A_9,A_3) = $$
$$= 5 + 2 + 2 + 5 + 7 = 21,$$
$$C(A_4) = D(A_2,A_4) + D(A_4,A_4) + D(A_5,A_4) + D(A_8,A_4) + D(A_9,A_4) = $$
$$= 6 + 0 + 5 + 8 + 8 = 27,$$
$$C(A_5) = D(A_2,A_5) + D(A_4,A_5) + D(A_5,A_5) + D(A_8,A_5) + D(A_9,A_5) = $$
$$= 1 + 5 + 0 + 3 + 7 = 16,$$
$$C(A_6) = D(A_2,A_6) + D(A_4,A_6) + D(A_5,A_6) + D(A_8,A_6) + D(A_9,A_6) = $$
$$= 3 + 4 + 10 + 1 + 5 = 23,$$
$$C(A_7) = D(A_2,A_7) + D(A_4,A_7) + D(A_5,A_7) + D(A_8,A_7) + D(A_9,A_7) = $$
$$= 2 + 3 + 1 + 6 + 3 = 15,$$
$$C(A_8) = D(A_2,A_8) + D(A_4,A_8) + D(A_5,A_8) + D(A_8,A_8) + D(A_9,A_8) = $$
$$= 5 + 8 + 3 + 0 + 9 = 25,$$
$$C(A_9) = D(A_2,A_9) + D(A_4,A_9) + D(A_5,A_9) + D(A_8,A_9) + D(A_9,A_9) = $$
$$= 1 + 8 + 7 + 9 + 0 = 25.$$

Из всех вычисленных сумм наименьшая равна 13, и достигается она при $A = A_2$, следовательно, медиана Кемени - это A_2 .

Обратим внимание на то, что минимум может достигаться не в одной точке, а в нескольких. Поэтому медиана Кемени - это, вообще говоря, не элемент соответствующего пространства, а его подмножество. Поэтому более правильно сказать, что данных табл.3 медиана Кемени - это множество $\{A_2\}$, состоящее из одного элемента A_2 , т.е. в условиях примера

$$\mathrm{Arg\ min} \sum_{i=1}^{p} D\,(A_i\,,A) = \{A_2\}.$$

В общем случае вычисление медианы Кемени - задача целочисленного программирования. В частности, для ее нахождения используется различные алгоритмы дискретной оптимизации, в частности, основанные на методе ветвей и границ. Применяют также алгоритмы, основанные на идее случайного поиска, поскольку для каждого бинарного отношения нетрудно найти множество его соседей.

Разработано весьма много различных методов экспертного оценивания (см., например, обзор []).

Цитированная литература

1. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. - М.: Наука, 1983. - - 416 с.
2. Шрейдер Ю.А. Равенство, сходство, порядок. М.: Наука, 1971.
3. Горский В.Г., Орлов А.И., Гриценко А.А. Метод согласования кластеризованных ранжировок // Автоматика и телемеханика. 2000. №3. С. 159-167.
4. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979. - 296 с.
5. Кемени Дж., Снелл Дж. Кибернетическое моделирование: Некоторые приложения. - М.: Советское радио, 1972. - 192 с.
6. Орлов А.И. Экспертные оценки // Заводская лаборатория. 1996. Т.62. № 1. С.54-60.

Глава 13. Эконометрические методы управления качеством и сертификации продукции

Одна из наиболее важных прикладных областей эконометрики, приносящих наибольший доход в денежном выражении - это эконометрика качества, основанная на применении статистических методов для обеспечения надлежащего качества продукции. Японцы считают: "Все, начиная от председателя Совета Директоров и до рядового рабочего в цехе, обязаны знать хотя бы основы статистических методов." Основам эконометрических методов управления качеством и сертификации продукции посвящена настоящая глава.

13.1. Основы статистического контроля качества продукции

Сначала дадим общие сведения о месте статистических методов в управлении качеством и сертификации продукции. Затем рассмотрим центральную тему эконометрики качества - статистический контроль качества продукции. Продемонстрируем его высокую экономическую эффективность.

Качество продукции и рыночная экономика. Руководители и специалисты промышленных предприятий хотят не только выжить, но и выиграть в борьбе с конкурентами. Более частными задачами, которые они хотят решить, обычно являются:

- выйти на международный рынок;
- поднять качество продукции до японского уровня;
- полностью ликвидировать рекламации, и т.д.

Для решения этих задач им надо повышать качество продукции. Все руководители и специалисты промышленных предприятий это хорошо знают, слова "сертификация", "международные стандарты ИСО серии 9000 по системам качества" уже навязли в зубах. Менее осознано, что управление качеством - это прежде всего применение современных статистических методов. На Западе (США) и на Востоке (Япония) это - аксиома. Вот типичное высказывание японского менеджера и инженера: "Методы статистики - именно то средство, которое необходимо изучить, чтобы внедрить управление качеством. Они - наиболее важная составная часть комплексной системы всеобщего управления качеством на фирме. В японских корпорациях все, начиная от председателя Совета Директоров и до рядового рабочего в цехе, обязаны знать хотя бы основы статистических методов." Так считает Каору Исикава, президент промышленного института Мусаси, заслуженный профессор Токийского университета (цитируется по японскому пособию по статистическим методам обеспечения качества [1, с.15]).

Раз все японские работники знают про статистические методы - значит, их научили в школе. Во всем мире - в США, Японии и Ботсване - школьники учат статистические методы как один из обязательных школьных предметов, вместе с физикой, химией, математикой и историей. ЮНЕСКО регулярно проводит конференции по преподаванию статистики в средней школе. И вот всем виден результат - качество компьютеров IBM и японских телевизоров. А отечественные бюрократы десятилетиями "боролись за качество" (вспомните, была "пятилетка эффективности и качества"!), "внедряли" кипы бумаг - КС УКП ... (популярное сочетание в 1970-е и 1980-е годы: КС УКП - это Комплексная Система Управления Качеством Продукции; имелись областные варианты - горьковская, львовская, днепропетровская и т.д.).

Справедливости ради надо отметить, что популярные ныне международные стандарты ИСО серии 9000 ничем принципиально не отличаются от давних документов КС УКП, а в некоторых отношениях КС УКП были более прогрессивными, чем нынешние стандарты ИСО

9000. В очередной раз придуманное в нашей стране попало на Запад, было там оформлено по-другому, а потом стало внедряться у нас как последнее достижение западной цивилизации…

ИТОГ НА СЕГОДНЯ: весь мир, кроме нас, знает статистические методы и повсеместно применяет их для повышения качества. Мы вынуждены догонять. Очевидно, овладение основами статистического контроля качества продукции - неотъемлемая часть экономического и тем более эконометрического образования.

О сертификации. Вслед за т. н. развитыми странами в России намечается всё расширяющаяся тенденция к сертификации продукции, т.е. к официальной гарантии поставки производителем продукции, удовлетворяющей установленным требованиям. Средства массовой информации отмечают, что в условиях рыночной экономики поставщики и продавцы должны иметь сертификаты качества на предлагаемые ими товары и услуги. Маркетинг, т.е. "производственная и коммерческая политика, нацеленная на получение максимальной прибыли на основе изучения рынка, создания конкурентоспособной продукции и её полной реализации" (определение взято из выпущенной нашим Центром брошюры [2, с.64-65]), включает в себя работы по сертификации. За новыми терминами зачастую скрываются хорошо известные понятия, несколько модернизированные в соответствии с современной обстановкой. Так, целесообразно связать комплексную систему управления качеством продукции с маркетингом: "маркетинг в широком смысле - это усовершенствованная, ориентированная на рыночную экономику КС УКП" [2,с.61].

Есть несколько уровней сертификации. Говоря о сертификации продукции, могут иметь в виду качество конкретной её партии. В ряде случаев это оправдано - рядового потребителя интересует качество лишь той единицы продукции, которую он сам приобрёл. Однако установление долговременных хозяйственных связей целесообразно лишь в случае, когда поставщик гарантирует высокое качество не одной, а всех партий своей продукции. Очевидно, для этого должны быть проведены оценка и сертификация технологических процессов и производств, обеспечивающих выпуск этой продукции.

Ещё больше повышается доверие к поставщику, если не только отдельные технологические процессы, но и всё предприятие в целом гарантированно выпускает продукцию высокого качества. Это обеспечивается действующей на предприятии системой качества, удовлетворяющей требованиям Международной организации по стандартизации (*International Standartization Organization,* сокращенно *ISO*, по-русски - ИСО), выраженным в системе стандартов ИСО 9000, о которых уже шла речь.

В условиях рыночной экономики основная характеристика товара - его конкурентоспособность. Очевидно, производителю необходимо уметь оценивать конкурентоспособность перед запуском продукции в производство или началом работы по продвижению на зарубежный рынок. Одним из основных компонентов конкурентоспособности является технический уровень продукции. Фирма, обладающая патентом или новой научно-технической разработкой, имеет более высокий излишек производителя по сравнению с другими фирмами. При выборе направления инвестиционных вложений одна из основных учитываемых характеристик - технический уровень продукции.

Из сказанного вытекает, что сертификация продукции - это современная форма управления качеством продукции. На Западе общепринято, что основная составляющая в управлении качеством продукции - это статистические методы (см., например, отчет Комитета ИСО по изучению принципов стандартизации [3]). В нашей стране внедрение комплексных систем управления качеством (КС УКП), как уже отмечалось, сводилось во многом к подготовке документации организационного характера. Статистические методы использовались в промышленности недостаточно, а государственные стандарты по этой тематике зачастую содержали грубейшие ошибки (см. ниже).

Подготовка предприятий к сертификации продукции, технологических процессов и производств, систем качества требует приложения труда квалифицированных специалистов, причем в достаточно большом объеме. Подобную работу обычно проводят специализированные организации.

О развитии статистических методов сертификации в России. Около 150 лет статистические методы применяются в России для проверки соответствия продукции установленным требованиям, т.е. для сертификации. Так, еще в 1846 г. действительный член Петербургской академии наук М.В. Остроградский рассматривал задачу статистического контроля партий мешков муки или штук сукна армейскими поставщиками. С тех пор в России в статистическом контроле качества было сделано многое, особенно в области теории: Так, солидные монографии проф. Ю.К. Беляева и проф. Я.П. Лумельского можно смело назвать классическими. Был выпущен и длинный ряд практических руководств, в основном переводных.

С начала 1970-х годов стали разрабатываться государственные стандарты по статистическим методам. В связи с обнаружением в них грубых ошибок 24 из 31 государственного стандарта по статистическим методам были отменены в 1986-87 гг. (перечень стандартов и описание ошибок приведены в работе [4]). К сожалению, потеряв правовую силу как нормативные документы, ошибочные стандарты продолжают использоваться как научно-технические издания.

В 1989 г. был организован Центр статистических методов и информатики (ЦСМИ) для работ по развитию и внедрению современных статистических методов. Уже к середине 1990 г. ЦСМИ были разработаны 7 диалоговых систем по современным статистическим методам управления качеством, а именно, СПК, АТСТАТ-ПРП, СТАТКОН, АВРОРА-РС, ЭКСПЛАН, ПАСЭК, НАДИС (описания этих систем даны в работе [5]).

Параллельно ЦСМИ вел работу по объединению статистиков. В апреле 1990 г. в Большом Актовом Зале Московского Энергетического института прошла Учредительная конференция Всесоюзной организации по статистическим методам и их применениям. На Учредительном съезде Всесоюзной статистической ассоциации (ВСА) в октябре 1990 г. в Московском экономико-статистическом институте эта организация вошла в состав ВСА в качестве секции статистических методов. В 1992г. после развала СССР и фактического прекращения работы ВСА на основе секции статистических методов ВСА организована Российская ассоциация по статистическим методам (РАСМ), а затем и Российская академия статистических методов, существующие и в настоящее время. В мероприятиях секции статистических методов ВСА и РАСМ активно участвовали несколько сот человек. Основной тематикой работ многих из этих специалистов являются статистические методы в сертификации (управлении качеством). В ЦСМИ и РАСМ, объединивших большинство ведущих российских специалистов, коллективными усилиями разработан единый подход к проблемам применения статистических методов в сертификации и управлении качеством.

Статистический контроль - это выборочный контроль на научной основе. Контроль качества продукции всем знаком хотя бы по названию - им обычно занимается отдел технического контроля (ОТК) предприятия. Есть различные виды контроля - входной контроль, приемочный контроль (готовой продукции), и контроль при передаче полуфабрикатов и комплектующих из цеха в цех. Кроме сплошного контроля всех изделий подряд применяют выборочный, когда о качестве партии продукции судят по результатам контроля некоторой части - выборки.

Зачем нужен выборочный контроль? Чтобы проверить качество спички - надо чиркнуть ею. Загорится - должное качество, не загорится - брак. Но повторно однажды зажженную спичку использовать уже нельзя. Поэтому партию спичек можно контролировать только выборочно. Партии консервов, лампочек, патронов - тоже. Т.е. при разрушающем контроле необходимо пользоваться выборочными методами и судить о качестве партии продукции по результатам контроля её части - выборки.

Выборочные методы контроля могут применяться и из экономических соображений, когда стоимость контроля высока по сравнению со стоимостью изделия. Например, вряд ли целесообразно визуально проверять качество каждой скрепки в каждой коробке.

Для проведения выборочного контроля необходимо сформировать выборку, выбрать план контроля. А если план имеется - полезно знать его свойства. Анализ и синтез планов проводят с помощью математического моделирования на основе теории вероятностей и математической статистики, применяя компьютерные диалоговые системы (пакеты программ).

Зачем нужны диалоговые системы по статистическому контролю? Раньше, действительно, ОТК формально применяли планы контроля из ГОСТов на конкретную продукцию, а реальное качество выпускаемых изделий никого не интересовало. Сейчас - ситуация начинает меняться. С декабря 1990 г. обязательность большинства ГОСТов отменена (в части основных показателей качества, кроме показателей безопасности). У промышленности сняты кандалы. Но - со становлением рыночной экономики появляются конкуренты. В том числе зарубежные. Руководителям производства приходится отлаживать систему контроля качества не для галочки, не по приказу обкома, а для повышения доходов предприятий. А потому - и собственных тоже.

Компьютерные диалоговые системы позволяют прежде всего проводить анализ и синтез планов контроля. Пусть перед Вами - прежний ГОСТ на продукцию, в нем есть раздел "Правила приемки" с планами контроля. Хороша эта система планов или плоха? С помощью диалоговых систем Вы найдете характеристики конкретного плана, приемочный и браковочный уровни дефектности (см. ниже) и т.д. Можно провести и синтез планов, т.е. компьютер подберет план, удовлетворяющий Вашим условиям.

Российской ассоциацией статистических методов были проанализированы сотни стандартов на конкретную продукцию (разделы "Правила приемки") и ГОСТы по статистическим методам. Обнаружено, что более половины и тех и других стандартов содержат грубые ошибки, пользоваться ими нельзя. Причины этого печального положения проанализированы в статье [4]. В отличие от ГОСТов, диалоговым системам ЦСМИ по статистическому контролю верить можно и нужно. И экономически выгодно. По оценкам, полученным в работе [6], применение современных статистических методов позволяет в среднем вдвое сократить трудозатраты на контрольные операции (как известно, на них расходуют примерно 10% от стоимости машиностроительной продукции). Следовательно, от внедрения современных эконометрических методов обеспечения качества продукции Россия может получить более 5 миллиардов долларов США дополнительного дохода в год.

Приведем ещё два сообщения о высокой экономической эффективности статистического контроля. "Мы документально зафиксировали экономию от применения методов статистического контроля и методов разрешения проблем, которым обучили наших сотрудников. Мы приближаемся к степени окупаемости около 30 долларов на 1 вложенный доллар. Вот почему мы получили такую серьезную поддержку от высшего руководства", - сообщает Билл Виггенхорн, ответственный за подготовку специалистов фирмы "Моторола" (цитируем по статье [7]).

По подсчетам профессора Массачусетского технологического института Фримена (см. монографию [8]), только статистический приемочный контроль давал промышленности США 4 миллиарда долларов в 1958 г. (это более 20 миллиарда долларов в ценах 2001 г.), т.е. 0,8% ВВП - Валового Внутреннего Продукта.

На наш взгляд, российским предпринимателям и менеджерам промышленных предприятий целесообразно равняться на японских коллег - знать хотя бы основы статистических методов, т.е. эконометрики, и активно их применять, постоянно консультируясь со специалистами-эконометриками.

Основы статистического контроля. Выборочный контроль, построенный на научной основе, т.е. исходящий из теории вероятностей и математической статистики, называют статистическим контролем. Предпринимателя и менеджера выборочный контроль может интересовать не только в связи с качеством продукции, но и в связи, например, с контролем экологической обстановки, поскольку зафиксированные государственными органами экологические нарушения влекут штрафы и иные "неприятные" последствия. Обсудим основные подходы статистического контроля.

При статистическом контроле решение о генеральной совокупности – об экологической обстановке в данном регионе или о партии продукции - принимается по выборке, состоящей из некоторого количества единиц (единиц экологического контроля или единиц продукции). Следовательно, выборка должна представлять партию, т.е. быть репрезентативной (представительной). Как эти слова понимать, как проверить репрезентативность? Ответ может быть дан лишь в терминах вероятностных моделей выборки.

Наиболее распространенными являются две вероятностные модели—биномиальная и гипергеометрическая. В биномиальной модели предполагается, что результаты контроля n единиц можно рассматривать как совокупность n независимых одинаково распределенных случайных величин $X_1, X_2,\ldots,X_n$, где $X_i = 1$, если i-ое измерение показывает, что есть нарушение, т.е. превышено ПДК (предельная норма концентрации) или i-ое изделие дефектно, и $X_i= 0$, если это не так. Тогда число X превышений ПДК или дефектных единиц продукции в партии равно

$$X= X_1+ X_2+\ldots+ X_n \quad .(1)$$

Из формулы (1) и Центральной Предельной Теоремы теории вероятностей вытекает, что при увеличении объема выборки n распределение X сближается с нормальным распределением. Известно, что распределение X имеет вид

$$P(X= k) = C_n^k\ p^k (1—p)^{n-k}\ ,\ (2)$$

где C_n^k - число сочетаний из n элементов по k, а p —уровень дефектности (в другой предметной области - доля превышений ПДК в генеральной совокупности), т.е. $p = P(X_i= 1)$. Формула (2) задает так называемое биномиальное распределение.

Гипергеометрическое распределение соответствует случайному отбору единиц в выборку. Пусть среди N единиц, составляющих генеральную совокупность, имеется D дефектных. Случайность отбора означает, что каждая единица имеет одинаковые шансы попасть в выборку. Мало того, ни одна пара единиц не должна иметь при отборе в выборку преимущества перед любой другой парой. То же самое —для троек, четверок и т.д. Это условие выполнено тогда и только тогда, когда каждое из C_N^n сочетаний по n единиц из N имеет одинаковые шансы быть отобранным в качестве выборки. Вероятность того, что будет отобрано заранее заданное сочетание, равна, очевидно, $1/\ C_N^n$.

Отбор случайной выборки согласно описанным правилам организуют при проведении различных лотерей. Пусть Y —число дефектных единиц в случайной выборке, организованной таким образом. Известно, что тогда $P\ (Y = k)$ – гипергеометрическое распределение, т.е.

$$P(Y = k) = \frac{C_n^k C_{N-n}^{D-k}}{C_N^n}. \qquad (3)$$

Замечательный математический результат состоит в том, что биномиальная и гипергеометрическая модели *весьма близки*, когда объем генеральной совокупности (партии) по крайней мере в 10 раз превышает объем выборки. Другими словами, можно принять, что

$$P(X = k) = P (Y = k), \qquad (4)$$

если объем выборки мал по сравнению с объемом партии. При этом в качестве p в формуле (4) берут D/N. Близость результатов, получаемых с помощью биномиальной и гипергеометрической моделей, весьма важна с философской точки зрения. Дело в том, что эти модели исходят из принципиально различных философских предпосылок. В биномиальной модели случайность *присуща каждой единице* - она с какой-то вероятностью дефектна, а с какой-то - годна. В то же время в гипергеометрической модели качество определенной единицы детерминировано, задано, а случайность проявляется лишь *в отборе*, вносится экологом или экономистом при составлении выборки. В науках о человеке противоречие между аналогичными моделями выборки еще более выражено. Биномиальная модель предполагает, что поведение человека, в частности, выбор им определенного варианта при ответе на вопрос, определяется с участием случайных причин. Например, человек может случайно сказать «да», случайно—«нет». Некоторые философы отрицают присущую человеку случайность. Они верят в причинность и считают поведение конкретного человека практически полностью детерминированным. Поэтому они принимают гипергеометрическую модель и считают, что случайность отличия ответов в выборке от ответов во всей генеральной совокупности определяется всецело случайностью, вносимой при отборе единиц наблюдения в выборку.

Соотношение (4) показывают, что во многих случаях нет необходимости принимать чью-либо сторону в этом споре, поскольку обе модели дают близкие численные результаты. Отличия проявляются при обсуждении вопроса о том, какую выборку считать *представительной*. Является ли таковой выборка, составленная из 20 изделий, лежащих сверху в первом вскрытом ящике? В биномиальной модели - да, в гипергеометрической - нет.

Биномиальная модель легче для теоретического изучения, поэтому будем её рассматривать в дальнейшем. Однако при реальном контроле лучше формировать выборку, исходя из гипергеометрической модели. Это делают, выбирая номера изделий (для включения в выборку) с помощью датчиков псевдослучайных чисел на ЭВМ (см. главу 11) или с помощью таблиц псевдослучайных чисел. Алгоритмы формирования выборки встраивают в современные программные продукты по статистическому контролю.

Планы статистического контроля и правила принятия решений. Под планом статистического контроля понимают алгоритм, т.е. правила действий, на входе при этом—генеральная совокупность (партия продукции), а на выходе—одно из двух решений: «принять партию» либо «забраковать партию». Рассмотрим несколько примеров.

Одноступенчатые планы контроля (n,c): отобрать выборку объема n; если число дефектных единиц в выборке X не превосходит c, то партию принять, в противном случае забраковать. Число c называется приемочным.

Частные случаи: план $(n,0)$ —партию принять тогда и только тогда, когда все единицы в выборке являются годными; план $(n,1)$ - партия принимается, если в выборке все единицы являются годными или ровно одно - дефектное, во всех остальных случаях партия бракуется.

Двухступенчатый план контроля $(n,a,b) + (m,c)$: отобрать первую выборку объема n; если число дефектных единиц в первой выборке X не превосходит a , то партию принять; если число дефектных единиц в первой выборке X больше или равно b, то партию забраковать; во всех остальных случаях, т.е. когда X больше a, но меньше b, следует взять вторую выборку объема m; если число дефектных единиц во второй выборке Y не превосходит c, то партию принять, в противном случае забраковать.

Рассмотрим в качестве примера план (20, 0, 2) + (40, 0). Сначала берется первая выборка объема 20. Если все единицы в ней - годные, то партия принимается. Если две или больше - дефектные, партия бракуется. А если только одно - дефектное? В реальной ситуации в таких случаях начинаются споры между представителями предприятия и экологического контроля, или поставщика и потребителя. Говорят, например, что дефектная единица случайно попала в партию, что ее подсунули конкуренты или что при контроле случайно сделан неправильный вывод. Поэтому, чтобы споры пресечь, берут вторую выборку объема 40 (вдвое большего, чем в первый раз). Если все единицы во второй выборке - годные, то партию принимают, в противном случае - бракуют.

В реальной нормативно-технической документации - договорах на поставку, стандартах, технических условиях, инструкциях по экологическому контролю и т.д. - не всегда четко сформулированы планы статистического контроля и правила принятия решений. Например, при описании двухступенчатого плана контроля вместо задания приемочного числа c может стоять загадочная фраза "результат контроля второй выборки считается окончательным". Остается гадать, как принимать решение по второй выборке. Менеджер, администратор (государственный служащий), эколог или экономист, занимающийся вопросами экологического контроля или контроля качества, должен первым делам добиваться кристальной ясности в формулировках правил принятия решений, иначе ошибочные и необоснованные решения, а потому и убытки неизбежны.

Оперативная характеристика плана статистического контроля. Каковы свойства плана статистического контроля? Они, как правило, определяются с помощью функции $f(p)$, связывающей вероятность p дефектности единицы контроля с вероятностью $f(p)$ положительной оценки экологической обстановки (приемки партии) по результатам контроля. При этом вероятность p того, что конкретная единица дефектна, называется входным уровнем дефектности, а указанная функция называется оперативной характеристикой плана контроля. Если дефектные единицы отсутствуют, $p = 0$, то партия всегда принимается, т.е. $f(0) = 1$. Если все единицы дефектные, $p = 1$, то партия наверняка бракуется, $f(1) = 0$. Между этими крайними значениями p функция $f(p)$ монотонно убывает.

Вычислим оперативную характеристику плана (n,0). Поскольку партия принимается тогда и только тогда, когда все единицы являются годными, а вероятность того, что конкретная единица—годная, равна $(1-p)$, то оперативная характеристика имеет вид

$$f(p) = P(X=0) = (1-p)^n. \qquad (5)$$

Для плана (n,1) оперативная характеристика, как легко видеть, такова:

$$f(p) = P(X=0)+P(X=1) = (1-p)^n + n\,(1-p)^{n-1} \qquad (6)$$

Оперативные характеристики для конкретных планов статистического контроля не всегда имеют такой простой вид, как в случае формул (5) и (6). Рассмотрим в качестве примера план (20, 0, 2) + (40, 0). Сначала найдем вероятность того, что партия будет принята по результатам контроля первой партии. Согласно формуле (5) имеем:

$$f_1(p) = P(X=0) = (1-p)^{20}.$$

Вероятность того, что понадобится контроль второй выборки, равна

$$P(X=1) = 20(1-p)^{19}.$$

При этом вероятность того, что по результатам её контроля партия будет принята, равна

$$f_2(p) = P(X=0) = (1-p)^{40}.$$

Следовательно, вероятность того, что партия будет принята со второй попытки, т.е. что при контроле первой выборки обнаружится ровно одна дефектная единица, а затем при контроле второй—ни одной, равна

$$f_3(p) = P(X=1)\,f_2(p) = 20(1-p)^{19}(1-p)^{40}= 20(1-p)^{59}.$$

Следовательно, вероятность принятия партии с первой или со второй попытки равна

$$f(p) = f_1(p) + f_3(p) = (1-p)^{20}+ 20(1-p)^{59}.$$

При практическом применении методов статистического приемочного контроля для нахождения оперативных характеристик планов контроля вместо формул, имеющих обозримый вид лишь для отдельных видов планов, применяют численные компьютерные алгоритмы или заранее составленные таблицы.

Риск поставщика и риск потребителя, приемочный и браковочный уровни дефектности. С оперативной характеристикой связаны важные понятия *приемочного и браковочного уровней дефектности*, а также понятия *"риск поставщика"* и *"риск потребителя"*. Чтобы ввести эти понятия, на оперативной характеристике выделяют две характерные точки, делящие входные уровни дефектности на три зоны—*А*, *Б* и *В*. В зоне *А* все почти всегда хорошо, а именно - почти всегда экологическая обстановка признается благополучной, почти все партии принимаются. В зоне *В*, наоборот, почти всегда все плохо, а именно - почти всегда экологический контроль констатирует экологические нарушения, почти все партии бракуются. Зона. *Б* - буферная, переходная, промежуточная, в ней как вероятность приемки, так и вероятность браковки заметно отличаются от 0 и 1. Для задания границ между зонами выбирают два малых числа—риск поставщика (производителя, предприятия) α и риск потребителя (заказчика, системы экологического контроля) β, при этом границы между зонами задают два уровня дефектности - приемочный $p_{\text{пр}}$ и браковочный $p_{\text{бр}}$, определяемые из уравнений

$$f(p_{\text{пр}}) = 1 — \alpha\,,\ f(p_{\text{бр}}) = \beta\,. \qquad (7)$$

Таким образом, если входной уровень дефектности не превосходит $p_{\text{пр}}$, то вероятность забракования партии мала, т.е. не превосходит α. Приемочный уровень дефектности выделяет зону *А* значений входного уровня дефектности, в которой нарушения экологической безопасности почти всегда не отмечаются, партии почти всегда принимаются, т.е. соблюдаются интересы проверяемого предприятия (в экологии), поставщика (при контроле качества). Это - зона комфортности для поставщика. Если он обеспечивает работу (уровень дефектности) в этой зоне, то его никто не потревожит.

Если же входной уровень дефектности больше браковочного уровня дефектности $p_{\text{бр}}$, то нарушения почти наверняка фиксируются, партия почти всегда бракуется, т.е. экологи узнают о нарушениях, потребитель оказывается защищен от попадания к нему партий со столь высоким уровнем брака. Поэтому можно сказать, что в зоне *В* соблюдаются интересы потребителей - брак к ним не попадает.

При выборе плана контроля часто начинают с выбора приемочного и браковочного уровней дефектности. При этом выбор конкретного значения приемочного уровня дефектности отражает интересы поставщика, а выбор конкретного значения браковочного уровня дефектности - интересы потребителя. Можно доказать, что для любых положительных чисел α и β, и любых входных уровней дефектности $p_{\text{пр}}$ и $p_{\text{бр}}$, причем $p_{\text{пр}}$ меньше $p_{\text{бр}}$, найдется план контроля *(n,c)* такой, что его оперативная характеристика *f(p)* удовлетворяет неравенствам

$$f(p_{\text{пр}}) > 1 - \alpha\,,\ f(p_{\text{бр}}) < \beta\,.$$

При практических расчетах обычно принимают α = 0,05 (т.е. 5%) и β = 0,1 (т.е. 10%).

Вычислим приемочный и браковочный уровни дефектности для плана (n,0). Из формул (5) и (7) вытекает, что

$$(1 - p_{\text{пр}})^n = 1 - \alpha\,,\ p_{\text{пр}} = 1 - (1 - \alpha)^{1/n}.$$

Поскольку риск поставщика α мал, то из известного соотношения математического анализа

$$\sqrt[n]{1-\alpha} = 1 - \frac{\alpha}{n} + O\left(\frac{\alpha^2}{n^2}\right)$$

вытекает приближенная формула

$$p_{\text{пр}} \approx \frac{\alpha}{n}.$$

Для браковочного уровня дефектности имеем

$$p_{\text{бр}} = 1 - \beta^{1/n}.$$

При практическом применении методов статистического приемочного контроля для нахождения приемочных и браковочных уровней дефектности планов контроля вместо формул, имеющих обозримый вид лишь для отдельных видов планов, применяют численные компьютерные алгоритмы или заранее составленные таблицы, имеющиеся в нормативно-технической документации или научно-технических публикациях.

Предел среднего выходного уровня дефектности. Обсудим судьбу забракованной партии продукции. В зависимости от ситуации эта судьба может быть разной. Партия может быть утилизирована. Например, забракованная партия гвоздей может быть направлена на переплавку. У партии может быть понижена сортность, и она может быть продана по более низкой цене (при этом результаты выборочного контроля будут использованы не для проверки того, что выдержан заданный уровень качества, а для оценки реального уровня качества). Наконец, партия продукции может быть подвергнута сплошному контролю (для этого обычно привлекают инженеров из всех заводских служб). При сплошном контроле все дефектные изделия обнаруживаются и либо исправляются на месте, либо извлекаются из партии. В результате в партии остаются только годные изделия. Такая процедура называется "контроль с разбраковкой".

При среднем входном уровне дефектности p и применении контроля с разбраковкой с вероятностью $f(p)$ партия принимается (и уровень дефектности в ней по-прежнему равен p) и с вероятностью ($1 - f(p)$) бракуется и подвергается сплошному контролю, в результате чего к потребителю поступают только годные изделия. Следовательно, по формуле полной вероятности средний выходной уровень дефектности равен

$$f_1(p) = pf(p) + 0(1 - f(p)) = pf(p).$$

Средний выходной уровень дефектности $f_1(p)$ равен 0 при $p=0$ и $p=1$, положителен на интервале (0;1), а потому достигает на нем максимума, который в теории статистического контроля называется пределом среднего выходного уровня дефектности (сокращенно ПСВУД):

$$\text{ПСВУД} = \max_{0 \le p \le 1} f_1(p).$$

Пример. Рассмотрим план (n,0). Для него $f(p) = (1 - p)^n$ и $f_1(p) = p(1-p)^n$. Чтобы найти ПСВУД, надо приравнять 0 производную среднего выходного уровня дефектности по среднему входному уровню дефектности:

$$\frac{df_1(p)}{dp} = \left(p(1-p)^n\right)' = (1-p)^n + pn(1-p)^{n-1} =$$

$$= (1-p)^{n-1}(1 - p - pn) = (1-p)^{n-1}(1 - (n+1)p) = 0.$$

В полученном уравнении корень $p = 1$ соответствует минимуму, а не максимуму. Поскольку непрерывная функция на замкнутом отрезке достигает максимума, то максимум достигается при

$$p_n = \frac{1}{n+1}.$$

Следовательно,

$$\text{ПСВУД} = p_n(1-p_n)^n = \frac{1}{n+1}\left(1 - \frac{1}{n+1}\right)^n. \qquad (8)$$

По выражению (8) могут быть проведены конкретные расчеты. Однако оно довольно громоздко. Его можно упростить, используя один замечательный предел из курса математического анализа, а именно:

$$\lim_{n\to\infty}\left(1-\frac{1}{n+1}\right)^{n+1} = e^{-1} = \frac{1}{2{,}718281828...} \approx 0{,}368. \qquad (9)$$

Сравнивая соотношения (8) и (9), видим, что

$$\text{ПСВУД} = \left(\frac{\frac{1}{n+1}}{1-\frac{1}{n+1}}\right)\left(1-\frac{1}{n+1}\right)^{n+1}.$$

Первая скобка равна $1/n$, а вторая согласно соотношению (9) приближается к 0,368 при росте объема выборки. Поэтому получаем простую асимптотическую формулу

$$\text{ПСВУД} \approx \frac{0{,}368}{n}.$$

Для более сложных планов ПСВУД рассчитывают с помощью более или менее сложных компьютерных программ.

При рассмотрении основ статистического контроля в настоящем пункте расчетные формулы удалось получить лишь для простейших планов, в основном для планов вида $(n,0)$. Если ослабить требования и рассчитывать не на точные формулы, а на асимптотические, при $n \to \infty$, то можно справиться и с одноступенчатыми планами вида (n, c).

13.2. Асимптотическая теория одноступенчатых планов статистического контроля

Пусть X - число дефектных единиц продукции в выборке объема n. Как уже отмечалось, распределение X является биномиальным и имеет вид

$$P(X=k) = C_n^k p^k (1-p)^{n-k},$$

где C_n^k - число сочетаний из n элементов по k, а p - входной уровень дефектности.

Пусть используется одноступенчатый план контроля (n, c). Тогда оперативная характеристика этого плана имеет вид

$$f(p) = \sum_{1\le k\le c} P(X=k) = \sum_{1\le k\le c} C_n^k p^k (1-p)^{n-k}.$$

Пусть $n \to \infty$. Тогда по Закону Больших Чисел теории вероятностей (по теореме Бернулли)

$$\frac{X}{n} \to p$$

(сходимость по вероятности). Значит, если c/n окажется заметно меньше входного уровня дефектности p, то партии будут почти всегда приниматься, а если c/n окажется заметно больше входного уровня дефектности p, то партии будут почти всегда отклоняться. Ситуация будет нетривиальной только там, где величины c/n и p близки друг к другу.

Хотя оперативная характеристика приближается с помощью сумм биномиальных вероятностей, целесообразно найти для нее приближение с помощью теоремы Муавра-Лапласа. Имеем цепочку тождественных преобразований:

$$f(p) = P(X \le c) = P\left(\frac{X-np}{\sqrt{np(1-p)}} \le \frac{c-np}{\sqrt{np(1-p)}}\right).$$

Однако справа строит именно то выражение, которое участвует в теореме Муавра-Лапласа. Воспользовавшись равномерной сходимостью в этой теореме, можно записать, что

$$\lim_{n\to\infty} f(p) = \Phi\left(\frac{c-np}{\sqrt{np(1-p)}}\right),$$

где $\Phi(x)$ - функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. Поскольку параметры в этой формуле связаны соотношением

$$p \approx \frac{c}{n},$$

то можно указать альтернативный вариант асимптотического выражения для оперативной характеристики:

$$f(p) \approx \Phi\left(\frac{(c-np)\sqrt{n}}{\sqrt{c(n-c)}}\right).$$

Последняя формула позволяет без труда написать асимптотические выражения для приемочного и браковочного уровней дефектности. Действительно, согласно определениям этих понятий

$$\Phi\left(\frac{(c-np_{пр})\sqrt{n}}{\sqrt{c(n-c)}}\right) = 1-\alpha, \quad \Phi\left(\frac{(c-np_{бр})\sqrt{n}}{\sqrt{c(n-c)}}\right) = \beta,$$

откуда с помощью элементарных преобразований получаем, что

$$p_{пр} = \frac{c}{n} - \frac{1}{\sqrt{n}}\sqrt{\frac{c}{n}\left(1-\frac{c}{n}\right)}\Phi^{-1}(1-\alpha), \quad p_{бр} = \frac{c}{n} - \frac{1}{\sqrt{n}}\sqrt{\frac{c}{n}\left(1-\frac{c}{n}\right)}\Phi^{-1}(\beta).$$

Поскольку при практическом применении статистического приемочного контроля, как уже отмечалось, принимают α= 0,05, β=0,10, то в предыдущие формулы следует подставить $\Phi^{-1}(0{,}95)=1{,}64$ и $\Phi^{-1}(0{,}10)=-1{,}28$. Итак, итоговые формулы для приемочного и браковочного уровней дефектности имеют вид

$$p_{пр} = \frac{c}{n} - \frac{1{,}64}{\sqrt{n}}\sqrt{\frac{c}{n}\left(1-\frac{c}{n}\right)}, \quad p_{бр} = \frac{c}{n} + \frac{1{,}28}{\sqrt{n}}\sqrt{\frac{c}{n}\left(1-\frac{c}{n}\right)}. \qquad (10)$$

Из формул (10) следует, в частности, что

$$\frac{\frac{c}{n} - p_{пр}}{p_{бр} - \frac{c}{n}} = \frac{1{,}64}{1{,}28} = 1{,}28. \qquad (11)$$

Следовательно, оценкой приемочной доли (отношения приемочного числа к объему выборки) является

$$\left(\frac{c}{n}\right)^{*} = \frac{p_{пр} + 1{,}28\,p_{бр}}{2{,}28}. \qquad (12)$$

Из формулы (10) следует, что

$$\sqrt{n}(p_{бр} - p_{пр}) = 2{,}92\sqrt{\frac{c}{n}\left(1-\frac{c}{n}\right)}. \qquad (13)$$

Следовательно, из формул (12) и (13) вытекает способ оценивания необходимого объема выборки:

$$n^* = \frac{8{,}53\left(\frac{c}{n}\right)^*\left(1-\left(\frac{c}{n}\right)^*\right)}{(p_{бр}-p_{пр})^2}. \qquad (14)$$

Итак, по формуле (12) можно рассчитать оценку выборочной доли, затем по формуле (14) - объем выборки, после чего, вернувшись к выборочной доле, найти приемочное число. Необходимо отметить, что результаты расчетов по рассматриваемым асимптотическим формулам отнюдь не всегда дают натуральные числа, поэтому необходима корректировка полученных результатов.

Рассматриваемые формулы позволяют решить сформулированную выше задачу - по заданным приемочному и браковочному уровням дефектности подобрать такой одноступенчатый план контроля, что его оперативная характеристика $f(p)$ удовлетворяет неравенствам

$$f(p_{пр}) \ge 1 - \alpha\ ,\ f(p_{бр}) \le \beta\ .$$

Поэтому при практической работе корректировка асимптотических результатов должна быть направлена на выполнение указанных неравенств.

Пример. Пусть $p_{пр} = 0{,}02$, $p_{бр} = 0{,}09$. Тогда по формуле (12) приемочная доля равна

$$\left(\frac{c}{n}\right)^* = \frac{p_{пр}+1{,}28p_{бр}}{2{,}28} = \frac{0{,}02+1{,}28\times 0{,}09}{2{,}28} = 0{,}0593.$$

Необходимый объем выборки рассчитывается по формуле (14):

$$n^* = \frac{8{,}53\left(\frac{c}{n}\right)^*\left(1-\left(\frac{c}{n}\right)^*\right)}{(p_{бр}-p_{пр})^2} = \frac{8{,}53\times 0{,}0593\times 0{,}9407}{0{,}0049} = 97{,}11.$$

Полученное число не является натуральным, поэтому вполне естественно откорректировать объем выборки до ближайшего целого, т.е. до 97. Тогда

$$c = 0{,}0593\times 97 = 5{,}75.$$

Заменив c на ближайшее натуральное число, получаем в результате асимптотических расчетов одноступенчатый план (97, 6).

13.3. Некоторые практические вопросы статистического контроля качества продукции и услуг

Познакомившись с некоторыми основными понятиями, подходами, и идеями теории статистического контроля качества, обсудим более практические стороны этой технико-экономической области.

Анализ и синтез планов контроля. На основе теории статистического контроля можно проанализировать планы контроля качества, имеющиеся в нормативно-технической документации (стандартах, технических условиях) и в договорах на поставку продукции и оказание услуг. Достаточно часто оказывается, что формулировки соответствующих разделов (разделов "Правила приемки", "Методы контроля" и др.) имеют различные недостатки и неточности, что может послужить в дальнейшем причиной к возникновению арбитражных ситуаций (т.е. решаемых через арбитражные или иные суды).

Если обсуждаемая система контроля качества выдерживает чисто логическую проверку, то наступает вторая стадия - анализ с точки зрения теории статистического контроля. На этой стадии рассчитывают характеристики применяемых планов контроля. О некоторых из них уже шла речь - приемочный и браковочный уровни дефектности, предел среднего выходного уровня дефектности. Есть и иные показатели, например, средний используемый объем выборки, средняя стоимость контроля, и т.п. Особенно важна прогнозируемая доля арбитражных ситуаций (споров между предприятиями) при используемой системе контроля.

На стадии анализа возможны неожиданные "открытия". Например, может оказаться, что существующая система контроля качества, хотя и является формально безупречной, но защищает лишь от партий продукции, в которой более половины единиц продукции дефектно (т.е. для применяемых планов контроля браковочный уровень дефектности больше 0,5). Или что система контроля защищает интересы поставщиков, у которых каждое пятое изделие является бракованным (приемочный уровень дефектности равен 0,2).

Замечание. До сих пор постоянно говорилось о контроле единиц и партий продукции. Однако нет никакого принципиального отличия с контролем услуг (медицинских, туристических, транспортных, образовательных, банковских и иных) или документации. Поэтому теория и практика статистического контроля качества продукции дает полезные рекомендации для банковского дела и бухгалтерского аудита. Надо только аккуратно заменить слова, описывающие предметную область применения теории статистического контроля.

После анализа ситуации с системой контроля естественно перейти к улучшению этой системы, к обоснованному выбору планов, к этапу синтеза. В зависимости от конкретных условий используются разнообразные подходы к выбору планов. Например, задают приемочный и браковочный уровни дефектности. В случае контроля с разбраковкой естественно использовать ограничения на предел среднего выходного уровня дефектности.

Обсудим подробнее оптимизационные постановки в статистическом приемочном контроле. Очевидно, имеется три вида затрат и потерь:

- затраты непосредственно на проведение контроля единиц продукции, включенных в выборку,

- потери в случае неверного решения о забраковании партии продукции (в которой на самом деле доля дефектной продукции соответствует требованиям нормативно-технической документации):

- потери в случае неверного решения о принятии партии продукции (в которой на самом деле доля дефектной продукции не соответствует требованиям нормативно-технической документации).

При этом первые два вида затрат непосредственно связаны с деятельностью предприятия, на котором производится продукция, третий вид затрат формируется там, где она потребляется. С этим связана принципиальная сложность подсчета затрат третьего вида. Особенно эта сложность проявляется тогда, когда попадание к потребителю дефектных изделий может привести к авариям с человеческими жертвами. Тогда в очередной раз возникает уже обсуждавшийся вопрос: сколько стоит человеческая жизнь? Только оценив потери здоровья и жизни в денежных единицах, можно сформировать функционал качества плана статистического контроля и затем оптимизировать его. К счастью, для большинства видов продукции вопрос о денежной оценке человеческой жизни не возникает. Проблема обычно "всего лишь" в том, что выпущенная продукция используется разнообразными конечными потребителями, а потому оценить эффект повышения доли ее дефектности затруднительно.

Поэтому наряду с функционалом качества, включающим все три вида затрат, рассматривают "условный" функционал на основе затрат первых двух типов, а на вероятность принятия партии продукции, в которой доля дефектной продукции не соответствует требованиям нормативно-технической документации, накладывают ограничение, т.е., грубо говоря, третий вид затрат учитывают в качестве ограничения.

Естественно также по-разному проводить контроль у поставщика (производителя) и потребителя (заказчика). Пусть для определенности поставщик используют план $(n_1, 0)$, а потребитель - $(n_2, 0)$. Тогда естественно зафиксировать в договоре о поставке, что $n_1 >> n_2$. Такая договоренность обеспечит тщательный контроль со стороны изготовителя и почти автоматическое подтверждение приемки со стороны потребителя (т.е. отсутствие спора).

Одна из распространенных догм состоит в том, что изготовитель и потребитель должны проводить контроль по одним и тем же планам контроля. Если план контроля и входной уровень контроля таков, что ситуация контроля относится к буферной зоне Б, т.е. вероятность приемки партии заметно отличается от 0 и 1, то указанная догма приводит к высокой вероятности спорных ситуаций. Пусть, например, оперативная характеристика равна 0,5. Пусть изготовитель принял партию (с вероятностью 0,5). После этого при независимом контроле у потребителя с той же вероятностью 0,5 она может быть отклонена и с вероятностью 0,5 принята. Значит, общий итог таков: 59% за то, что партия будет забракована у поставщика, 25% - за спорную ситуацию (поставщик принял, потребитель забраковал), 25% - за принятие и поставщиком и потребителем. Конечно, рассмотрен крайний случай - наиболее частое появление спорных ситуаций. Но реальное появление 10-15% арбитражных споров - это типовая ситуация в 1980-е годы.

Один из вариантов выбора планов контроля поставщиком и потребителем выглядит так. Стороны договариваются о некотором "приемлемом" входном уровне дефектности p^*. Затем поставщик выбирает план контроля, используя p^* как браковочный уровень дефектности, а потребитель - рассматривая p^* как приемочный уровень дефектности. Подробнее об анализе, синтезе и оптимизации планов статистического контроля рассказано в специальной литературе, в частности, в работах [6,8].

Усеченные планы. Рассмотрим план статистического контроля (60, 3). Пусть при проверке единицы продукции появляются в таком порядке: дефектная, дефектная, дефектная, дефектная,… Четыре дефектные единицы подряд! Надо ли дальше проверять выборку? Исходя из здравого смысла - нет. Ведь совершенно неважно, каковы будут результаты по остальным 59-и единицам продукции, окажутся они годными или дефектными - 4 дефектные единицы уже есть, и партию следует забраковать. Контроль мог бы быть перекрашен и тогда, когда при проверке 60 единиц все 60 окажутся годными - независимо от качества остальных 3 партию надо принимать.

Усеченные планы - это планы статистического контроля, в которых контроль разрешается прекращать, если итог (принятие или забракование партии) становится ясен ранее, чем проведен контроль всех включенных в выборку единиц продукции. Усеченные планы применяются, когда единицы продукции поступают на контроль последовательно, одна за другой (или группа за группой). Это не всегда так. Если, например, план (60, 3) применяется для контроля качества электролампочек, и все 63 лампочки ввернуты в гнезда на испытательном стенде и одновременно включены, то подход на основе усеченных планов применить нельзя.

Возможность применения усеченных планов должна быть явным образом указана в нормативно-технической документации и в договорах на поставку. Опишем юридический казус, связанный с усеченными планами. В ГОСТе на штангенциркули был предусмотрен план контроля (20,0). Органы Госстандарта проверяли завод "Точнометр" (название изменено). Проверили первый штангенциркуль - дефектен, второй - дефектен,…, десятый - дефектен. На этом комиссия остановилась, вполне резонно (с точки зрения здравого смысла) решив, что партия штангенциркулей должна быть забракована. Органы Госстандарта наложили на завод "Точнометр" штраф за выпуск некачественной продукции (в соответствии с действующим в то время правопорядком). Однако завод опротестовал это решение в суд. И суд удовлетворил протест, ссылаясь на то, что порядок проведения контроля качества штангенциркулей был нарушен! Бракоделы не смогли бы уйти от наказания, если бы в соответствующих документах была бы прописана возможность использования усеченных планов.

Выделение единиц бесформенной (жидкой, газообразной) продукции. Во всем предыдущем изложении постоянно встречается термин "единица продукции". Он вполне ясен, если речь идет об отдельных изделиях - дискетах, коробках спичек, патронах, бутылках минеральной воды, электробритвах, или отдельных деталях - болтах, гвоздях, пластмассовых дисках… Совершенно ясно, что многие виды продукции имеют иной вид - газообразный, жидкий или, как говорят, бесформенный (порошкообразный, желеобразный,…). Как быть с ним? В работе [9] предложен подход, позволяющий применить к бесформенной продукции методы статистического контроля качества.

Основное - это выделить единицу продукции. Она не должна быть очень малой, поскольку ясно, что в бесформенной продукции свойства вещества в близких точках близки. Основная идея состоит в том, чтобы взять некоторое количество пар точек, отстоящих друг от друга на определенное расстояние, и выяснить, есть связь (т.е. значим ли ранговый коэффициент корреляции Спирмена - см. главу 5) между значениями изучаемого свойства в этих парах точек или нет. Если связь есть, значит, точки разнесены на недостаточное расстояние, другими словами, точки относятся к одной и той же единице продукции. Поэтому расстояние между точками надо увеличить. Если связь уже не обнаруживается, то это значит, что они относятся к разным единицам продукции. В процессе увеличения расстояния тем самым была оценена величина ребра куба, в виде которого условно представляем себе единицу бесформенной продукции. Разбив бесформенную продукцию на единицы, можно применять описанные выше подходы для контроля ее качества (подробнее см. [9]).

Отбор случайной выборки при статистическом контроле качества продукции. Как и при любом выборочном обследовании, при статистическом контроле качества продукции остро строит проблема отбора репрезентативной (представительной) выборки (см. главу 2 выше). Эта проблема усугубляется экономической заинтересованностью участников процесса. В соответствии с обсуждениями главы 11 наиболее научно-обоснованным является использование датчиков псевдослучайных чисел. С другой стороны, исходя из экономической и технической целесообразности, популярна схема многоступенчатой выборки. Например, из 15 вагонов отобрать вагон № 5, из него - контейнер №3 около двери (из 12 контейнеров), из контейнера №3 - ящики №№ 7, 15 и 23, а из этих ящиков - каждое пятое изделие. При этом описании составления выборки совершенно ясно, что реально классическая случайная выборка организуется лишь при контроле контейнера №3, и остается только надеяться, что он является типичным для всей партии.

13.4. Всегда ли нужен контроль качества продукции?

Чем выше достигнутый уровень качества, тем больше необходимый объем контроля - таков парадокс классической теории статистического контроля. Возможный выход состоит в переходе к расширению возможностей менеджера при выборе технической политики на основе учета экономических рисков. Перекладывание контроля на потребителя может быть экономически выгодно, если производитель организовал защиту от риска методом пополнения партий или путем развития технического обслуживания.

В государственных стандартах, технических условиях, другой нормативно-технической документации, относящейся к потребительским товарам и услугам, различным изделиям, веществам, материалам, иным видам продукции, а также в договорах между поставщиками и потребителями обычно присутствуют разделы "Правила приемки и методы контроля". Поэтому, в частности, методы статистического контроля качества продукции являются важной составной частью статистических методов сертификации, которым посвящена работа [4]. Как уже говорилось, имеется соответствующая вероятностно-статистическая теория, посвященная анализу и синтезу (выбору) планов контроля. Однако эта теория вообще не предусматривает отказа от контроля, поскольку игнорирует возможность перехода на иную стратегию организации взаимоотношений поставщика и потребителя, например, на стратегию технического обслуживания, при которой выходной контроль не проводится, а обнаруженные потребителями дефектные изделия заменяются годными или ремонтируются. Основная обсуждаемая в настоящем пункте идея - обоснование необходимости включения теории статистического приемочного контроля в более широкую технико-экономическую теорию взаимоотношений поставщиков и потребителей и целесообразности перехода при повышении качества продукции от контроля качества к иным способам защиты потребителя, например, к развитому техническому обслуживанию или к поставке запасных единиц продукции.

Использование экономических показателей при выборе планов статистического (выборочного) контроля пропагандировалось давно, но делалось это в рамках парадигмы обязательности контроля. Здесь рассматривается более широкая система взглядов, согласно которой контроль качества продукции - лишь один из способов урегулирования взаимоотношений между поставщиками и потребителями.

В более широком плане речь идет об отказе от получения детальной информации, если она стоит слишком дорого, и переходе к использованию иных механизмов управления. Так, качественные методы химического анализа часто используют именно потому, что соответствующие количественные методы более трудоемки и дороги, но не намного полезнее с практической точки зрения. Пример из всем знакомой области: в средней школе знания учащихся контролируются еженедельно, в высшей же - один или несколько раз в семестр, однако разница с точки зрения эффективности управления процессом обучения невелика. Другой пример: как показано в статистике интервальных данных (см. главу 9), из-за погрешностей измерений нецелесообразно увеличивать их число сверх некоторого "рационального объема выборки", а для увеличения точности оценивания характеристик вероятностных распределений необходимо использовать более точные средства измерения. С учетом сказанного описываемый в настоящем пункте подход представляется менее необычным.

Оценка снизу необходимого объема выборки. Как известно, в теории статистического приемочного контроля качества продукции разработано много подходов к выбору планов контроля:

- на основе приемочного и браковочного уровней дефектности;

- исходя из предела среднего выходного уровня дефектности (при контроле с разбраковкой);

- с использованием экономических показателей, относящихся к предприятию (см., например, ГОСТ 24660-81);

- с использованием экономических показателей, относящихся к народному хозяйству в целом; и т.д. (см. предыдущий пункт).

Имеется обширная литература, посвященная обоснованию и сравнению этих подходов, разработке соответствующей математической теории и программного обеспечения. Не углубляясь в эти проблемы, сосредоточим внимание на одном парадоксальном явлении: при повышении качества выпускаемой продукции теория рекомендует увеличивать объем контроля!

Действительно, при повышении качества выпускаемой продукции требования потребителя, очевидно, обеспечиваются все лучше. Следовательно, должен уменьшаться браковочный уровень дефектности, т.е. то значение входного уровня дефектности, при котором вероятность приемки партии равна риску потребителя. Из всех планов с общим объемом контроля n минимум вероятности приемки партии (т.е. оперативной характеристики) достигается на одноступенчатом плане *(n,0)*. (Напомним, что согласно этому плану партия принимается тогда и только тогда, когда из n проверенных единиц продукции все оказываются годными.) Другими словами, оперативная характеристика для плана *(n,0)* является огибающей (снизу) множества всех оперативных характеристик. Следовательно, из всех планов с общим объемом контроля n минимум браковочного уровня дефектности достигается также на плане *(n,0)*.

В дальнейшем будем исходить из биномиальной модели выборки, согласно которой число дефектных единиц продукции в выборке объема n имеет биномиальное распределение с параметрами n и p, где p - входной уровень дефектности. Как хорошо известно, эта модель является приближением для модели простой случайной выборки из партии, согласно которой указанное число имеет гипергеометрическое распределение. Напомним, что гипергеометрическая модель переходит в биномиальную, если объем партии безгранично возрастает, а доля дефектных единиц продукции в партии приближается к p. Если объем выборки составляет не более 10% объема партии, то с достаточной для практики точностью принимают, что соответствующее биномиальное распределение хорошо приближает гипергеометрическое.

Примем обычное предположение о том, что риск потребителя равен 0,10. Как известно, браковочный уровень дефектности $p_{бр}$ для плана *(n,0)* определяется из условия

$$(1 - p_{бр})^n = 0{,}10 \ .$$

Это соотношение дает возможность по заданному браковочному уровню дефектности $p_{бр}$ найти необходимый объем выборки:

$$n = \ln 0{,}10 \ / \ln (1 - p_{бр}) = - 2{,}30 \ / \ln (1 - p_{бр}) \ .$$

Поскольку в силу сказанного ранее представляют интерес малые значения браковочного уровня дефектности, воспользуемся тем, что при малых x согласно правилам математического анализа

$$\ln (1 + x) = x + O(x^2) \ .$$

Вторым слагаемым в правой части последней формулы, как обычно в асимптотических рассуждениях, можно пренебречь. Следовательно, необходимый объем выборки с достаточной точностью может быть найден по формуле

$$n = 2{,}30 \ / \ p_{бр} \ . \qquad (15)$$

(При конкретных расчетах надо, очевидно, правую часть округлить до ближайшего целого числа.) Например, при довольно низком (с точки зрения мирового рынка) качестве выпускаемой продукции можно задать $p_{бр}$ = 0,01, т.е. потребовать, чтобы почти все (точнее, не менее 90%) партии, в которых дефектных единиц больше, чем 1 из 100, были забракованы и не достигли потребителя. Тогда объем контроля должен составлять не менее n = 230.

Основной парадокс теории статистического приемочного контроля. Как следует из сказанного выше, необходимый объем выборки, определяемый для какого-либо плана контроля по заданному браковочному уровню дефектности $p_{бр}$, будет не меньше, чем для плана *(n,0),* т.е. не меньше, чем 2,30 / $p_{бр}$.Таким образом, если достигнут достаточно высокий уровень качества, такой, что потребителю может попасть не более 1 дефектной единицы продукции из 10000, т.е. $p_{бр}$ = 0,0001, то объем контроля должен быть не меньше n = 23000. Если же качество повысится в 100 раз, т.е. потребителю сможет попасть не более 1 дефектной единицы продукции из 1000000, то объем контроля и затраты на него возрастут также в 100 раз, и минимально необходимый объем контроля составит 2,3 миллиона единиц продукции. Поскольку объем партий большинства видов продукции существенно меньше этого числа, то проведенные выше расчеты говорят о необходимости перехода на сплошной контроль.

Итак, выводы парадоксальны: если качество выпускаемой продукции не очень хорошее, то целесообразно проводить статистический (выборочный) контроль, если же качество возрастает, то объем контроля и затраты на него увеличиваются, вплоть до перехода на сплошной контроль. Если это возможно, т.е. контроль не является разрушающим. А если невозможно, то попадаем в тупиковую ситуацию - высокое качество не может быть подтверждено.

В реальных ситуациях объемы контролируемых выборок - единицы или десятки, но обычно отнюдь не сотни и тысячи. Если контролируются 100 изделий, то согласно формуле (15) браковочный уровень дефектности равен 2,3 %. И это - предел для реально используемых объемов контроля. Следовательно, статистический приемочный контроль (в том числе выходной или входной) может быть применен для контроля лишь такой продукции, в которой из 50 изделий хотя бы одно дефектно. Другими словами, этот метод управления качеством предназначен лишь для продукции сравнительно низкого качества (входной уровень дефектности не менее 1-2%) или при обслуживании потребителя, согласного на довольно высокий браковочный уровень дефектности (не менее 2,3%).
Следовательно, для повышения качества необходимо использовать контрольные карты и другие методы статистического регулирования технологических процессов на предприятии (о них подробно рассказано, например, в монографиях [1,10]), методы "всеобщего (в другом переводе - тотального) контроля качества" и др. Недаром этим методам уделяется больше внимания в зарубежных методических изданиях, чем собственно статистическому приемочному контролю.

От контроля к пополнению партии. Рассмотрим простую идею: отказываемся от контроля качества вообще, но зато по первому требованию потребителя заменяем дефектную единицу продукции на новую. При этом экономим на контроле, но вместо этого тратим средства на замену продукции. Выгодно это или не выгодно?

Замена продукции может проводиться различными способами. Для многих видов товаров народного потребления это делается с помощью системы гарантийного обслуживания, гарантийных сроков и мастерских, через сеть розничной торговли и т.д.

Другой вариант - к партии поставляемой продукции добавляется некоторое количество единиц продукции для замены имеющихся, возможно, в ней дефектных единиц. Сначала обсудим подробнее именно этот вариант идеи замены продукции.

Пусть поставщик выпускает продукцию с известным ему уровнем дефектности p. Тогда число X дефектных единиц в партии объема N имеет биномиальное распределение с параметрами N и p. По теореме Муавра-Лапласа X не превосходит (при достаточно большом N) величины

$$D_0(t) = Np + t\,(Np(1-p))^{1/2}$$

с вероятностью Ф(t). где Ф(.) - функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. Поскольку Ф(4) = 0,999968329, то для практических целей достаточно положить t = 4, при этом более чем $D_0(4)$ дефектных единиц продукции попадет в партию лишь в 3 случаях из 100000.

Пусть C_0 - цена одной единицы продукции, C_1 - стоимость неразрушающего контроля одной единицы продукции (с исправлением дефектов при их обнаружении). Сравним сначала две стратегии технико-экономических отношений поставщика с потребителями:

сплошной контроль (затраты C_1N)

и пополнение партии дополнительными изделиями в числе $D_0(4)$ (затраты $C_0D_0(4)$). Вторая стратегия лучше (экономически выгоднее), если

$$C_1N > C_0D_0(4) = C_0(Np + 4\sqrt{Np(1-p)}). \qquad (16)$$

Поделим на C_0N, получим равносильное неравенство

$$\frac{C_1}{C_0} > p + 4\frac{\sqrt{p(1-p)}}{\sqrt{N}}.$$

Поскольку $p(1-p)$ не превосходит 1/4 при всех p, то из неравенства

$$C_1/C_0 > p + 2 / N^{1/2} \qquad (17)$$

вытекает неравенство (16). Ясно, что в случае, если

$$C_1/C_0 > p\ ,$$

неравенство (17) (а потому и неравенство (16)) выполняется при достаточно больших объемах партии, а именно, при

$$N > \{2\ C_0 / (C_1 - C_0\ p)\}^2 .$$

Например, если стоимость контроля составляет 10% от стоимости продукции (типовая ситуация в машиностроении), т.е. $C_1/C_0 = 0{,}1$, а уровень дефектности $p = 0{,}01$, то последнее неравенство дает N>493. В то же время нетрудно проверить, что неравенство (16) выполняется при

$$0{,}1 > 0.01 + 4\ (0.01*0{,}99)^{1/2} / N^{1/2}\ ,$$

т.е. при $N > 19$. Расхождение более чем на порядок (в 26 раз) объясняется заменой при переходе от формулы (16) к формуле (17) величины $p(1-p)$ на 1/4, т.е. на гораздо большую величину - при малом входном уровне дефектности p.

Выгодно ли введение статистического контроля? Пусть рассматривается описанная выше стратегия пополнения партий. Мы сравнивали ее со стратегией сплошного контроля, которая во многих случаях оказалась хуже. Может быть, поставщику имеет смысл использовать статистический контроль? Понятно, что речь может идти лишь о (неразрушающем) контроле с разбраковкой, поскольку только в этом случае меняется доля дефектности в потоке партий, направляемых потребителям.

Пусть используется план $(n,0)$ с приемочным уровнем дефектности, равным реально достигнутому предприятием уровню дефектности p. Как известно, тогда объем выборки определяется из условия

$$(1\text{-}p)^n = 0{,}95,$$

т.е.

$$n = \ln 0{,}95 / \ln (1 - p) = - 0{,}0513\ / \ln (1 - p)\ .$$

При малом p уже не раз применявшееся соотношение из математического анализа дает с достаточной для практики точностью

$$n = 0{,}05 / p\ .$$

С вероятностью $(1\text{-}p)^n = 0{,}95$ партия принимается, с вероятностью 0,05 подвергается разбраковке. В первом случае партия поступает к потребителю с тем же уровнем дефектности, что и до контроля, но при этом добавляются затраты на контроль, равные C_1n. Партию необходимо пополнить $D_0(4)$ изделиями (затраты $C_0D_0(4)$), общие затраты (в среднем на одну выпущенную партию) равны

$$0{,}95\ (C_1n + C_0\ D_0\ (4))\ .$$

Во втором случае фактически проводится сплошной контроль с исправлением дефектов и затратами C_1N. Суммарные затраты при использовании выборочного контроля равны

$$0{,}95\ (C_1n + C_0\, D_0\,(4))\ + 0{,}05\ C_1N\,.$$

Он более выгоден, чем отсутствие контроля (с добавлением "запасных" изделий), в случае справедливости неравенства

$$0{,}95\ (C_1n + C_0\, D_0\,(4))\ + 0{,}05\ C_1N\ < C_0\, D_0\,(4),$$

что эквивалентно неравенству

$$19\ C_1n\ + C_1N < C_0\, D_0\,(4).$$

Сравнение с формулой (16) показывает, что если контроль не является разрушающим, то выборочный контроль менее выгоден, чем сплошной (по сравнению с формулой (16) добавляется первое слагаемое в левой части последней формулы), и тем более весьма проигрывает в экономической эффективности по сравнению с отсутствием контроля в сочетании с пополнением партии.

Итак, введение статистического контроля в схеме пополнения партии не выгодно.

От системы контроля к системе технического обслуживания. Вернемся к первому из указанных ранее вариантов замены продукции. Что выгоднее - сплошной контроль на предприятии или замена дефектных изделий, обнаруженных потребителями? Реальное перекладывание контроля на потребителей влечет потери, связанные с удовлетворением их претензий, но при малой доле дефектных изделий эти потери малы по сравнению с затратами на контроль.

Действительно, пусть W - средние потери поставщика, связанные с пропуском потребителю дефектной единицы продукции. Сюда входят, в частности, такие виды потерь:

- стоимость новой единицы продукции (при замене изделия или возврате его стоимости);

- расходы системы распределения продукции и гарантийного ремонта, включая издержки на устранение дефектов;

- потери из-за нежелательного изменения предпочтений потребителя, из-за снижения имиджа фирмы;

- затраты на возмещение ущерба, понесенного потребителем, страховые сборы, судебные издержки, и т.д.

Потери W в несколько раз (по экспертной оценке - обычно в 5-10 раз) превышают расходы C_0 на изготовление единицы продукции. Кроме того, для быстрого решения проблем потребителей, связанных с обнаружением дефектов, необходима развитая система технического обслуживания.

Пусть изготовлена партия продукции объема N. Тогда расходы на сплошной (неразрушающий) контроль составляют C_1N (при этом дефектные единицы продукции извлекаются и утилизируются, расходами на утилизацию или доходами от нее в настоящем изложении пренебрегаем). Пусть p - доля дефектных единиц продукции в партии. Тогда Np - математическое ожидание числа дефектных единиц продукции в партии, а WNp - математическое ожидание потерь. Если

$$WNp < C_1\, N, \quad p < C_1 / W, \quad (18)$$

то выгоднее отказаться от сплошного контроля. При повышении качества, т.е. снижении доли дефектности, целесообразно переходить к поиску и устранению дефектов не непосредственно на предприятии, а в пунктах системы технического обслуживания.

В формуле (18) участвует математическое ожидание WNp. Реальные потери могут быть больше, но не намного. Как и выше, с помощью теоремы Муавра-Лапласа можно утверждать, что практически наверняка они не превышают $WD_0(4)$, а потому преимущество решения об отказе от контроля неоспоримо при

$$WD_0(4) < C_1N, \quad p + 4\,(p(1-p))^{1/2} / N^{1/2} < C_1 / W. \qquad (19)$$

Аналогично выводу неравенства (17) заключаем, что неравенство (19) наверняка будет выполнено, если

$$p + 2 / N^{1/2} < C_1 / W. \quad (20)$$

Пусть $C_1 / W = 0{,}1$, выпускается партия объема $N = 1600$. Тогда согласно неравенству (20) отказ от контроля выгоден уже при $p < 0{,}05$, т.е. граничное значение соответствует довольно низкому уровню качества - 1 единица продукции из 20.

Выгодно ли в рассматриваемой ситуации вводить выборочный контроль? Пусть объем контроля равен n, приемочное число $c = 0$, с вероятностью y партия принимается, а с вероятностью $1 - y$ бракуется (и затем подвергается разбраковке). В первом случае расходы на контроль равны C_1n, а остальная часть партии содержит в среднем $(N - n)\,p$ дефектных единиц продукции, и средние издержки равны $y\{C_1n + W(N - n)p\}$. Во втором случае суммарные затраты равны $(1 - y)C_1N$. Следовательно, введение контроля выгодно, если

$$y\{C_1n + W(N - n)p\} + (1 - y)C_1\,N < WNp\,.$$

Преобразуем это неравенство к виду

$$yn\{C_1 - Wp\}(1 - y)^{-1} + C_1N < WNp. \quad (21)$$

Если выполнено неравенство $p<C_1/W$, то второе слагаемое в левой части неравенства (21) больше правой части этого неравенства, в то время как первое слагаемое в левой части (21) положительно. Следовательно, неравенство (21) неверно, и введение выборочного контроля нецелесообразно - как и в разобранном ранее случае метода пополнения партий.

Выше приведен базовый (простейший, исходный) метод сравнения различных систем взаимоотношений поставщиков и потребителей. Целесообразно дальнейшее его развитие, которое предоставляем читателю.

Отметим в заключение, что реально статистический контроль качества продукции, осуществляемый поставщиком (выходной контроль), решает две основные задачи: обеспечение интересов потребителя и обнаружение разладок собственных технологических процессов (по результатам контроля последовательности партий). Как показано выше, для решения первой из этих задач он не всегда оптимален. Вторую из названных задач также часто эффективнее решать с помощью иных методов, например, обнаруживать разладку технологических процессов с помощью тех или иных контрольных карт. Таким образом, область применения методов статистического приемочного контроля является довольно ограниченной. Очевидно, однако, что нельзя исключать эти методы из арсенала менеджеров по качеству, в частности, при использовании концепции "всеобщего управления качеством (*TQM - Total Quality Management)"*. Хотя бы потому, что они незаменимы при использовании разрушающих методов контроля.

Наиболее перспективным представляется использование результатов настоящего пункта в рамках концепции контроллинга - современной концепции системного управления организацией, в основе которой лежит стремление обеспечить ее долгосрочное эффективное существование (см., например, [11-13]).

Итак, в настоящем пункте сформулирован основной парадокс теории статистического приемочного контроля - повышение качества выпускаемой продукции приводит к увеличению объема контроля. Описан способ разрешения этого парадокса на основе перехода от чисто технической политики выбора плана контроля к технико-экономической, основанной на сравнении по экономическим показателям схем контроля и схем технического обслуживания и пополнения партий. Проанализирован базовый метод такого сравнения, позволяющий выделить область экономического преимущества схемы пополнения партий и схемы технического обслуживания по сравнению со схемой контроля.

13.5. Статистический контроль по двум альтернативным признакам и метод проверки их независимости по совокупности малых выборок

В настоящем пункте рассмотрим статистический приемочный контроль по двум альтернативным признакам одновременно. Обсуждается соотношение входного уровня дефектности изделия в целом с входными уровнями дефектности отдельных контролируемых параметров. На основе результатов статистики объектов нечисловой природы (глава 8) рассмотрен метод проверки независимости двух альтернативных признаков. Метод нацелен на применение прежде всего в задачах статистического контроля качества продукции. При этом проверка независимости проводится по совокупности малых выборок, т.е. в так называемой асимптотике А.Н.Колмогорова, когда число неизвестных параметров распределения не является постоянным, а растет пропорционально объему данных.

При статистическом контроле качества продукции, в частности, при сертификации, чаще всего используют контроль по альтернативным признакам. При этом устанавливается, соответствует ли контролируемый параметр единицы продукции (изделия, детали) заданным в нормативно-технической документации требованиям или не соответствует. Если соответствует - единица продукции признается годной. Примем для определенности, что в этом случае результат контроля кодируется символом 0. Если же не соответствует - единица продукции признается дефектной, а результат контроля кодируется символом 1.

Таким образом, в рассматриваемой нами математической модели контроля альтернативный признак - это функция $X = X(w)$, определенная на множестве единиц продукции $W = \{w\}$ и принимающая два значения 0 и 1, причем $X(w) = 0$ означает, что единица продукции w является годной, а $X(w) = 1$ - что она является дефектной.

Методы статистического контроля, в частности, включенные в государственные стандарты и иную нормативно-техническую документацию (НТД), как правило, используют контроль по одному признаку. В НТД указывают правила выбора планов контроля и расчета различных их характеристик, приводят графики оперативных характеристик и т.п.

Однако на производстве контроль нередко проводится по нескольким альтернативным признакам. Возникает проблема выбора плана контроля и расчета его характеристик. В настоящее время для решения этой проблемы нет достаточно обоснованных и общепринятых рекомендаций.

Рассмотрим сначала контроль по двум альтернативным признакам $X(w)$ и $Y(w)$. В вероятностной модели $X(w)$ и $Y(w)$ - случайные величины, принимающие два значения - 0 и 1. Пусть, пользуясь стандартной терминологией,

$$p_1 = P\ (X(w) = 1)$$

- входной уровень дефектности для первого признака, а

$$p_2 = P\ (Y(w) = 1)$$

- для второго. Вероятности результатов контроля по двум признакам одновременно описываются четырьмя числами:

$$P\ (X(w) = 0,\ Y(w) = 0) = p_{00}\ ,\ P\ (X(w) = 1,\ Y(w) = 0) = p_{10}\ ,$$

$$P\ (X(w) = 0,\ Y(w) = 1) = p_{01}\ ,\ \ P\ (X(w) = 1,\ Y(w) = 1) = p_{11}\ ,$$

при этом справедливы соотношения:

$$p_{00} + p_{10} + p_{01} + p_{11} = 1,\ \ p_{10}\ + p_{11} =\ p_1\ ,\ p_{01} + p_{11} = p_2\ .$$

С прикладной точки зрения наиболее интересна вероятность p_{00} того, что единица продукции является годной (по всем параметрам), и вероятность ее дефектности (1-p_{00}), т.е. входной уровень дефектности для изделия в целом.

В табл.1 сведены вместе введенные выше вероятности.

Табл. 1. Вероятности результаты испытаний
при контроле по двум альтернативным признакам

	X=0	X=1	Всего
Y=0	p_{00}	p_{10}	$1-p_2$
Y=1	p_{01}	p_{11}	p_2
Всего	$1-p_1$	p_1	1

Есть три важных частных случая - поглощения, несовместности и независимости дефектов, другими словами, поглощения, несовместности и независимости событий *{w: X(w) = 1} и {w: Y(w) = 1}*. В случае поглощения одно из этих событий содержит другое, а потому

$$p_{00} = 1 - max\ (p_1 , p_2)\ .$$

В случае несовместности

$$p_{00} = 1 - p_1 - p_2\ .$$

В случае независимости

$$p_{00} = (1 - p_1)(1 - p_2) = 1 - p_1 - p_2 + p_1p_2\ .$$

Ояевидно, что вероятность годности изделия всегда заключена между значениями, соответствующими случаям поглощения и несовместности. Кроме того, известно, что при большом числе признаков и малой вероятности дефектности по каждому из них случаи поглощения и независимости дают (в асимптотике) крайние значения для вероятности годности изделия, т.е. формулы, соответствующие независимости и несовместности, асимптотически совпадают.

Рассмотрим несколько примеров. Пусть некоторая продукция, скажем, гвозди, контролируются по двум альтернативным признакам, для определенности, по весу и длине. Пусть результаты контроля 1000 единиц продукции представлены в табл.2

Табл. 2. Результаты 1000 испытаний
по двум альтернативным признакам (случай поглощения)

	X=0	X=1	Всего
У=0	952	0	952
У=1	0	48	48
Всего	952	48	1000

Судя по данным табл.2, дефекты всегда встречаются парами - если есть один, то есть и другой. Входной уровень дефектности как по каждому показателю, так и по обоим вместе - один и тот же, а именно, 0,048. Получив по результатам статистического наблюдения данные типа приведенных в табл.2, целесообразно перейти к контролю только одного показателя, а не двух. Каково именно? Видимо, того, контроль которого дешевле. Однако совсем иная ситуация в случае несовместности дефектов (табл.3).

Табл. 3. Результаты 1000 испытаний
по двум альтернативным признакам (случай несовместности)

	X=0	X=1	Всего
У=0	904	48	952
У=1	48	0	48
Всего	952	48	1000

Судя по данным табл.3, дефекты всегда встречаются поодиночке - если есть один, то другого нет. В результате входной уровень дефектности по каждому признаку по-прежнему равен 0,048, в то время как доля дефектных изделий (т.е. имеющих хотя бы один дефект) вдвое выше, т.е. входной уровень дефектности для изделия в целом равен 0,096.

Случай независимости результатов контроля по двум независимым признакам (табл.4) лежит между крайними случаями поглощения и несовместности. Независимость альтернативных признаков обосновывается путем статистической проверки с помощью описанного ниже критерия $n^{1/2}V$, значение которого для данных табл.4 равно 1,866.

Табл. 4. Результаты 1000 испытаний
по двум альтернативным признакам (случай независимости)

	X=0	X=1	Всего
У=0	909	43	952
У=1	43	5	48
Всего	952	48	1000

Согласно данным табл.4, входной уровень дефектности для каждого из двух альтернативных признаков по-прежнему равен 0,048, в то время как для изделий в целом он равен 0,091, т.е. на 5,5% меньше, чем в случае несовместности, и на 47% больше, чем в случае поглощения.

Проблема состоит в том, что таблицы и стандарты по статистическому приемочному контролю относятся обычно к случаю одного контролируемого параметра. А как быть, если контролируемых параметров несколько? Приведенные выше примеры показывают, что входной уровень дефектности изделия в целом не определяется однозначно по входным уровням дефектности отдельных его параметров.

Как должны соотноситься характеристики планов контроля по отдельным признакам с характеристиками плана контроля по двум (или многим) признакам одновременно? Рассмотрим распространенную рекомендацию - складывать уровни дефектности, т.е. считать, что уровень дефектности изделия в целом равен сумме уровней дефектности по отдельным его параметрам. Она, очевидно, опирается на гипотезу несовместности дефектов, а потому во многих случаях преувеличивает дефектность, а потому ведет к использованию излишне жестких планов контроля, что экономически невыгодно.

Зная специфику применяемых технологических процессов, в ряде конкретных случаев можно предположить, что дефекты по различным признакам возникают независимо друг от друга. Это предположение необходимо обосновывать по статистическим данным. Если же оно обосновано, следует рассчитывать входной уровень дефектности по формуле

$$1 - p_{00} = p_1 + p_2 - p_1 p_2 ,$$

соответствующей независимости признаков.

Итак, необходимо уметь проверять по статистическим данным гипотезу независимости двух альтернативных признаков. Речь идет о статистической проверке нулевой гипотезы

$$H_0:\ p_{11} = p_1 p_2 \qquad (22)$$

(что эквивалентно проверке равенства $p_{00} = (1 - p_1)(1 - p_2)$). Нетрудно проверить, что гипотеза о справедливости равенства (22) эквивалентна гипотезе

$$H_0 : p_{00} p_{11} - p_{10} p_{01} = 0. \qquad (23)$$

В простейшем случае предполагается, что проведено *n* независимых испытаний (X_i , Y_i), $i = 1,2,...,n$, в каждом из которых проконтролированы два альтернативных признака, а вероятности результатов контроля не меняются от испытания к испытанию. Общий вид статистических данных приведен в табл.5.

Табл. 5. Общий вид результатов контроля
по двум альтернативным признакам.

	X=0	X=1	Всего
У=0	*a*	*b*	*a+b*
У=1	*c*	*d*	*c+d*
Всего	*a+c*	*b+d*	*n*

В табл.5 величина a - число испытаний, в которых $(Xi\ ,\ Y_i)$ = (0,0), величина b - число испытаний, в которых $(Xi\ ,\ Y_i)$ = (1,0), и т.д.

Случайный вектор $(a,\ b,\ c,\ d)$ имеет мультиномиальное распределение с числом испытаний n и вектором вероятностей исходов $(p_{00}\ ,\ p_{10}\ ,\ p_{01}\ ,\ p_{11}\ )$. Состоятельными оценками этих вероятностей являются дроби $a/n,\ b/n,\ c/n,\ d/n$ соответственно. Следовательно, критерий проверки гипотезы (23) может быть основан на статистике

$$Z = ad - bc\ . \qquad (24)$$

Как вытекает из известной формулы для ковариаций мультиномиального вектора (см., например, формулу (6.3.5) в учебнике С.Уилкса [14] на с. 153),

$$M(Z) = n\ (p_{10}\,p_{01}\ - p_{00}\,p_{11}), \quad (25)$$

что равно 0 при справедливости гипотезы независимости (23).

Связь между переменными X и Y обычно измеряется коэффициентом, отличающимся от Z нормирующим множителем:

$$V = (ad - bc)\ \{\ (a + b)(a + c)(b + d)(c + d)\ \}^{-\ 1/2} \qquad (26)$$

(см. классическую монографию М. Дж. Кендалла и А. Стьюарта [15, с.723], на которую уже были ссылки, в частности, в главе 5). При справедливости гипотезы H_0 и больших n случайная величина nV^2 имеет хи-квадрат распределение с одной степенью свободы, а $n^{1/2}V$ имеет стандартное нормальное распределение с математическим ожиданием 0 и дисперсией 1 (см. [15, с.736]).

Рассмотрим еще один пример. Пусть проведено 100 испытаний, результаты которых описаны в табл.6. Тогда

$$V = (50 \cdot 20 - 10 \cdot 20)\ (60 \cdot 70 \cdot 30 \cdot 40)^{-1/2} =$$
$$= (1000 - 200) \cdot 5940000^{-1/2} = 800\ /\ 2245 = 0{,}35635,$$
$$n^{1/2}V = 3{,}5635\ \ .$$

Табл. 6. Результаты 100 испытаний
по двум альтернативным признакам.

	X=0	X=1	Всего
У=0	50	10	60
У=1	20	20	40
Всего	70	30	100

Поскольку полученное значение $n^{1/2}V$ превышает критическое значение при любом применяемом в статистике уровне значимости, то гипотезу о независимости признаков необходимо отклонить.

К сожалению, приведенный простой метод годится не всегда. При статистическом анализе реальных данных возникают проблемы, связанные с отсутствием достаточно больших однородных выборок, т.е. выборок, в которых постоянны параметры вероятностных распределений. Реально единицы продукции представляются на контроль партиями, из каждой партии контролируются лишь несколько изделий, т.е. малая выборка. При этом от партии к партии меняются параметры p_{00}, p_{10}, p_{01}, p_{11}, описывающие уровень дефектности. Поэтому необходимы статистические методы, позволяющие проверять гипотезу независимости признаков по совокупности малых выборок. Построим один из возможных методов.

Рассмотрим вероятностную модель совокупности k малых выборок объемов n_1 , n_2 ,..., n_k соответственно. Пусть j -я выборка (X_{jt}, Y_{jt}), $t = 1, 2,..., n_j$, имеет распределение, задаваемое вектором параметров $(p_{00j}, p_{10j}, p_{01j}, p_{11j})$ в соответствии с ранее введенными обозначениями, $j = 1,2,...,k$. Будем проверять гипотезу

$$H_0: \ p_{11j} = (p_{10j} + p_{11j})(p_{01j} + p_{11j}),\ j = 1,2,...,k, \qquad (27)$$

или в эквивалентной формулировке

$$H_0: \ p_{11j}\, p_{00j} - p_{10j}\, p_{01j}\ ,\ j = 1,2,...,k\ . \qquad (28)$$

Основная идея состоит в нахождении асимптотического распределения статистики типа $n^{1/2}V$ при росте числа k малых выборок, а именно, статистики

$$S = g_1 Z_1 + g_2 Z_2 + ... + g_k Z_k\ , \qquad (29)$$

где Z_1 , Z_2 ,..., Z_k - статистики, рассчитанные по формуле (24) для каждой из k выборок, т.е. $Z_j = a_j d_j - b_j c_j$, $j = 1,2,...,k,$ а g_1 , g_2 , ... , g_k - некоторые весовые коэффициенты, которые, в частности, могут совпадать. Поскольку

$$M(S) = g_1 M(Z_1) + g_2 M(Z_2) + ... + g_k M(Z_k), \qquad (30)$$

то при справедливости гипотезы независимости (27) - (28) имеем $M(S) = 0$ согласно соотношению (25). Поскольку слагаемые в сумме (29) независимы, то при росте k случайная величина S в силу Центральной Предельной Теоремы является асимптотически нормальной. Дисперсия этой величины равна сумме дисперсий слагаемых:

$$D(S) = g_1^2 D(Z_1) + g_2^2 D(Z_2) + ... + g_k^2 D(Z_k)\ . \qquad (31)$$

Для оценивания дисперсии S необходимо использовать **несмещенные** оценки дисперсий в каждой из k выборок (и в этом одна из основных "изюминок" разбираемого метода). Предположим, что построены статистики T_j такие, что

$$M(T_j) = D(Z_j)\ ,\ j = 1,2,...,k\ . \qquad (32)$$

Тогда при некоторых математических "условиях регулярности", на которых нет необходимости здесь останавливаться, несмещенная оценка дисперсии статистики S, имеющая согласно формулам (31) и (32) вид

$$L = g_1^2 T_1 + g_2^2 T_2 + ... + g_k^2 T_k\ , \qquad (33)$$

в силу закона больших чисел такова, что дробь $D(S) / L$ приближается к 1 при росте числа выборок (сходимость по вероятности). Отсюда следует, что распределение случайной величины $Q = S L^{-1/2}$ приближается при росте числа выборок к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1. Следовательно, критерий проверки гипотезы (27) - (28) независимости признаков, состоящий в том, что при $-1,96 < Q < 1,96$ гипотеза принимается, а при Q , выходящих за пределы интервала (- 1,96; 1,96) , гипотеза отклоняется, имеет уровень значимости, приближающийся к 0,05 при росте числа выборок. Мощность этого критерия зависит от величины $M(S)D(S)^{-1/2}$ при альтернативе.

Для реализации намеченного плана осталось научиться несмещенно оценивать $D(Z_j)$. К сожалению, в литературе по несмещенному оцениванию не рассматривают случай мультиномиального распределения, поэтому кратко опишем процедуру построения несмещенной оценки $D(Z_j)$. Поскольку согласно формулам (24) и (25)

$$D(Z_j) = M(Z_j^2) - (M(Z_j))^2 = M (a_j^2 d_j^2) - 2 M (a_j b_j c_j d_j) +$$
$$+ M (b_j^2 c_j^2) + n_j^2 (p_{11j}\, p_{00j} - p_{10j}\, p_{01j})^2, \quad (34)$$

то для вычисления $D(Z_j)$ достаточно найти входящие в правую часть формулы (34) начальные смешанные моменты мультиномиального распределения (четвертого порядка). Теоретически это просто - известен вид характеристической функции мультиномиального распределения (см., например, формулу (6.3.4) в монографии [14, с.152]), а начальные смешанные моменты равны значениям ее соответствующих производных в 0, деленным на нужную степень мнимой единицы (формула (5.2.3) в монографии [4, с.131]). Например, с помощью описанной процедуры после некоторых вычислений получаем, что (для упрощения записи здесь и далее опустим индекс j)

$$M (a^2 d^2) = n(n-1)(n-2)(n-3)p_{11}{}^2 p_{00}{}^2 + n(n-1)(n-2)(p_{11}{}^2 p_{00} +$$
$$+ p_{11}\, p_{00}{}^2) + n(n-1)p_{11}\, p_{00}\, . \quad (35)$$

Формула (35) показывает, что начальные смешанные моменты мультиномиального распределения являются многочленами от параметров p_{11}, p_{00}, p_{10}, p_{01} этого распределения, однако конкретный вид этих многочленов достаточно громоздок, поэтому не будем их здесь выписывать, ограничившись формулой (35) в качестве образца.

Как вытекает из формул (34) и (35), для построения несмещенной оценки $D(Z_j)$ достаточно научиться несмещенно оценивать произведения типа $p_{11}{}^r p_{00}{}^m$, где целые неотрицательные числа r, m не превосходят 2. Эта задача решается, начиная с меньших степеней. Известно, что для ковариации мультиномиального вектора

$$M (ad) = - n\, p_{00}\, p_{11} \quad (36)$$

(см., например, формулу (6.3.5) в монографии [14, с.153]), а потому несмещенной оценкой для $p_{00}\, p_{11}$ является ($- ad / n$). Далее, поскольку справедлива аналогичная (35) формула

$$M(a^2 d) = n(n-1)(n-2)\, p_{11} p_{00}{}^2 + n(n-1)p_{11} p_{00}\, , \quad (37)$$

то с помощью формулы (36) преобразуем формулу (37) к виду

$$M(a^2 d + (n-1)ad) = n(n-1)(n-2)p_{11} p_{00}{}^2\, , \quad (38)$$

т.е. несмещенной оценкой $p_{11} p_{00}{}^2$ является $ad(a + n-1)\{n(n-1)(n-2)\}^{-1}$.

Следующий шаг - аналогичным образом с помощью формул (36) и (38) получаем несмещенную оценку для $p_{11}{}^2 p_{00}{}^2$, а затем и для $D(Z_j)$. Промежуточные формулы опущены из-за громоздкости. Окончательный результат таков:

$$T_j = (b_j + d_j)(c_j + d_j)(a_j + c_j)(a_j + b_j)(n-1)^{-1}\, . \quad (39)$$

Как легко видеть,

$$Z_j / T_j^{-1/2} = (n_j -1)^{1/2}\, V_j \, ,$$

т.е. в случае одной выборки предлагаемый метод совпадает с классическим.

Общая идея рассматриваемого метода проверки гипотез по совокупности малых выборок состоит в том, что подбирается статистика, математическое ожидание которой для каждой малой выборки равно 0 при справедливости проверяемой гипотезы. Затем для каждой выборки строится несмещенная оценка дисперсии этой статистики. Итоговая статистика критерия для проверки гипотезы - это сумма рассматриваемых статистик для всех малых выборок, деленная на квадратный корень из суммы всех несмещенных оценок дисперсий рассматриваемых статистик. При справедливости нулевой гипотезы эта итоговая статистика имеет в асимптотике стандартное нормальное распределение (при выполнении некоторых математических "условий регулярности", которые обычно выполняются при анализе реальных статистических данных).

Впервые такой способ проверки гипотез по совокупности малых выборок был предложен в монографии [16, раздел 4.5]. Нестандартность постановки состоит в том, что число неизвестных параметров растет пропорционально объему данных, т.е. имеет место т.н. "асимптотика Колмогорова", или асимптотика растущей размерности. Дальнейшее развитие применительно к данных типа "да"-"нет" (или "годен" - "дефектен") шло в рамках теории люсианов как части статистики объектов нечисловой природы (см. главу 8).

13.6. Эконометрика качества и сертификация

Как уже отмечалось, вслед за экономически развитыми странами в России намечается всё расширяющаяся тенденция к сертификации продукции, т.е. к официальной гарантии поставки производителем продукции, удовлетворяющей установленным требованиям. Поставщики и продавцы должны иметь сертификаты качества на предлагаемые ими товары и услуги. Маркетинг, т.е. производственная и коммерческая политика, нацеленная на получение максимальной прибыли на основе изучения рынка, создания конкурентоспособной продукции и ее полной реализации, включает в себя работы по сертификации.

Не будем останавливаться на быстро меняющейся организационной стороне процесса сертификации и соответствующих отечественных и зарубежных нормативных документах, а также на различных системах сертификации. Как общие проблемы сертификации, так и выбор схемы сертификации для конкретной продукции активно обсуждаются в печати. Приведем лишь несколько замечаний, необходимых для дальнейшего изложения.

Напомним, что, говоря о сертификации продукции, могут иметь в виду качество конкретной ее партии. В ряде случаев это оправдано - рядового потребителя интересует качество лишь той единицы продукции, которую он приобрел. Однако установление долговременных хозяйственных связей целесообразно лишь в случае, когда поставщик гарантирует высокое качество не одной, а всех партий своей продукции. Другими словами, должны быть проведены оценка и сертификация технологических процессов и производств.

Еще больше повышается доверие к поставщику, если не только отдельные технологические процессы, но и всё предприятие в целом гарантированно выпускает продукцию высокого качества. Это обеспечивается действующей на предприятии системой качества, удовлетворяющей требованиям Международной организации по стандартизации ИСО.

В условиях рыночной экономики одна из основных характеристик товара - его конкурентоспособность. Очевидно, производителю необходимо уметь оценивать конкурентоспособность перед запуском продукции в производство или началом работы по продвижению на зарубежный рынок (подробнее см. рекомендации [2]). Следует отметить, что в литературе имеются различные мнения по поводу понятия "конкурентоспособность". В частности, нельзя согласиться с крайне упрощенным подходом в учебнике [17], в котором конкурентоспособность сводится к соотношению цен на внутреннем и внешнем рынках. Достаточно напомнить о таких приемах конкурентной борьбы, как демпинг и (недобросовестная) реклама, таможенные пошлины и квоты.

Одним из основных компонентов конкурентоспособности продукции является ее технический уровень. В западных учебниках справедливо отмечают, что фирма, обладающая патентом или новой научно-технической разработкой, имеет более высокий “излишек производителя” по сравнению с другими фирмами (см., например, учебники [18-20]). В частности, согласно одному из наиболее популярных западных учебников [21] при выборе направления инвестиционных вложений одна из основных учитываемых характеристик - технический уровень продукции.

Из сказанного вытекает, что сертификация материалов и других видов продукции - это современная форма управления качеством продукции. На Западе общепринято, что основная составляющая в управлении качеством продукции - это статистические методы (см., например, отчет Комитета ИСО по изучению принципов стандартизации [3]). В нашей стране внедрение комплексных систем управления качеством (КС УКП), к сожалению, сводилось во многом всего лишь к подготовке документации организационного характера. Статистические методы использовались в промышленности недостаточно, прежде всего из-за недостаточной подготовки кадров, а государственные стандарты по этой тематике зачастую содержали грубейшие ошибки (см. ниже). Ситуация в области применения статистических методов и причины нашего отставания достаточно подробно разобраны в публикациях [4, 22].

За новыми терминами зачастую скрываются хорошо известные понятия, несколько модернизированные в соответствии с современной обстановкой. Так, "маркетинг в широком смысле - это усовершенствованная, ориентированная на рыночную экономику КС УКП" (см. рекомендации [2, с.61]). Другими словами, идея КС УКП была хороша, а вот ее реализация…

Подготовка предприятий к сертификации продукции, технологических процессов и производств, систем качества требует приложения труда квалифицированных специалистов, причем в достаточно большом объеме. Подобную работу обычно проводят специализированные организации на основе системы методических материалов, охватывающих все стороны подготовки предприятия к сертификации, в частности, с целью выхода на международный рынок.

Как отмечалось в начале главы, около 150 лет статистические методы применяются в России для проверки соответствия продукции установленным требованиям, т.е. для сертификации. С начала 1970-х годов стали разрабатываться государственные стандарты по статистическим методам. В связи с обнаружением в них грубых ошибок в 1985 г. была организована "Рабочая группа по упорядочению системы стандартов по прикладной статистике и другим статистическим методам". В ее работе приняли участие 66 специалистов, в том числе 15 докторов и 36 кандидатов наук. Выводы Рабочей группы кратко отражены в статьях [4, 22]. В соответствии с рекомендациями Рабочей группы 24 из 31 государственного стандарта по статистическим методам были отменены в 1986-87 гг. К сожалению, потеряв правовую силу как нормативные документы, ошибочные стандарты продолжают использоваться инженерами как научно-технические издания. Полученные Рабочей группой результаты и выводы не были широко и подробно опубликованы, ошибки в государственных стандартах не были публично вскрыты, и авторы дальнейших публикаций продолжают ссылаться на издания с грубейшими ошибками. Так, в многочисленных работах пропагандируются ошибочные стандарты, посвященные применению контрольных карт при статистическом регулировании технологических процессов. Продолжает широко использоваться грубо ошибочный ГОСТ 11.006-74 (СТ СЭВ 1190-78) "Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим", хотя разбору ошибок в этом стандарте посвящена уже давняя статья [23] (см. также главу 4). Перечисленные факты делают целесообразным популяризацию результатов и выводов Рабочей группы и в настоящее время, через 15 лет после окончания анализа стандартов по статистическим методам.

В 1988-89 гг. наиболее активная часть Рабочей группы (10 докторов и 15 кандидатов наук) составили "Аванпроект комплекса методических документов и пакетов программ по статистическим методам стандартизации и управления качеством". Это обширное сочинение (около 1600 стр.) и на настоящий момент является наиболее полным руководством по рассматриваемой тематике. Информация о нем приложена к переводу книги японских авторов по аналогичной тематике [1].

К сожалению, Госстандарт не пожелал финансировать реализацию заказанного им "Аванпроекта". Тогда решено было действовать самостоятельно. В 1989 г. был организован Центр статистических методов и информатики (ЦСМИ; в настоящее время - Институт высоких статистических технологий и эконометрики). К середине 1990 г. в ЦСМИ были разработаны 7 диалоговых систем по современным статистическим методам управления качеством, а именно, СПК, АТСТАТ-ПРП, СТАТКОН, АВРОРА-РС, ЭКСПЛАН, ПАСЭК, НАДИС (описания этих систем приведены в работе [5]). В работе участвовали 128 специалистов. В дальнейшем к ЦСМИ присоединялись новые группы научно-технических работников, уже к концу 1991 г. нас было более 300. Информация о программных продуктах и другой деятельности ЦСМИ постоянно помещалась в журналах "Заводская лаборатория" и "Надежность и контроль качества". Программные продукты, разработанные Центром статистических методов и информатики, использовались более чем в 100 организациях и предприятиях. Среди них - производственные объединения "Уралмаш", "АвтоВАЗ", "Пластик", ЦНИИ черной металлургии им. Бардина, НИИ стали, ВНИИ эластомерных материалов и изделий, НИИ прикладной химии, ЦНИИ химии и механики, НПО "Орион", НИЦентр по безопасности атомной энергетики, ВНИИ экономических проблем развития науки и техники, ВНИИ нефтепереработки, МИИТ, Казахский политехнический институт, Ульяновский политехнический институт, Донецкий государственный университет и др.

Как уже отмечалось, параллельно ЦСМИ вел работу по объединению статистиков и эконометриков. В апреле 1990 г. в Большом Актовом Зале Московского Энергетического института прошла Учредительная конференция Всесоюзной организации по статистическим методам и их применениям. На Учредительном съезде Всесоюзной статистической ассоциации (ВСА) в октябре 1990 г. в Московском экономико-статистическом институте эта организация вошла в состав ВСА в качестве секции статистических методов (подробнее о создании и задачах ВСА рассказано, например, в статьях [24, 25]). В 1992 г. после развала СССР и фактического прекращения работы ВСА на основе секции статистических методов ВСА организована Российская ассоциация по статистическим методам (РАСМ), а затем и Российская академия статистических методов, действующие и в настоящее время. В мероприятиях секции статистических методов ВСА и РАСМ активно участвовали несколько сот специалистов по статистическим методам и эконометрике. А одной из основных тематик этих специалистов являются, как следует из сказанного выше, статистические методы в сертификации (управлении качеством). В ЦСМИ и РАСМ, объединивших большинство ведущих российских специалистов, коллективными усилиями разработан единый подход к проблемам применения статистических методов в сертификации и управлении качеством. Именно с позиций ЦСМИ и РАСМ написана настоящая глава.

Классификация статистических методов сертификации. Рассмотрим два основания для классификации. Первый - по виду статистических методов. Второй - по этапам жизненного цикла продукции, на которых соответствующий метод применяется. Первое основание привычно для специалистов по разработке статистических методов и соответствующего программного обеспечения, второе - для тех, кто эти методы применяет на конкретных предприятиях.

В ЦСМИ сложилось пятичленное деление по первому основанию (в скобках указаны наименования диалоговых систем ЦСМИ, рассмотренных, в частности, в работе [5]):

а) прикладная статистика - иногда с дальнейшим выделением статистики случайных величин, многомерного статистического анализа, статистики случайных процессов и временных рядов, статистики объектов нечисловой природы (Система Регрессионного

Статистического Моделирования СРСМ, или СТАТМАСТЕР; АДДА, ГРАНТ, КЛАМС, ЭКОНОМЕТРИК, РЕГРЕССИЯ, ЛИСАТИС, ЭКОСТАТ, РЕСТ);

б) статистический приемочный контроль (СПК, АТСТАТ-ПРП, КОМПЛАН);

в) статистическое регулирование технологических процессов, в частности, методом контрольных карт (СТАТКОН, АВРОРА-РС);

г) планирование эксперимента (ПЛАН, ЭКСПЛАН, ПАСЭК, ПЛАНЭКС);

д) надежность и испытания (НАДИС, ОРИОН, СЕНС).

Быстрое развитие компьютерной техники имеет и свою оборотную сторону. Вполне добротные программные продукты устаревают и выходят из обращения просто потому, что они сделаны на отработавшем свой срок операционном (системном) программном обеспечении. "Выжить" может только то программное обеспечение, которое поддерживается соответствующей фирмой и постоянно совершенствуется с чисто программистской точки зрения. Важна система технической поддержки, обучение и, конечно, реклама. При этом чисто научная сторона дела отходит на задний план. Эти простые соображения объясняют, почему за 10 лет (с 1991 г. по 2001 г.) отечественный рынок программных продуктов по эконометрике и статистическим методам стал гораздо более бедным по числу продуктов, научный уровень явно понизился, зато дизайн явно стал более привлекательным.

Перейдем ко второму основанию классификации методов сертификации. Согласно п.5.1 "Петля качества" стандарта ИСО 9004 "Общее руководство качеством и элементы системы качества. Руководящие указания" система качества функционирует "... одновременно со всеми остальными видами деятельности, влияющими на качество продукции или услуг, и взаимодействует с ними. Ее воздействие распространяется на все этапы от первоначального определения и до конечного удовлетворения требований и потребностей потребителя. Эти этапы и виды деятельности включают:

1) маркетинг, поиски и изучение рынка;

2) проектирование и/или разработку технических требований, разработку продукции (опытного образца);

3) поиски поставщиков и оптовых покупателей, организацию материально-технического снабжения (решение задач логистики);

4) подготовку и разработку производственных (технологических) процессов;

5) непосредственно производство продукции;

6) контроль качества продукции, проведение испытаний и обследований;

7) упаковку и хранение продукции;

8) реализацию (сбыт) и распределение (доставку) продукции;

9) монтаж и эксплуатацию продукции у потребителей;

10) технические помощь и обслуживание;

11) утилизацию после использования.

Подробное рассмотрение применения основных типов статистических методов на перечисленных этапах жизненного пути продукции не входит в задачу настоящей книги. Сводка, приведенная в табл. 7, показывает, что статистические методы широко применяются на всех этапах жизненного пути продукции.

Табл. 7. Применение статистических методов на различных этапах жизненного цикла продукции по ИСО 9004

Номер этапа	а	Б	в	г	д	Специаль-ные модели
1	+	-	-	+	-	+

2	+	-	-	+	+	+
3	+	-	-	-	-	+
4	+	+	+	+	+	+
5	+	+	+	+	-	+
6	+	+	+	+	+	+
7	+	+	+	+	+	+
8	+	+	-	-	-	+
9	+	+	+	+	+	+
10	+	-	-	-	-	+
11	+	+	+	+	-	+

Помимо компьютерных диалоговых систем широкого назначения, на каждом конкретном предприятии и на любом конкретном этапе жизненного пути продукции могут быть использованы специальные модели, например, на этапе 3 “материально-техническое снабжение” - модели управления запасами (см. о них, например, главу 5 монографии [16]).

Среди диалоговых систем по статистическому анализу выделим пакеты, ориентированные на восстановление зависимостей (СТАТМАСТЕР, он же СРСМ - Система Регрессионного Статистического Моделирования, и его развитие ЭКОНОМЕТРИК, а также РЕГРЕССИЯ), анализ нечисловых данных на основе методов статистики объектов нечисловой природы (АДДА, КЛАМС, а также ориентированный на экспертное оценивание ГРАНТ, на анализ интервальных данных РЕСТ), прогнозирование (ЛИСАТИС и его развитие ЭКОСТАТ, а также относящиеся к временным рядам разделы пакета АВРОРА-РС - Анализ Временных Рядов и Обнаружение РАзладки).

Для регулярного решения обширных комплексов задач сертификации и управления качеством на конкретном предприятии в ряде случаев целесообразно создать диалоговую систему, предназначенную для использования именно на этом предприятии. В частности, для решения задач этапа 4 используют созданные для конкретного предприятия программные системы, соединяющие в себе банки данных и пакеты статистических методов анализа этих данных. Примерами являются "Автоматизированное рабочее место материаловеда (АРМ материаловеда)" и "Автоматизированное рабочее место математика (АРМ математика)", разработанные Центром статистических методов и информатики для ВНИИ эластомерных материалов и изделий.

Для объединения типовых пакетов в индивидуальную систему полезно программное средство ИНТЕГРАТОР - универсальный инструмент, предназначенный для создания интегрированных программных систем и обеспечивающий возможность совместного использования различных пакетов прикладных программ на персональных компьютерах IBM PC. Так, с помощью ИНТЕГРАТОРА был разработан АРМ математика на основе пакетов СРСМ, ПЛАН, АТСТАТ-ПРП, соответствующей базы данных и ряда программ, ориентированных на специфику ВНИИ эластомерных материалов и изделий.

На всех этапах жизненного цикла продукции, особенно на этапах 3, 8, 10, часто используют специализированные вероятностно-статистические модели, в том числе модели управления запасами (см., например, монографию [16, гл.5]), массового обслуживания и др. Такие модели и их программное обеспечение, как правило, разрабатываются для конкретного предприятия и потому хорошо приспособлены к особенностям этого предприятия.

Как избежать ошибок в нормативно-технической документации и инструктивно-методической документации? Как уже отмечалось, многие ошибочные государственные стандарты по статистическим методам управления качеством были отменены (хотя результаты анализа, проведенного Рабочей группой, так и не были вовремя и полностью опубликованы). Однако эти стандарты продолжают и до сих пор использоваться как авторитетные научно-технические публикации. Почему так происходит? Как вообще могли появиться ошибки в нормативно-технических документах и почему в течение ряда лет эти документы использовались, несмотря на очевидные для специалистов ошибки?

Дело в том, что инженеру, экономисту, менеджеру, работнику прикладной науки (короче - инженеру) несвойственно менять свою специальность, становиться математиком и самостоятельно повторять выкладки и рассуждения, положенные в основу ГОСТа. Поэтому инженер обычно не может самостоятельно обнаружить математические ошибки в ГОСТе, даже грубейшие. Главное - он не хочет этим заниматься. С другой стороны, математику несвойственно анализировать нормативно-техническую документацию. Он также обычно не хочет этим заниматься. "Рабочая группа по упорядочению системы стандартов по прикладной статистике и другим статистическим методам" была уникальным примером совместной работы математиков и инженеров, именно поэтому ей удалось сопоставить тексты стандартов с результатами современной науки.

Вполне естественно, что невежды и бюрократы Госстандарта сделали всё, чтобы помешать признанию и исправлению допущенных ими ошибок в государственных стандартах. До сих пор продолжаются попытки навязать промышленности (в качестве нормативных документов!) тексты, грубая ошибочность которых давно установлена. Кроме того, государственные стандарты, отмененные как нормативно-технические документы, продолжают физически существовать как издания (брошюры) и использоваться при проведении инженерных расчетов, проектировании систем контроля и т.д. Все это делает необходимым пропаганду выводов Рабочей группы относительно ГОСТов по статистическим методам. Они приведены в работе [4].

Обратим внимание только на результаты анализа государственных стандартов по общим вопросам статистических методов управления качеством (они не были отменены в результате деятельности Рабочей группы).

ГОСТ 15895-77 (СТ СЭВ 547-77, СТ СЭВ 3404-81). Статистические методы управления качеством продукции. Термины и определения.

Терминологический стандарт содержит огромное количество грубейших ошибок. Достойно удивления и сожаления, что подготовленные на столь низком научно-техническом уровне документы, как стандарты по терминологии и организации внедрения статистических методов, оказались утвержденными не только в СССР (и затем в России), но и в рамках международной стандартизации (в рамках СЭВ). На чиновников Госстандарта не подействовали ни заключения ведущих специалистов, ни научные публикации. Только общие решения по переводу подобных стандартов на уровень рекомендательных документов избавил академиков и профессоров - авторов учебников по теории вероятностей и математической статистике, по статистическим методам и эконометрике - от "преследования по закону" (!) за использование определений и обозначений, отличающихся от включенных в рассматриваемый стандарт. Наличие подобных "стандартов" - одна из причин появления терминологического "Приложения 1" в настоящей книге. В свое время этот текст был разработан взамен негодного стандарта СТ СЭВ 3404-81, однако справиться с невеждами вовремя не удалось…

ГОСТ 23853-79 (СТ СЭВ 3946-82). Организация внедрения статистических методов анализа, регулирования технологических процессов и статистического приемочного контроля качества продукции. Основные положения.

Документ описывает статистические методы управления качеством лишь в том небольшом и во многом устаревшем объеме, в котором они были стандартизированы к моменту его разработки. В документах Рабочей группы приведены 33 конкретные замечания, показывающие чрезвычайно низкий научно-технический уровень подготовки рассматриваемого документа. Даже в момент разработки его нельзя использовать ни как методический, ни тем более как нормативный документ.

Подведем итоги настоящей главы. В России активно разрабатываются теоретические, программные и практические аспекты эконометрические и статистических методов сертификации и управления качеством продукции. Некоторые из них кратко рассмотрены выше. Ранее разработанные нормативно-техническая и методическая документация, диалоговые компьютерные системы по статистическим методам продолжают использоваться, несмотря на политические преобразования. В частности, стандарты СССР и СЭВ продолжают оставаться широко известными методическими документами, хотя СССР и СЭВ уже нет. Большое значение имеет работа по устранению ошибок в нормативно-технических и инструктивно-методических документах с целью уменьшения числа ошибок в практической работе. Важно создать такую систему, чтобы никто не мог навязать стране свои ошибки в качестве стандартов, проигнорировав протесты ведущих специалистов. При этом условии внедрение современных эконометрических методов сертификации и управления качеством продукции могут дать нашей стране экономический эффект, измеряемый миллиардами долларов США в год.

Цитированная литература

1. Статистические методы повышения качества. Перевод с японского. / Под ред. Х. Кумэ. - М.: Финансы и статистика, 1990.- 301 с.
2. Организационно-методические материалы по маркетингу на предприятии. - М.: Всесоюзный центр статистических методов и информатики, 1991. - 91 с.
3. Цели и принципы стандартизации. / Под ред. Т. Сандерса. - М.: Изд-во стандартов, 1974. - 132 с.
4. Орлов А.И. Сертификация и статистические методы // Заводская лаборатория. - 1997. - Т.63 - No.3. - С.55-62.
5. Орлов А.И. Внедрение современных статистических методов с помощью персональных компьютеров // Качество и надежность изделий. No.5(21). - М.: Знание, 1992. - С.51-78.
6. Орлов А.И. Об оптимизации выборочного контроля качества продукции // Стандарты и качество. - 1989. - No. 3. - С. 91-94.
7. Broody M. Helping Workers Work Smarter / Fortune. June 8. 1987. Pp.86-88.
8. Гнеденко Б.В. Математика и контроль качества продукции. - М.: Знание, 1978. - 64 с.
9. Кравченко Г.Г., Орлов А.И. О статистическом приемочном контроле порошкообразных материалов // Надежность и контроль качества. 1991. No.2. С.37-39.
10. Мердок Дж. Контрольные карты. / Пер. с англ. - М.: Финансы и статистика, 1986. - 132 с.
11. Фалько С.Г., Носов В.М. Контроллинг на предприятии. - М.: Об-во "Знание" России, 1995. - 80 с.
12. Хан Д. Планирование и контроль: концепция контроллинга. / Пер. с нем. - М.: Финансы и статистика, 1997. - 800 с.
13. Карминский А.М., Оленев Н.И., Примак А.Г., Фалько С.Г. Контроллинг в бизнесе. Методологические и практические основы построения контроллинга в организациях. - М.: Финансы и статистика, 1998. - 256 с.
14. Уилкс С. Математическая статистика. - М.: Наука, 1967. - 632 с.
15. Кендалл М.Дж., Стьюарт А. Статистические выводы и связи. - М.: Наука, 1973. - 900 с.
16. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979. - 296 с.
17. Лэйард Р. Макроэкономика. Курс лекций для российских читателей. - М.: "Джон Уайли энд Санз", 1994. - 160 с.
18. Пиндайк Р., Рубинфельд Д. Микроэкономика. - М.: "Экономика" - “Дело", 1992. - 510 с.

19. Varian H.R. Intermediate Microeconomics. A Modern Approach. - New York: W.W.Norton & Company, 1993. - 623 pp.
20. Begg D., Fischer S., Dornbusch R. Economics. - London: McGraw-Hill Book Company, 1991. - 667 pp.
21. Brealey R.A., Myers S.C. Principles of Corporate Finance. - New York: McGraw-Hill, Inc., 1991. - 924 pp.
22. Орлов А.И. О современных проблемах внедрения прикладной статистики и других статистических методов. // Заводская лаборатория. - 1992. - Т.58.- №1. - С. 67-74.
23. Орлов А.И. Распространенная ошибка при использовании критериев Колмогорова и омега-квадрат. // Заводская лаборатория. 1985. Т.51. № 1. С.60-62.
24. Орлов А.И. Создана единая статистическая ассоциация. // Вестник Академии наук СССР. 1991. № 7. С.152-153.
25. Орлов А.И. Всесоюзная статистическая ассоциация - гарантия успешного внедрения современных статистических методов. // Надежность и контроль качества. 1991. № 6. С.54-55.

Глава 14. Эконометрика прогнозирования и риска

Последствия решений менеджера, экономиста, инженера проявятся в будущем. А будущее неизвестно. Мы обречены принимать решения в условиях неопределенности. Мы всегда рискуем, поскольку нельзя исключить возможность нежелательных событий. Но можно сократить вероятность их появления. Для этого необходимо спрогнозировать дальнейшее развитие событий, в частности, последствия принимаемых решений.

14.1. Методы социально-экономического прогнозирования

Кратко рассмотрим различные методы прогнозирования (предсказания, экстраполяции), используемые в социально-экономической области. По вопросам прогнозирования имеется большое число публикаций (см., например, книги [1-9]). Как часть эконометрики существует научная и учебная дисциплина "Математические методы прогнозирования". Ее целью является разработка, изучение и применение современных математических методов эконометрического (в частности, статистического, экспертного, комбинированного) прогнозирования социально-экономических явлений и процессов, причем методы должны быть проработаны до уровня, позволяющего их использовать в практической деятельности экономиста, инженера и менеджера. К основным задачам этой дисциплины относятся разработка, изучение и применение современных математико-статистических методов прогнозирования (в том числе непараметрических методов наименьших квадратов с оцениванием точности прогноза, адаптивных методов, методов авторегрессии и др.), развитие теории и практики экспертных методов прогнозирования, в том числе методов анализа экспертных оценок на основе статистики нечисловых данных, методов прогнозирования в условиях риска и комбинированных методов прогнозирования с использованием совместно экономико-математических и эконометрических (как статистических, так и экспертных) моделей. Теоретической основой методов прогнозирования являются математические дисциплины (прежде всего, теория вероятностей и математическая статистика, дискретная математика, исследование операций), а также экономическая теория, экономическая статистика, менеджмент, социология, политология и другие социально-экономические науки.

Как общепринято со времен основоположника научного менеджмента Анри Файоля, прогнозирование и планирование - основа работы менеджера. Сущность эконометрического прогнозирования состоит в описании и анализе будущего развития, в отличие от планирования, при котором директивным образом задается будущее движение. Например, вывод прогнозиста может состоять в том, что за час мы сможем отойти пешком от точки А не более чем на 5 км, а указание плановика - в том, что через час необходимо быть в точке Б. Ясно, что если расстояние между А и Б не более 5 км, то план реален (осуществим), а если более 10 км - не может быть осуществлен в заданных условиях. Необходимо либо отказаться от нереального плана, либо перейти на иные условия его реализации, например, двигаться не пешком, а на автомашине. Рассмотренный пример демонстрирует возможности и ограниченность методов прогнозирования. А именно, эти методы могут быть успешно применены при условии некоторой стабильности развития ситуации и отказывают при резких изменениях.

Один из вариантов применения методов прогнозирования - выявление необходимости изменений путем "приведения к абсурду". Например, если население Земли каждые 50 лет будет увеличиваться вдвое, то нетрудно подсчитать, через сколько лет на каждый квадратный метр поверхности Земли будет приходиться по 10000 человек. Из такого прогноза следует, что закономерности роста численности населения должны измениться.

Учет нежелательных тенденций, выявленных при прогнозировании, позволяет принять необходимые меры для их предупреждения, а тем самым помешать осуществлению прогноза.

Есть и самоосуществляющиеся прогнозы. Например, если в вечерней телевизионной передаче будет сделан прогноз о скором банкротстве определенного банка, то наутро многие вкладчики этого банка пожелают получить свои деньги, у входа в банк соберется толпа, а банковские операции придется остановить. Такую ситуацию журналисты описывают словами: "Банк лопнул". Обычно для этого достаточно, чтобы в один "прекрасный" (для банка) момент вкладчики пожелали изъять заметную долю (скажем, 30%) средств с депозитных счетов.

Прогнозирование - частный вид моделирования как основы познания и управления.

Роль прогнозирования в управлении страной, отраслью, регионом, предприятием очевидна. Необходимы учет СТЭП-факторов (социальных, технологических, экономических, политических), факторов конкурентного окружения и научно-технического прогресса, а также прогнозирование расходов и доходов предприятий и общества в целом (в соответствии с жизненным циклом продукции - во времени и по 11-и стадиям международного стандарта ИСО 9004). Проблемы внедрения и практического использования математических методов эконометрического прогнозирования связаны прежде всего с отсутствием в нашей стране достаточно обширного опыта подобных исследований, поскольку в течение десятилетий планированию отдавался приоритет перед прогнозированием.

Статистические методы прогнозирования. Простейшие методы восстановления используемых для прогнозирования зависимостей исходят из заданного временного ряда, т.е. функции, определенной в конечном числе точек на оси времени. Задачам анализа и прогноза временных рядов посвящена глава 6 выше. Временной ряд при этом часто рассматривается в рамках вероятностной модели, вводятся иные факторы (независимые переменные), помимо времени, например, объем денежной массы (агрегат М2). Временной ряд может быть многомерным, т.е. число откликов (зависимых переменных) может быть больше одного. Основные решаемые задачи - интерполяция и экстраполяция. Метод наименьших квадратов в простейшем случае (линейная функция от одного фактора) был разработан К.Гауссом более двух столетий назад, в 1794-1795 гг. (см. главу 5). Могут оказаться полезными предварительные преобразования переменных.

Опыт прогнозирования индекса инфляции и стоимости потребительской корзины накоплен в Институте высоких статистических технологий и эконометрики. При этом оказалось полезным преобразование (логарифмирование) переменной - текущего индекса инфляции. Характерно, что при стабильности условий точность прогнозирования оказывалась достаточно удовлетворительной - 10-15 %. Однако спрогнозированное на осень 1996 г. значительное повышение уровня цен не осуществилось. Дело в том, что руководство страны перешло к стратегии сдерживания роста потребительских цен путем массовой невыплаты зарплаты и пенсий. Условия изменились - и статистический прогноз оказался непригодным. Влияние решений руководства Москвы проявилось также в том, что в ноябре 1995 г. (перед парламентскими выборами) цены в Москве упали в среднем на 9,5%, хотя обычно для ноября характерен более быстрый рост цен, чем в другие месяцы года, кроме декабря и января.

Наиболее часто используется метод наименьших квадратов при нескольких факторах (2-5). Метод наименьших модулей и другие методы экстраполяции применяются реже, хотя их статистические свойства зачастую лучше. Большую роль играет традиция и общий невысокий уровень знаний об эконометрических методах прогнозирования.

Оценивание точности прогноза - необходимая часть процедуры квалифицированного прогнозирования. При этом обычно используют вероятностно-статистические модели восстановления зависимости, например, строят наилучший прогноз по методу максимального правдоподобия. Разработаны параметрические (обычно на основе модели нормальных ошибок) и непараметрические оценки точности прогноза и доверительные границы для него (на основе Центральной Предельной Теоремы теории вероятностей). Так, нами предложены и изучены методы доверительного оценивания точки наложения (встречи) двух временных рядов и их применения для оценки динамики технического уровня собственной продукции и продукции конкурентов, представленной на мировом рынке.

Применяются также эвристические приемы, не основанные на какой-либо теории: метод скользящих средних, метод экспоненциального сглаживания.

Адаптивные методы прогнозирования позволяют оперативно корректировать прогнозы при появлении новых точек. Речь идет об адаптивных методах оценивания параметров моделей и об адаптивных методах непараметрического оценивания. Отметим, что с развитием вычислительных мощностей компьютеров проблема сокращения объемов вычисления теряет свое значение.

Многомерная регрессия, в том числе с использованием непараметрических оценок плотности распределения - основной на настоящий момент эконометрический аппарат прогнозирования. Подчеркнем, что нереалистическое предположение о нормальности погрешностей измерений и отклонений от линии (поверхности) регрессии использовать не обязательно. Однако для отказа от предположения нормальности необходимо опереться на иной математический аппарат, основанный на многомерной центральной предельной теореме теории вероятностей и эконометрической технологии линеаризации. Он позволяет проводить точечное и интервальное оценивание параметров, проверять значимость их отличия от 0 в непараметрической постановке, строить доверительные границы для прогноза.

Весьма важна проблема проверки адекватности модели, а также проблема отбора факторов. Дело в том, что априорный список факторов, оказывающих влияние на отклик, обычно весьма обширен, желательно его сократить, и крупное направление современных эконометрических исследований посвящено методам отбора "информативного множества признаков". Однако эта проблема пока еще окончательно не решена. Проявляются необычные эффекты. Так, в главе 5 установлено, что обычно используемые оценки степени полинома имеют геометрическое распределение. Перспективны непараметрические методы оценивания плотности вероятности и их применения для восстановления регрессионной зависимости произвольного вида. Наиболее общие постановки в этой области получены с помощью подходов статистики нечисловых данных (см. главу 8).

К современным статистическим методам прогнозирования относятся также модели авторегрессии, модель Бокса-Дженкинса, системы эконометрических уравнений, основанные как на параметрических, так и на непараметрических подходах.

Для установления возможности применения асимптотических результатов при конечных (т.н. "малых") объемах выборок полезны компьютерные статистические технологии (см. главу 11). Они позволяют также строить различные имитационные модели. Отметим полезность методов размножения данных (бутстреп-методов). Системы прогнозирования с интенсивным использованием компьютеров объединяют различные методы прогнозирования в рамках единого автоматизированного рабочего места прогнозиста.

Прогнозирование на основе данных, имеющих нечисловую природу, в частности, прогнозирование качественных признаков основано на результатах статистики нечисловых данных (см. главу 8). Весьма перспективными для прогнозирования представляются регрессионный анализ на основе интервальных данных, включающий, в частности, определение и расчет нотны и рационального объема выборки (см. главу 8), а также регрессионный анализ нечетких данных, разработанный в монографии [10]. Общая постановка регрессионного анализа в рамках статистики нечисловых данных и ее частные случаи - дисперсионный анализ и дискриминантный анализ (распознавание образов с учителем), давая единый подход к формально различным методам, полезна при программной реализации современных статистических методов прогнозирования.

Экспертные методы прогнозирования. Необходимость и общее представление о применении экспертных методов прогнозирования при принятии решений на различных уровнях управления - на уровне страны, отрасли, региона, предприятия - вытекает из рассмотрений главы 12. Отметим большое практическое значение экспертиз при сравнении и выборе инвестиционных и инновационных проектов, при управлении проектами, экологических экспертиз. Роли лиц, принимающих решения (ЛПР), и специалистов (экспертов) в процедурах принятия решений, критерии принятия решений и место экспертных оценок в процедурах принятие решений рассмотрены выше. В качестве примеров конкретных экспертных процедур, широко используемых при прогнозировании, укажем метод Дельфи и метод сценариев. На их основе формируются конкретные процедуры подготовки и принятия решений с использованием методов экспертных оценок, например, процедуры распределения финансирования научно-исследовательских работ (на основе балльных оценок или парных сравнений), технико-экономического анализа, кабинетных маркетинговых исследований (противопоставляемых "полевым" выборочным исследованиям - см. главу 2), оценки, сравнения и выбора инвестиционных проектов.

В соотнесении с задачами прогнозирования напомним о некоторых аспектах планирования и организации экспертного исследования. Должны быть сформированы Рабочая группа и экспертная комиссия. Основные этапы проведения экспертного исследования рассмотрены в главе 12. Весьма ответственными этапами являются формирование целей экспертного исследования (сбор информации для ЛПР и/или подготовка проекта решения для ЛПР и др.) и формирование состава экспертной комиссии (методы списков (реестров), "снежного кома", самооценки, взаимооценки) с предварительным решением проблемы априорных предпочтений экспертов. Различные варианты организации экспертного исследования, различающиеся по числу туров (один, несколько, не фиксировано), порядку вовлечения экспертов (одновременно, последовательно), способу учета мнений (с весами, без весов), организации общения экспертов (без общения, заочное, очное с ограничениями ("мозговой штурм") или без ограничений) позволяют учесть специфику конкретного экспертного исследования. Компьютерное обеспечение деятельности экспертов и Рабочей группы, экономические вопросы проведения экспертного исследования важны для успешного проведения экспертного исследования.

Напомним, что экспертные оценки могут быть получены в различных математических формах. Наиболее часто используются количественные или качественные (порядковые, номинальные) признаки, бинарные отношения (ранжировки, разбиения, толерантности), интервалы, нечеткие множества, результаты парных сравнений, тексты и др. (см. главы 8 и 9) Основные понятия (репрезентативной) теории измерений: основные типы шкал, допустимые преобразования, адекватные выводы и др. - важны применительно к экспертному оцениванию. Необходимо использовать средние величины, соответствующие основным шкалам измерения. Применительно к различным видам рейтингов репрезентативная теория измерений позволяет выяснить степень их адекватности прогностической ситуации, предложить наиболее полезные для целей прогнозирования (см. главу 3).

Например, анализ рейтингов политиков по степени их влиятельности, публиковавшийся одной из известных центральных газет, показал, что из-за неадекватности используемого математического аппарата лишь первые 10 мест, возможно, имеют некоторое отношение к реальности (они не меняются при переходе к другому способу анализа данных, т.е. не зависят от субъективизма членов Рабочей группы), остальные - "информационный шум", попытки опираться на них при прогностическом анализе могут привести лишь к ошибкам. Что же касается начального участка рейтинга этой газеты, то он также может быть подвергнут сомнению, но по более глубоким причинам, например, связанным с составом экспертной комиссии.

Основными процедурами обработки прогностических экспертных оценок являются проверка согласованности, кластер-анализ и нахождение группового мнения.

Проверка согласованности мнений экспертов, выраженных ранжировками, проводится с помощью коэффициентов ранговой корреляции Кендалла и Спирмена, коэффициента ранговой конкордации Кендалла и Бэбингтона Смита. Используются параметрические модели парных сравнений - Терстоуна, Бредли-Терри-Льюса - и непараметрические модели теории люсианов (о люсианах см. главу 8). В главе 12 рассмотрена процедура согласования ранжировок и классификаций; построения согласующих бинарных отношений по ГОГ - методу (т. е. методу Горского-Орлова-Гриценко).

При отсутствии согласованности разбиение мнений экспертов на группы сходных между собой проводят методом ближайшего соседа или другими методами кластерного анализа (автоматического построения классификаций, распознавания образов без учителя). Классификация люсианов осуществляется на основе вероятностно-статистической модели.

Используют различные методы построения итогового мнения комиссии экспертов. Своей простотой выделяется метод средних рангов. Компьютерное моделирование (см. работу [11]) позволило установить ряд свойств медианы Кемени, часто рекомендуемой для использования в качестве итогового (обобщенного, среднего) мнения комиссии экспертов. Интерпретация закона больших чисел для нечисловых данных в терминах теории экспертного опроса такова: итоговое мнение устойчиво, т.е. мало меняется при изменении состава экспертной комиссии, и при росте числа экспертов приближается к "истине". При этом в соответствии с принятым в монографии [12] подходом предполагается, что ответы экспертов можно рассматривать как результаты измерений с ошибками, все они - независимые одинаково распределенные случайные элементы, вероятность принятия определенного значения убывает по мере удаления от некоторого центра - "истины", а общее число экспертов достаточно велико.

Проблемы применения методов прогнозирования в условиях риска. Многочисленны примеры ситуаций, связанных с социальными, технологическими, экономическими, политическими, экологическими и другими рисками. Именно в таких ситуациях обычно и необходимо прогнозирование. Известны различные виды критериев, используемых в теории принятия решений в условиях неопределенности (риска). Из-за противоречивости решений, получаемых по различным критериям, очевидна необходимость применения оценок экспертов.

В конкретных задачах прогнозирования необходимо провести классификацию рисков, поставить задачу оценивания конкретного риска, провести структуризацию риска, в частности, построить деревья причин (ы другой терминологии, деревья отказов) и деревья последствий (деревья событий). Центральной задачей является построение групповых и обобщенных показателей, например, показателей конкурентоспособности и качества. Риски необходимо учитывать при прогнозировании экономических последствий принимаемых решений, поведения потребителей и конкурентного окружения, внешнеэкономических условий и макроэкономического развития России, экологического состояния окружающей среды, безопасности технологий, экологической опасности промышленных и иных объектов. Метод сценариев незаменим применительно к анализу технических, экономических и социальных последствий аварий.

Имеется некоторая специфика применения методов прогнозирования в ситуациях, связанных с риском. Велика роль функции потерь и методов ее оценивания, в том числе в экономических терминах. В конкретных областях используют вероятностный анализ безопасности (для атомной энергетики) и другие специальные методы. Эконометрике риска посвящен ряд дальнейших разделов настоящей главы.

Современные компьютерные технологии прогнозирования. Перспективны интерактивные методы прогнозирования с использованием баз эконометрических данных, имитационных (в том числе на основе применения метода Монте-Карло, т.е. метода статистических испытаний) и экономико-математических динамических моделей, сочетающих экспертные, статистические и моделирующие блоки. Обратим внимание на сходство и различие методов экспертных оценок и экспертных систем. Можно сказать, что экспертная система моделирует поведение эксперта путем формализации его знаний по специальной технологии. Но интуицию "живого эксперта" нельзя заложить в ЭВМ, а при формализации мнений эксперта (фактически - при его допросе) наряду с уточнением одних его представлений происходит и огрубление других. Другими словами, при использовании экспертных оценок непосредственно обращаются к опыту и интуиции высококвалифицированных специалистов, а при применении экспертных систем имеют дело с компьютерными алгоритмами расчетов и выводов, при создании которых когда-то давно привлекались эксперты как источник данных и типовых заключений.

Обратим внимание на возможность использования в прогнозировании производственных функций, статистически описывающих связь выпуска с факторами производства, на различные способы учета научно-технического прогресса, в частности, на основе анализа трендов и с помощью экспертного выявления точек роста. Примеры экономических прогнозов всех видов имеются в литературе. К настоящему времени разработаны компьютерные системы и программные средства комбинированных методов прогнозирования. Одна из первых таких систем была создана в 70-е годы в ИМЭМО АН СССР под руководством С.А.Петровского.

14.2. Основные идеи технологии сценарных экспертных прогнозов

Как уже отмечалось в пункте 14.1, социально-экономическое прогнозирование, как и любое прогнозирование вообще, может быть успешным лишь при некоторой стабильности условий. Однако решения органов власти, отдельных лиц, иные события меняют условия, и события развиваются по-иному, чем ранее предполагалось. Объективно имеются точки выбора (фуркации), после которых рассматриваемое прогнозистами развитие может пойти по одному из нескольких возможных путей (эти пути и называют обычно сценариями). Выбор может делаться на разных уровнях - конкретной личностью (перейти на другую работу или остаться), менеджером (выпускать ту или иную марку продукции), конкурентами (сотрудничество или борьба), властными структурами (выбор той или иной системы налогообложения), населением страны (выбор президента), "международным сообществом" (вводить или нет санкции против России).

Рассмотрим пример. Вполне очевидно, что после первого тура президентских выборов 1996 г. о дальнейшем развитии социально-экономических событий можно было говорить лишь в терминах сценариев: если победит Б.Н. Ельцин, то будет то-то и то-то, если победит Г.А. Зюганов, то события пойдут так-то и так-то.

Например, работа [13] имела целью прогноз динамики валового внутреннего продукта (ВВП) на 9 лет (1999-2007). При ее проведении было ясно, что за это время произойдут различные политические события, в частности, по крайней мере два цикла парламентских и президентских выборов (при условии сохранения нынешней политической структуры), результаты которых нельзя предсказать однозначно. Поэтому прогноз динамики ВВП мог быть сделан лишь по отдельности для каждого сценария из некоторой гаммы, охватывающей возможные пути социально-экономической динамики России.

Метод сценариев необходим не только в социально-экономической области. Например, при разработке методологического, программного и информационного обеспечения анализа риска химико-технологических проектов необходимо составить детальный каталог сценариев аварий, связанных с утечками токсических химических веществ. Каждый из таких сценариев описывает аварию своего типа, со своим индивидуальным происхождением, развитием, техническими, экономическими и социальными последствиями, возможностями предупреждения.

Таким образом, метод сценариев - это метод декомпозиции (разделения на части) задачи прогнозирования, предусматривающий выделение набора отдельных вариантов развития событий (сценариев), в совокупности охватывающих все возможные варианты развития. При этом каждый отдельный сценарий должен допускать возможность достаточно точного прогнозирования, а общее число сценариев должно быть обозримо.

Возможность подобной декомпозиции не очевидна. При применении метода сценариев необходимо осуществить два этапа исследования:

- построение исчерпывающего, но обозримого набора сценариев;
- прогнозирование в рамках каждого конкретного сценария с целью получения ответов на интересующие исследователя вопросы.

Каждый из этих этапов лишь частично формализуем. Существенная часть рассуждений проводится на качественном уровне, как это принято в общественно-экономических и гуманитарных науках. Одна из причин заключается в том, что стремление к излишней формализации и математизации приводит к искусственному внесению определенности там, где ее нет по существу, либо к использованию громоздкого математического аппарата. Так, рассуждения на словесном уровне считаются доказательными в большинстве ситуаций принятия решений, в то время как попытка уточнить смысл используемых слов с помощью, например, теории нечетких множеств приводит к весьма громоздким математическим моделям и расчетам.

Для построения исчерпывающего, но обозримого набора сценариев необходимо предварительно проанализировать динамику социально-экономического развития рассматриваемого экономического агента и его окружения. Корни будущего - в настоящем и прошлом, причем зачастую - в весьма далеком прошлом. Кроме макроэкономических и микроэкономических характеристик, известных лишь с погрешностями, которые нельзя считать случайными или малыми, необходимо учитывать состояние и динамику отечественного массового сознания, политических, в то числе внешнеполитических реалий, поскольку на обычно рассматриваемом интервале времени (до 10 лет) экономика зачастую следует за политикой, а не наоборот.

Так, например, к началу 1985 г. экономика СССР находилась в достаточно стабильном состоянии с ежегодным ростом в среднем 3-5%. Если бы руководство страны находилось в руках иных людей, то развитие продолжалось бы в прежних условиях и к концу тысячелетия ВВП СССР увеличился бы на 50% и составил бы примерно 150 % от уровня 1985 г. Реально же из-за политических причин ВВП России за эти 15 лет упал примерно в 2 раза, т.е. составил около 50 % по сравнению с 1985 г., или в 3 раза меньше, чем можно было бы ожидать из чисто экономических причин при сохранении стабильных условий 1985 г.

Набор сценариев должен быть обозрим. Приходится исключать различные маловероятные события - прилет инопланетян, падение астероида, массовые эпидемии ранее неизвестных болезней, и т.д.

Само по себе создание набора сценариев - предмет экспертного исследования, проводимого в соответствии с описанной выше методологией. Кроме того, эксперты могут оценить вероятности реализации того или иного сценария. Ясно, что эти оценки не являются надежными.

Часто используют упрощенный подход к прогнозированию методом сценариев. А именно, формулируют три сценария - оптимистический, вероятный и пессимистический. При этом для каждого из сценариев достаточно произвольно выбирают значения параметров, описывающих производственно-экономическую ситуацию (по-английски - case). Цель такого подхода - рассчитать интервалы разброса для характеристик и "коридоры" для временных рядов, интересующих исследователя (и заказчика исследования). Например, прогнозируют финансовый поток (по-английски - cash flow) и чистую текущую стоимость (по-английски - net present value или NPV) инвестиционного проекта.

Ясно, что такой упрощенный подход не может дать максимального или минимального значения характеристики, он дает лишь представление о порядке количественной меры разброса. Однако его развитие приводит к байесовской постановке в теории принятия решений. Например, если сценарий описывается элементом конечномерного евклидова пространства, то любое вероятностное распределение на множестве исходных параметров преобразуется в распределение интересующих исследователя характеристик. Расчеты могут быть проведены с помощью современных информационных технологий метода статистических испытаний. Надо в соответствии с заданным распределением на множестве параметров выбирать с помощью датчика псевдослучайных чисел конкретный вектор параметров и рассчитывать для него итоговые характеристики. В результате получится эмпирическое распределение на множестве итоговых характеристик, которое можно разными способами анализировать, находить оценку математического ожидания, разброса и др. Остается только неясным, как задавать распределение на множестве параметров. Естественно, для этого можно использовать экспертов.

Прогнозирование в рамках каждого конкретного сценария с целью получения ответов на интересующие исследователя вопросы также осуществляется в соответствии с описанной выше методологией прогнозирования. При стабильных условиях могут быть применены статистические методы прогнозирования временных рядов. Однако этому обычно предшествует анализ с помощью экспертов, причем зачастую прогнозирование на словесном уровне является достаточным (для получения интересующих исследователя и ЛПР выводов) и не требующим количественного уточнения.

Вопрос об использовании результатов прогнозирования относится не к эконометрике, а к смежной науке - теории принятия решений. Как известно, при принятии решений на основе анализа ситуации, в том числе результатов прогнозных исследований, можно исходить из различных критериев. Так, можно ориентироваться на то, что ситуация сложится наихудшим, или наилучшим, или средним (в каком-либо смысле) образом. Можно попытаться наметить мероприятия, обеспечивающие минимально допустимые полезные результаты при любом варианте развития ситуации, и т.д.

Итак, рассмотрена концепция современной методики экспертного оценивания методом сценариев. Она использовалась, например, для прогнозирования социально-экономического развития России (см. работу [13]).

14.3. Различные виды рисков

Будущее нам неизвестно. А потому неизвестны и будущие доходы и расходы, мы можем лишь прогнозировать их с той или иной степенью уверенности. Как описывать неопределенность будущего? Чем мы рискуем и что вообще понимать под "риском"? Как отражается неопределенность будущего на финансовых потоках (потоках платежей и поступлений), их характеристиках и выводах об эффективности управляющих воздействий на те или иные экономические процессы и других решениях? Как уменьшить возможные потери и защититься от рисков?

Чтобы управлять рисками, надо сначала знать риски. Поскольку на деятельность любой организации непосредственно либо потенциально влияют риски различной природы, необходима классификация рисков. Возможно, для различных целей понадобятся различные классификации, основанные на различных методологических принципах.

Для построения такой классификации необходимо какой-либо упорядочивающий принцип. Возьмем за основу движение от частного к общему. Тогда естественно выделить:

производственные риски, связанные непосредственно с деятельностью предприятия;

коммерческие риски, вызванные неполной предсказуемостью динамики рынка, т.е. действий потребителей и конкурентов;

финансовые риски, определяемые макроэкономической ситуацией;

риски, возникающие на уровне государства и Земли в целом.

Затем необходимо изучить степень их влияния на показатели эффективности деятельности организации с целью выделения наиболее значимых.

После этого целесообразно провести изучение различных способов оценки финансовых и иных рисков в случаях, когда они моделируются с помощью тех или иных математических структур. В частности, распространено моделирование рисков с помощью вероятностей и случайных величин. При этом используются такие характеристики случайной величины, как математическое ожидание, дисперсия, квантили, коэффициент вариации, линейные комбинации математического ожидания и среднего квадратического отклонения и др. Подчеркнем, что эти характеристики следует рассматривать в непараметрической постановке, поскольку нет никаких оснований предполагать, что распределение характеристики риска входит в то или иное из известных параметрических семейств.

Перспективной представляется разработка методов описания рисков с помощью теории нечетких множеств, лингвистических переменных, качественных признаков, интервальных математических и эконометрических моделей и др.

Существенно, что описание может быть многомерным. Например, каждая координата может соответствовать своему виду воздействия (нарушения, происшествия) и описываться количественным либо качественным признаком. Тогда дополнительно возникает задача агрегирования (сведения вместе) показателей риска. Для агрегирования могут быть использованы различные методы, разработанные в теории оценки технического уровня и в теории экспертных оценок.

Следующий этап - разработка методологии применения различных методов управления рисками с использованием экспертных оценок, современных методов прогнозирования, эконометрических и экономико-математических моделей с целью повышения эффективности деятельности организации в условиях риска. При этом необходимо научиться практически решать проблему многокритериальности (согласования оценок рисков, полученных по различным основаниям, с целью эффективного управления риском).

К настоящему времени накоплена огромная литература по вопросам риска, как общая, например, теория статистического риска, так и по отдельным вопросам – по экологическим рискам, статистическим методам обеспечения качества, финансовым рискам и др.

Производственные риски. К ним можно прежде всего отнести риски, связанные с выпуском дефектной продукции. Хорошо известно, что при массовом производстве невозможно обеспечить выпуск продукции без дефектов. Поэтому действуют отделы технического контроля (ОТК), службы (бюро) качества и другие подразделения, осуществляющие контроль качества продукции. Известно, что в машиностроении стоимость контрольных операций составляет в среднем около 10% от стоимости продукции. Часть риска компенсируется службами технического обслуживания продукции, уже находящейся у потребителя. Постоянно используемыми терминами в этой области являются «риск поставщика» и «риск потребителя». Вопросам управления качеством посвящена обширная литература (см. главу 13). Одна из важных групп показателей качества – надежность.

Другой вид рисков связан с осуществлением действующих технологических процессов. Речь идет об авариях различной степени тяжести, от незначительных нарушений технологических процессов до катастроф с человеческими жертвами. Здесь целесообразно обратить внимание на экологические риски, в частности, связанные с аварийными сбросами в реки технологических жидкостей, выбросами в атмосферу газов и взвешенных частиц и др. За подобные действия предприятия обычно обязаны платить штрафы согласно предписаниям экологических органов.

Отметим риски, относящиеся к проектируемым продукции или технологическим процессам. Они могут быть связаны с ошибками разработчиков или физической невозможностью осуществления того или иного процесса. Так, в течение всей второй половины XX века физики постоянно говорили о появлении в ближайшее время неиссякаемого источника энергии на основе преобразования тяжелой воды с помощью управляемого термоядерного синтеза. Эта пропаганда, несомненно, сдерживала финансирование и развитие ресурсосберегающих технологий. Еще в начале XX в. Д.И. Менделеев говорил, что сжигать нефть – это то же самое, что топить печь ассигнациями. Тем не менее и сейчас нефть используют как топливо, разведанных запасов остается все меньше. Излишний оптимизм физиков нам всем еще дорого обойдется.

Среди производственных рисков есть и социальные, связанные с теми или иными конфликтами. Здесь надо разделять конфликты между службами (отделами, цехами), с которыми можно бороться, оптимизируя организационную структуру предприятия; различного происхождения конфликты между менеджерами высшего звена; конфликты между профсоюзами и администрацией по поводу заработной платы или условий труда, и др. Современные методы управления персоналом позволяют заранее спрогнозировать многие из таких конфликтов и предложить пути их разрешения.

Коммерческие риски. Речь идет о рисках, связанных с неопределенностью будущей рыночной ситуации в стране. В частности, о будущих действий поставщиков в связи с меняющимися предпочтениями потребителей. Напомним, например, о быстрых изменениях на рынке вычислительной техники в связи с появлением персональных компьютеров. Мода в той или иной степени отражается на поведении потребителей во многих областях.

Весьма существенны риски, связанные с деятельностью партнеров организации - участников экономической жизни (в том числе их законопослушностью как налогоплательщиков), в частности, с их деловой активностью, финансовым положением, отношением к соблюдению обязательств. Особенно надо отметить роль конкурентного окружения, от действий которого зависит многое в судьбе конкретного предприятия. В частности, важны информационные риски, связанные с промышленным шпионажем и возможностями проникновения конкурентов в коммерческие тайны и иного воздействия на внутренние дела организации, в частности, через компьютерные сети типа Интернет.

К этому же типу можно отнести риски, связанные с социальными и административными факторами в конкретных регионах, с взаимоотношениями рассматриваемой организации с органами местной и региональной власти, как официальными, так и криминальными.

Финансовые риски. Отметим прежде всего риски, связанные с колебаниями цен на товары и услуги (динамикой инфляции), ставки рефинансирования Центрального банка, норм банковских процентов по кредитам и депозитам, валютных курсов и других макроэкономических показателей, в том числе котировок государственных и частных (корпоративных) ценных бумаг. Часть этих рисков носит объективный, а часть – число спекулятивный характер. К этому же разделу можно отнести риски, связанные с нестабильностью законодательства и текущей экономической политики (т.е. с деятельностью руководства страны, министерств и ведомств). Дополнительные проблемы создает множественность нормативно-правовых актов, регулирующих хозяйственно-экономическую деятельность организации (порядка 10^4 , если считать не только федеральные нормативно-правовые акты, но и нормативно-правовые акты субъектов федерации, например, г. Москвы), зачастую противоречащих друг другу, что вызывает необходимость в участии в работе организации юристов, в том числе в судебных процессах.

Риски, возникающие на уровне государства и Земли в целом. К этому типу отнесем риски, связанные с политической ситуацией в целом, действиями партий, профсоюзов, экологических и других организаций в масштабе страны. Типичным примером являются риски, связанные с заметным изменением курса страны в результате тех или иных выборов. Другой пример – российский кризис, начавшийся в августе 1998 г. и непосредственно вызванный решением трех чиновников. Большое значения имеют риски, связанные с социальной борьбой («рельсовая война», забастовки, массовые столкновения, терроризм, и др.)…

Внешнеэкономические риски, например, связанные с динамикой цены на нефть, крупномасштабными зарубежными финансовыми (в Юго-Восточной Азии) или военными (Югославия) кризисами и т.д., могут оказать существенное воздействие на рассматриваемую организацию (предприятие).

Большое число рисков связано с природными явлениями. Их можно объединить под именем «экологические». К ним относятся, в частности, риски, связанные с неопределенностью ряда природных явлений. Типичным примером является погода, от которой зависят урожайность (а потому и цены на сельскохозяйственные товары), расходы на отопление и уборку улиц, доходы от туризма и др. Обратим внимание на риски, связанные с недостаточными знаниями о природе (например, нам неизвестен точный объем полезных ископаемых в том или ином месторождении, а потому мы не можем точно предсказать развитие добывающей промышленности и объем налоговых поступлений от ее предприятий). Нельзя забывать о рисках экологических бедствий и катастроф, типа ураганов, смерчей, землетрясений, цунами, селей и др.

Каждый из перечисленных видов рисков может быть структуризован далее. Так, имеются крупные развернутые разработки по анализу рисков технологических аварий, в частности, на химических производствах и на атомных электростанциях (соответствующая теория именуется ВАБ – вероятностный анализ безопасности). Ясно, что аварии типа Чернобыльской существенно влияют на значения СТЭП-факторов (принятое сокращение для комплекса социальных, технологических, экономических и политических факторов, действующих на организацию) и тем самым на поступления и выплаты из бюджета как на местном, так и на федеральном уровне (что существенно, если «организация» – это муниципальный или государственный орган власти или его подразделение типа налоговой инспекции).

Подходы к учету неопределенности и описанию рисков. В настоящее время при компьютерном и математическом моделировании для описания неопределенностей чаще всего используют такие математические средства, как:

- вероятностно-статистические методы,
- методы статистики нечисловых данных, в том числе интервальной статистики и интервальной математики, а также методы теории нечеткости,
- методы теории конфликтов (теории игр).

Они применяются в имитационных, эконометрических, экономико-математических моделях, реализованных обычно в виде программных продуктов.

Некоторые виды неопределенностей связаны с безразличными к организации силами - природными (погодные условия) или общественными (смена правительства). Если явление достаточно часто повторяется, то его естественно описывать в вероятностных терминах. Так, прогноз урожайности зерновых вполне естественно вести в вероятностных терминах. Если событие единично, то вероятностное описание вызывает внутренний протест, поскольку частотная интерпретация вероятности невозможна. Так, для описания неопределенности, связанной с исходами выборов или со сменой правительства, лучше использовать методы теории нечеткости, в частности, интервальной математики (интервал – удобный частный случай описания нечеткого множества). Наконец, если неопределенность связана с активными действиями соперников или партнеров, целесообразно применять методы анализа конфликтных ситуаций, т.е. методы теории игр, прежде всего антагонистических игр, но иногда полезны и более новые методы кооперативных игр, нацеленных на получение устойчивого компромисса.

Подходы к оцениванию рисков. Понятие "риск", как уже отмечалось, многогранно. Например, при использовании статистических методов управления качеством продукции риски - это вероятности некоторых событий (в статистическом приемочном контроле риск поставщика - это вероятность забракования партии продукции хорошего качества, а риск потребителя - приемки "плохой" партии; при статистическом регулировании процессов рассматривают риск незамеченной разладки и риск излишней наладки). Тогда оценка риска – это оценка вероятности, точечная или интервальная, по статистическим данных или экспертная. В таком случае для управления риском задают ограничения на вероятности нежелательных событий. Иногда под уменьшением риска понимают уменьшение дисперсии случайной величины, поскольку при этом уменьшается неопределенность. В теории принятия решений риск - это плата за принятие решения, отличного от оптимального, он обычно выражается как математическое ожидание. В экономике плата измеряется обычно в денежных единицах, т.е. в виде финансового потока (потока платежей и поступлений) в условиях неопределенности.

Методы математического моделирования позволяют предложить и изучить разнообразные методы оценки риска. Широко применяются два вида методов - статистические, основанные на использовании эмпирических данных, и экспертные, опирающиеся на мнения и интуицию специалистов.

Чтобы продемонстрировать сложность проблемы оценивания риска и различные существующие подходы, рассмотрим простейший случай. Пусть неопределенность носит вероятностный характер, а потери описываются случайной величиной (не вектором и не процессом). Тогда минимизация риска может состоять:

1) в минимизации математического ожидания (ожидаемых потерь),

2) в минимизации квантиля распределения (например, медианы функции распределения потерь или квантиля порядка 0,99, выше которого располагаются большие потери, встречающиеся крайне редко - в 1 случае из 100),

3) в минимизации дисперсии (т.е. показателя разброса возможных значений потерь),

4) в минимизации суммы математического ожидания и утроенного среднего квадратического отклонения (на основе известного "правила трех сигм"), или иной линейной комбинации математического ожидания и среднего квадратического отклонения (используют в случае близости распределения потерь к нормальному как комбинацию подходов, нацеленных на минимизацию средних потерь и разброса возможных значений потерь),

5) в максимизации математического ожидания функции полезности (в случае, когда полезность денежной единицы меняется в зависимости от общей располагаемой суммы, как предполагается в учебном пособии по микроэкономике [14], в частности, когда необходимо исключить возможность разорения экономического агента), и т.д.

Обсудим пять перечисленных постановок. Первая из них – минимизация средних потерь – представляется вполне естественной, если все возможные потери малы по сравнению с ресурсами предприятия. В противном случае первый подход неразумен. Рассмотрим условный пример. У человека имеется 10000 рублей. Ему предлагается подбросить монету. Если выпадает «орел», то он получает 50000 рублей. Если же выпадает «цифра», он должен уплатить 20000 рублей. Стоит ли данному человеку участвовать в описанном пари? Если подсчитать математическое ожидание дохода, то, поскольку каждая сторона монеты имеет одну и ту же вероятность выпасть, равную 0,5, оно равно 50000 х 0,5 + (-20000) х 0,5 = 15000. Казалось бы, пари весьма выгодно. Однако большинство людей на него не пойдет, поскольку с вероятностью 0,5 они лишатся всего своего достояния и останутся должны 10000 рублей, другими словами, разорятся. Здесь проявляется психологическая оценка ценности рубля, зависящая от общей имеющейся суммы – 10000 рублей для человека с обычным доходом значит гораздо больше, чем те же 10000 руб. для миллиардера.

Второй подход нацелен как раз на минимизацию больших потерь, на защиту от разорения. Другое его применение – исключение катастрофических аварий, например, типа Чернобыльской. При втором подходе средние потери могут увеличиться (по сравнению с первым), зато максимальные будут контролироваться.

Третий подход нацелен на минимизацию разброса окончательных результатов. Средние потери при этом могут быть выше, чем при первом, но того, кто принимает решение, это не волнует – ему нужна максимальная определенность будущего, пусть даже ценой повышения потерь.

Четвертый подход сочетает в себе первый и третий, хотя и довольно примитивным образом. Проблема ведь в том, что управление риском в рассматриваемом случае – это по крайней мере двухкритериальная задача – желательно средние потери снизить (другими словами, математическое ожидание доходов повысить), и одновременно уменьшить показатель неопределенности – дисперсию. Хорошо известны проблемы, возникающие при многокритериальной оптимизации.

Наиболее продвинутый подход – пятый. Но для его применения необходимо построить функцию полезности. Это – большая самостоятельная задача. Обычно ее решают с помощью специально организованного эконометрического исследования.

Если неопределенность носит интервальный характер, т.е. описывается интервалами, то естественно применить методы статистики интервальных данных (как части интервальной математики), рассчитать минимальный и максимальный возможный доходы и потери, и т.д.

Разработаны различные способы уменьшения экономических рисков, связанные с выбором стратегий поведения, в частности, диверсификацией, страхованием и др. Причем эти подходы относятся не только к отдельным организациям. Так, применительно к системам налогообложения диверсификация означает использование не одного, а системы налогов, чтобы нейтрализовать действия налогоплательщиков, нацеленные на уменьшение своих налоговых платежей. Однако динамика реальных экономических систем такова, что любые формальные модели дают в лучшем случае только качественную картину. Например, не существует математических моделей, позволяющих достаточно точно спрогнозировать инфляцию вообще и даже реакцию экономики на одноразовое решение типа либерализации цен.

Необходимость применения экспертных оценок при оценке и управлении рисками. Из сказанного выше вытекает, что разнообразные формальные методы оценки рисков и управления ими во многих случаях (реально во всех нетривиальных ситуациях) не могут дать однозначных рекомендаций. В конце процесса принятия решения - всегда человек, менеджер, на котором лежит ответственность за принятое решение.

Поэтому процедуры экспертного оценивания естественно применять не только на конечном, но и на всех остальных этапах анализа рассматриваемого организацией проекта, используя при этом весь арсенал теории и практики экспертных оценок.

При этом нецелесообразно полностью отказываться от использования формально-экономических методов, например, основанных на вычислении чистых текущих (приведенных, дисконтированных) потерь и других характеристик. Использование соответствующих программных продуктов полезно для принятия обоснованных решений. Однако нельзя абсолютизировать формально-экономические методы. На основные вопросы типа: достаточно ли высоки доходы, чтобы оправдать риск, или: что лучше - быстро, но мало, или долго, но много - ответить могут только менеджеры с помощью экспертов.

Поэтому система поддержки принятия решений в организации должна сочетать формально-экономические и экспертные процедуры.

Разработка системы поддержки принятия решений в организации, нацеленной на оценивание рисков и управление ими – не простое дело. Укажем несколько проблем, связанных с подобной работой. Совершенно ясно, что система должна быть насыщена конкретными численными данными об экономическом состоянии региона, страны, возможно и мира в целом. Добыть такие данные нелегко, в частности, потому, что сводки Российского статистического агентства (ранее – Госкомстата РФ) искажены (подробнее о состоянии теории и практики статистики в России см. главу 1 и статью [15]). В частности, мы занялись изучением инфляции именно потому, что наши данные по этому показателю превышали данные Госкомстата РФ примерно в 2 раза (см. главу 7). Зарубежные источники типа учебного пособия [16] также содержат неточности. Так при составлении балансовых соотношений для макроэкономических показателей по данным [16] выяснилось, что государство должно иметь дополнительный источник доходов в несколько сотен миллиардов долларов, а доходы бизнеса имеют излишек в 30 миллиардов долларов. Другими словами, популярное учебное пособие [16] содержит данные, не согласующиеся друг с другом. Ошибка ли это авторов или сознательная фальсификация с целью скрыть от читателей характеристики американской экономики – не будем здесь обсуждать.

При решении рассматриваемых вопросов могут оказаться полезными известные публикации по методам учета финансового риска [17, 18]. При использовании широкого арсенала статистических методов необходимо учитывать особенности их развития в России и СССР, наложившие свой отпечаток на современное состояние в области кадров и литературных источников.

14.4. Подходы к управлению рисками

Чтобы управлять, надо знать цель управления и иметь возможность влиять на те характеристики риска, которые определяют степень достижения цели.

Обычно можно выделить множество допустимых управляющих воздействий, описываемое с помощью соответствующего множества параметров управления. Тогда указанная выше возможность влиять на те характеристики риска, которые определяют степень достижения цели, формализуется как выбор значения управляющего параметра. При этом управляющий параметр может быть числом, вектором, быть элементом конечного множества или иметь более сложную математическую природу.

Основная проблема – корректная формулировка цели управления рисками. Поскольку существует целый спектр различных характеристик риска (например, если потери от риска моделируются случайной величиной), то оптимизация управления риском сводится к решению задачи многокритериальной оптимизации. Например, естественной является задача одновременной минимизации среднего ущерба (математического ожидания ущерба) и разброса ущерба (дисперсии ущерба).

Как известно, для любой многокритериальной задачи целесообразно рассмотреть множество решений (т.е. значений параметра управления), оптимальных по Парето. Эти решения оптимальны в том смысле, что не существует возможных решений, которые бы превосходили бы Парето-оптимальные решения одновременно по всем критериям. Точнее, превосходили бы хотя бы по одному критерию, а по остальным были бы столь же хорошими. Теория Парето - оптимальных решений хорошо развита (см., например, монографию [19]).

Ясно, что для практической реализации надо выбирать одно из Парето - оптимальных решений. Как выбирать? Разработан целый спектр подходов, из которых выбор может быть сделан только субъективным образом. Таким образом, снова возникает необходимость применения методов экспертных оценок.

Эксперты могут выбирать непосредственно из множества Парето - оптимальных решений, если оно состоит лишь из нескольких элементов. Или же они могут выбирать ту или иную процедуру сведения многокритериальной задачи к однокритериальной.

Как пытаются решать многокритериальные задачи? Один из подходов – выбрать т.н. «главный критерий», по которому проводить оптимизацию, превратив остальные критерии в ограничения. Например, минимизировать средний ущерб, потребовав, чтобы дисперсия ущерба не превосходила заданной величины.

Иногда задача многокритериальной оптимизации допускает декомпозицию. Найдя оптимальное значение для главного критерия, можно рассмотреть область возможных значений для остальных критериев, выбрать из них второй по важности и оптимизировать по нему, и т.д.

Что же делают эксперты? Они выбирают главный критерий (или упорядочивают критерии по степени важности), задают численные значения ограничений, иногда точность или время вычислений.

Второй основной подход – это свертка многих критериев в один интегральный и переход к оптимизации по одному критерию. Например, рассматривают линейную комбинацию критериев. Строго говоря, метод «главного критерия» – один из вариантов свертки, в котором вес главного критерия равен 1, а веса остальных – 0. Построение свертки, в частности, задание весов, целесообразно осуществлять экспертными методами.

Используют также методы, основанные на соображениях устойчивости (наиболее общий подход к изучению устойчивости рассмотрен в монографии [12]). При этом рассматривают область значений управляющих параметров, в которых значение оптимизируемого одномерного критерия (главного параметра или свертки) отличается от оптимального не более чем на некоторую заданную малую величину. Такая область может быть достаточно обширной. Например, если в линейном программировании одна из граней многогранника, выделенного ограничениями, почти параллельна плоскости равных значений оптимизируемого критерия, то вся эта грань войдет в рассматриваемую область. В выделенной области можно провести оптимизацию другого параметра, и т.д. При таком подходе эксперты выбирают допустимое отклонение для основного критерия, выделяют второй критерий, задают ограничения и т.д.

Отметим, что рассмотренные выше вероятностно-статистические подходы к оцениванию рисков предполагают использование в качестве критериев таких характеристик случайной величины, как математическое ожидание, медиана, квантили, дисперсия и др. Эти характеристики определяются функцией распределения случайного ущерба, соответствующего рассматриваемому риску. При практическом использовании этого подхода перечисленные характеристики оцениваются по статистическим данным. Они оцениваются по выборке, состоящей из наблюденных величин ущерба. Согласно правилам главы 4 при этом необходимо вычислять доверительные интервалы, содержащие оцениваемые теоретические характеристики с заданной доверительной вероятностью. Таким образом, критерий, на использовании которого основана оптимизация, всегда определен лишь с некоторой точностью, а именно, лишь с точностью до полудлины доверительного интервала. Таким образом, мы приходим к постановке, рассмотренной в предыдущем абзаце.

Необходимо обратить внимание на существенное изменение ситуации в области вычислительной оптимизации за последние 40 лет. Если в 1960-е годы из-за маломощности тогдашних компьютеров большое значение имела разработка быстрых методов счета, то в настоящее время внимание переносится на постановки задач и интерпретацию результатов. По нашим наблюдениям, это объясняется не только наличием различных программных продуктов по оптимизации, но и тем, что почти любую практическую задачу оптимизации можно решить простейшими методами типа переборных (перебирая возможные значения управляющих параметров с маленьким шагом), либо методом случайного поиска, поскольку быстродействие современных компьютеров позволяет это сделать.

Цитированная литература

1. Бестужев-Лада И.В. Окно в будущее: Современные проблемы социального прогнозирования. - М.: Мысль, 1970. - 269 с.
2. Гаврилец Ю.Н. Социально-экономическое планирование: Системы и модели. - М.: Экономика, 1974. - 174 с.
3. Загоруйко Н.Г. Эмпирическое предсказание. - Новосибирск: Наука, 1979. - 124 с.
4. Нейлор Т. Машинные имитационные эксперименты с моделями экономических систем. - М.: Мир, 1975.
5. Сидельников Ю.В. Теория и организация экспертного прогнозирования. - М.: ИМЭМО АН СССР, 1990. - 196 с.
6. Тейл Г. Эконометрические прогнозы и принятие решений. - М.: Статистика, 1971. - 488 с.
7. Френкель А.А. Математические методы анализа динамики и прогнозирования производительности труда. - М.: Экономика, 1972. - 190 с.
8. Четыркин Е.М. Статистические методы прогнозирования. - М.: Статистика, 1977.
9. Янч Э. Прогнозирование научно-технического прогресса. - М.: Прогресс, 1990. - 568 с.
10. Орлов А.И. Задачи оптимизации и нечеткие переменные. - М.: Знание, 1980. - 64 с.
11. Жихарев В.Н., Орлов А.И. Законы больших чисел и состоятельность статистических оценок в пространствах произвольной природы. – В сб.: Статистические методы оценивания и проверки гипотез. Межвузовский сборник научных трудов. – Пермь: Изд-во Пермского государственного университета, 1998. С.65-84.
12. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979. - 296 с.
13. Орлов А.И. Сценарии социально-экономического развития России до 2007 г. - Журнал «Обозреватель-Observer». 1999. No.10 (117). С.47-50.
14. Пиндайк Р., Рубинфельд Д. Микроэкономика. - М.: "Экономика" - "Дело", 1992.
15. Орлов А.И. О перестройке статистической науки и ее применений - Вестник статистики, 1990, № 1, с.65-71.
16. Макконнелл К.Р., Брю С.Л. Экономикс: Принципы, проблемы и политика. В 2 т.: Пер. с англ. 11-го изд. - М.: Республика, 1992.
17. Первозванский А.А., Первозванская Т.Н. Финансовый рынок: расчет и риск. - М.: Инфра-М, 1994.
18. Четыркин Е.М. Методы экономических расчетов. - М.: Гамма, 1992.
19. Подиновский В.В., Ногин В.Д. Парето-оптимальные решения многокритериальных задач. - М.: Наука, 1982.

Глава 15. Современные эконометрические методы

Подведем итоги и наметим перспективы развития эконометрических методов. В настоящей главе дается критический анализ современного состояния эконометрики и прикладной статистики, обсуждаются тенденции развития статистических методов, выделяются пять основных «точек роста». Рассматриваются основные нерешенные эконометрические проблемы. В связи с внедрением современных эконометрических методов обосновывается полезность понятия "высокие статистические технологии".

15.1. О развитии эконометрических методов

Современное состояние в эконометрике, как и в других областях, определяется прошлым. Кратко рассмотрим историю эконометрики и прикладной статистики, начав с их практической пользы.

Что дает прикладная статистика народному хозяйству? Так называлась статья [1], в которой приводились многочисленные примеры успешного использования методов эконометрики и прикладной математической статистики при решении практических задач. Обширный перечень примеров приведен в предыдущих главах настоящей книги. Его можно продолжать практически безгранично. Так, в любом номере журнала "Заводская лаборатория" есть работы, в которых те или иные методы эконометрики и прикладной статистики применяются для решения прикладных технико-экономических задач.

Поэтому бесспорно совершенно, что методы эконометрики и прикладной статистики успешно применяются в различных отраслях народного хозяйства, практически во всех областях науки. Согласно докладу [2], в 1988 г. затраты в СССР на статистический анализ данных оценивались в 2 миллиарда рублей ежегодно.

Большая практическая значимость эконометрики и прикладной статистики, особенно в экономике, менеджменте, технических исследованиях и разработках, оправдывает целесообразность развития их методологии, в которых эти области научной и прикладной деятельности рассматривалась бы как целое, "с высоты птичьего полета". Чтобы иметь возможность обсуждения тенденций развития эконометрики и статистических методов в XXI веке, необходимо хотя бы кратко рассмотреть их историю.

Об истории эконометрики и прикладной статистики. Типовые примеры раннего этапа применения статистических методов описаны в Ветхом Завете (см., например, Книгу Чисел). С математической точки зрения они сводились к подсчетам числа попаданий значений наблюдаемых признаков в определенные градации. В дальнейшем результаты стали представлять в виде таблиц и диаграмм, как это и сейчас делают Госкомстат РФ (Российское статистическое агентство). Надо признать, что по сравнению с Ветхим Заветом есть прогресс - в Библии не было таблиц. Однако нет продвижения по сравнению с работами российских статистиков конца девятнадцатого - начала двадцатого века (типовой монографией тех времен можно считать книгу [3], которая в настоящее время ещё легко доступна).

Сразу после возникновения теории вероятностей (Паскаль, Ферма, 17 век) вероятностные модели стали использоваться при обработке статистических данных. Например, изучалась частота рождения мальчиков и девочек, было установлено отличие вероятности рождения мальчика от 0.5, анализировались причины того, что в парижских приютах эта вероятность не та, что в самом Париже, и т.д. Имеется достаточно много публикаций по истории теории вероятностей, однако в некоторых из них имеются неточные утверждения, что заставило одного из крупнейших ученых XX в. академика Украинской АН Б.В. Гнеденко включить в очередное издание своего курса [4] главу по истории математики случайного.

В 1794 г. (по другим данным - в 1795 г.) К. Гаусс разработал метод наименьших квадратов, один из наиболее популярных ныне статистических методов (см. главу 5 выше), и применил его при расчете орбиты астероида Церера - для борьбы с ошибками астрономических наблюдений. В X1X веке заметный вклад в развитие практической статистики внес бельгиец А. Кетле, на основе анализа большого числа реальных данных показавший устойчивость относительных статистических показателей, таких, как доля самоубийств среди всех смертей. Интересно, что основные идеи статистического приемочного контроля и сертификации продукции обсуждались академиком М.В. Остроградским и применялись в российской армии ещё в середине X1X в.. Статистические методы управления качеством, сертификации и классификации продукции и сейчас весьма актуальны (см. главу 13 выше).

Современный этап развития прикладной статистики можно отсчитывать с 1900 г., когда англичанин К. Пирсон основан журнал "Biometrika". Первая треть XX в. прошла под знаком параметрической статистики. Изучались методы, основанные на анализе данных из параметрических семейств распределений, описываемых кривыми из т.н. семейства Пирсона. Наиболее популярным было нормальное (гауссово) распределение. Для проверки гипотез использовались критерии Пирсона, Стьюдента, Фишера. Были предложены метод максимального правдоподобия, дисперсионный анализ, сформулированы основные идеи планирования эксперимента.

Разработанную в первой трети XX в. теорию называем параметрической статистикой, поскольку ее основной объект изучения - это выборки из распределений, описываемых одним или небольшим числом параметров. Наиболее общим является семейство кривых Пирсона, задаваемых четырьмя параметрами. Как правило, нельзя указать каких-либо веских причин, по которым конкретное распределение результатов наблюдений должно входить в то или иное параметрическое семейство (подробнее см. начало главы 4). Исключения хорошо известны: если вероятностная модель предусматривает суммирование независимых случайных величин, то сумму естественно описывать нормальным распределением; если же в модели рассматривается произведение таких величин, то итог, видимо, приближается логарифмически нормальным распределением, и т.д. Однако в подавляющем большинстве реальных ситуаций подобных моделей нет, и приближение реального распределения с помощью кривых из семейства Пирсона или его подсемейств - чисто формальная операция.

Именно из таких соображений критиковал параметрическую статистику академик АН СССР С.Н. Бернштейн в 1927 г. в своем докладе на Всероссийском съезде математиков [5]. Однако эта теория, к сожалению, до сих пор остается основой преподавания статистических методов и продолжает использоваться основной массой прикладников, остающихся далекими от новых веяний в статистике. Почему так происходит? Чтобы попытаться ответить на этот вопрос, обратимся к одной из статистических наук - наукометрии, в которой статистическими методами анализируется развитие научных исследований.

Наукометрия прикладной статистики. Проведенный несколько лет назад анализ прикладной статистики как области научно-практической деятельности (в рамках движения за создание Всесоюзной статистической ассоциации, учрежденной в 1990 г.) показал, в частности, что актуальными для специалистов в настоящее время являются не менее чем 100 тысяч публикаций (подробнее см. статьи [6,7]). Реально же каждый из них знаком с существенно меньшим количеством книг и статей. Так, в наиболее солидном и обширном из научных изданий в области эконометрики и прикладной статистики - трехтомнике Кендалла и Стьюарта [8-10] всего около 2 тысяч литературных ссылок. При всей очевидности соображений о многократном дублировании ценных идей в различных публикациях приходится признать, что каждый специалист по эконометрике и прикладной статистике владеет лишь небольшой частью накопленных в этой области знаний. Не удивительно, что приходится постоянно сталкиваться с игнорированием или повторением ранее полученных результатов, с уходом в тупиковые (с точки зрения практики) направления исследований, с беспомощностью при обращении к реальным данным, и т.д. Все это - одно из проявлений адапционного механизма торможения развития науки, которая оказывается не в состоянии даже осмыслить ранее полученные результаты. Об этом печальном явлении еще более 30 лет назад писали В.В.Налимов и другие науковеды (см., например, [11]).

Традиционный предрассудок состоит в том, что каждый новый результат, полученный исследователем - это кирпич, вложенный в непрерывно растущее здание науки, который непременно будет проанализирован и использован научным сообществом. Реальная ситуация - совсем иная. Как известно, большинство книг в центральных библиотеках никто никогда не читал. Так что с новым результатом, скорее всего, познакомятся лишь несколько человек, да и то поверхностно, а использовать его будут, в лучшем случае, сам автор в дальнейших работах и его ученики.

Основа профессиональных знаний экономиста, менеджера, исследователя и инженера закладывается в период обучения. Затем они пополняются в том узком направлении, в котором работает специалист. Следующий этап - их тиражирование новому поколению. В результате вузовские учебники отстоят от современного развития на десятки лет. Так, учебники по математической статистике, по нашей экспертной оценке, в основном соответствуют 40-60-м годам XX в. А потому тем же годам соответствует по своему научному и методологическому уровню большинство вновь публикуемых исследований и тем более - прикладных работ. Одновременно приходится признать, что результаты, которым не повезло, поскольку они не вошли в учебники, независимо от их научной и (или) прикладной ценности почти все забываются.

Активно продолжается развитие тупиковых направлений. Психологически это понятно. Приведу пример из своего опыта. В свое время по заказу Госстандарта я разработал методы оценки параметров гамма-распределения (см. государственный стандарт [12]. Поэтому мне близки и интересны работы по оцениванию параметров по выборкам из распределений, принадлежащих тем или иным параметрическим семействам, понятия функции максимального правдоподобия, эффективности оценок, использование неравенства Рао-Крамера и т.д. К сожалению, я знаю, что это - тупиковая ветвь, поскольку реальные данные не подчиняются каким-либо параметрическим семействам, надо применять иные статистические методы, о которых речь пойдет ниже. Понятно, что специалистам по параметрической статистике, потратившим многие годы на совершенствование в своей области, психологически трудно согласиться с подобным утверждением. В том числе и мне было трудно перейти на другую позицию, отраженную в настоящей книге и исходящую из потребностей прикладных работ.

15.2. Точки роста

Отечественная литература по эконометрике и прикладной статистике столь же необозрима, как и мировая. Только в секции "Математические методы исследования" журнала "Заводская лаборатория" с 1960-х годов опубликовано более 1000 статей. Не будем даже пытаться перечислять коллективы исследователей или основные монографии в этой области. Отметим только одно издание. По нашему мнению, наилучшей отечественной книгой по прикладной статистике является сборник статистических таблиц Л.Н. Большева и Н.В.Смирнова [13] с подробными комментариями, играющими роль сжатого учебника и справочника.

Основная цель настоящей главы - выделить и обсудить "точки роста" эконометрики и прикладной статистики, те их направления, которые представляются перспективными в будущем, в XXI веке, но пока в большинстве учебных изданий отодвинуты на задний план традиционными постановками.

При описании современного этапа развития эконометрических и статистических методов целесообразно выделить пять актуальных направлений, в которых развивается современная прикладная статистика, т.е. пять "точек роста": непараметрика (т.е. непараметрическая статистика), робастность, бутстреп, статистика интервальных данных, статистика нечисловых данных (в несколько иной терминологии - статистика объектов нечисловой природы). Обсудим их.

Непараметрическая статистика (см. также главу 4). В первой трети XX в., одновременно с параметрической статистикой, в работах Спирмена и Кендалла появились первые непараметрические методы, основанные на коэффициентах ранговой корреляции, носящих ныне имена этих статистиков. Но непараметрика, не делающая нереалистических предположений о том, что функции распределения результатов наблюдений принадлежат тем или иным параметрическим семействам распределений, стала заметной частью статистики лишь со второй трети XX века. В 30-е годы появились работы А.Н.Колмогорова и Н.В.Смирнова, предложивших и изучивших статистические критерии, носящие в настоящее время их имена. Эти критерии основаны на использовании так называемого эмпирического процесса. (Как известно, эмпирический процесс – это разность между эмпирической и теоретической функциями распределения, умноженная на квадратный корень из объема выборки.) В работе А.Н.Колмогорова 1933 г. изучено предельное распределение супремума модуля эмпирического процесса, называемого сейчас критерием Колмогорова. Затем Н.В. Смирнов исследовал супремум и инфимум эмпирического процесса, а также интеграл (по теоретической функции распределения) квадрата эмпирического процесса.

Следует отметить, что встречающееся иногда в литературе словосочетание "критерий Колмогорова-Смирнова" некорректно, поскольку эти два статистика никогда не печатались вместе и не изучали один и тот же критерий схожими методами. Корректно сочетание "критерий типа Колмогорова-Смирнова", применяемое для обозначения критериев, основанных на использовании супремума функций от эмпирического процесса.

После второй мировой войны развитие непараметрической статистики пошло быстрыми темпами. Большую роль сыграли работы Ф. Вилкоксона и его школы. К настоящему времени с помощью непараметрических методов можно решать практически тот же круг статистических задач, что и с помощью параметрических. Однако для обеспечения широкого внедрения непараметрических методов необходимо провести еще целый комплекс теоретических и пилотных (т.е. пробных) прикладных работ. Все большую роль играют непараметрические оценки плотности, непараметрические методы регрессии и распознавания образов (дискриминантного анализа). В нашей стране непараметрические методы получили достаточно большую известность после выхода в 1965 г. первого издания упомянутого выше сборника статистических таблиц Л.Н. Большева и Н.В.Смирнова [13], содержащего подробные таблицы для основных непараметрических критериев.

Тем не менее параметрические методы всё еще популярнее непараметрических, особенно среди тех прикладников, кто слабо знаком со статистическими методами. Неоднократно публиковались (см. начало гл.4) экспериментальные данные, свидетельствующие о том, что распределения реально наблюдаемых случайных величин, в частности, ошибок измерения, в подавляющем большинстве случаев отличны от нормальных (гауссовских). Тем не менее теоретики продолжают строить и изучать статистические модели, основанные на гауссовости, а практики - применять подобные методы и модели. Другими словами, "ищут под фонарем, а не там, где потеряли".

Устойчивость статистических процедур (робастность) (см. также главу 10). Если в параметрических постановках на данных накладываются слишком жесткие требования - их функции распределения должны принадлежать определенному параметрическому семейству, то в непараметрических, наоборот, излишне слабые - требуется лишь, чтобы функции распределения были непрерывны. При этом игнорируется априорная информация о том, каков "примерный вид" распределения. Априори можно ожидать, что учет этого "примерного вида" улучшит показатели качества статистических процедур. Развитием этой идеи является теория устойчивости (робастности) статистических процедур, в которой предполагается, что распределение исходных данных мало отличается от некоторого параметрического семейства. За рубежом эту теорию разрабатывали П.Хубер, Ф.Хампель и многие другие. Из монографий на русском языке, трактующих о робастности и устойчивости статистических процедур, самой ранней и наиболее общей была книга [14], следующей - монография [15]. Частными случаями реализации идеи робастности (устойчивости) статистических процедур являются статистика объектов нечисловой природы (см. главу 8) и статистика интервальных данных (см. главу 3)..

Имеется большое разнообразие моделей робастности в зависимости от того, какие именно отклонения от заданного параметрического семейства допускаются. Среди теоретиков наиболее популярной оказалась модель выбросов, в которой исходная выборка "засоряется" малым числом "выбросов", имеющих принципиально иное распределение. Однако эта модель представляется "тупиковой", поскольку в большинстве случаев большие выбросы либо невозможны из-за ограниченности шкалы прибора либо интервала изменения измеряемой величины, либо от них можно избавиться, применяя лишь статистики, построенные по центральной части вариационного ряда. Кроме того, в подобных моделях обычно считается известной частота засорения, что в сочетании со сказанным выше делает их малопригодными для практического использования.

Более перспективным представляется, например, модель малых отклонений распределений, в которой расстояние между распределением каждого элемента выборки и базовым распределением не превосходит заданной малой величины, и модель статистики интервальных данных.

Бутстреп (размножение выборок) (см. также главу 11). Другое из упомянутых выше направлений - бутстреп - связано с интенсивным использованием возможностей вычислительной техники. Основная идея состоит в том, чтобы теоретическое исследование заменить вычислительным экспериментом. Вместо описания выборки распределением из параметрического семейства строим большое число "похожих" выборок, т.е. "размножаем" выборку. Затем вместо оценивания характеристик (и параметров) и проверки гипотез на основе свойств теоретического распределения решаем эти задачи вычислительным методом, рассчитывая интересующие нас статистики по каждой из "похожих" выборок и анализируя полученные при этом распределения. Например, вместо того, чтобы теоретическим путем находить распределение статистики, доверительные интервалы и другие характеристики, моделируют большое число выборок, похожих на исходную, затем рассчитывают соответствующие значения интересующей исследователя статистики и изучают их эмпирическое распределение. Квантили этого распределения задают доверительные интервалы, и т.д.

Термин "бутстреп" мгновенно получил широкую известность после первой же статьи Б.Эфрона 1979 г. по этой тематике. Он сразу же стал обсуждаться в массе публикаций, в том числе и научно-популярных. В "Заводской лаборатории" № 10 за 1987 г. была помещена подборка статей по бутстрепу. На русском языке выпущен сборник статей Б. Эфрона [16]. Основная идея бутстрепа по Б. Эфрону состоит в том, что методом Монте-Карло (статистических испытаний) многократно извлекаются выборки из эмпирического распределения. Эти выборки, естественно, являются вариантами исходной, напоминают ее.
Сама по себе идея "размножения выборок" была известна гораздо раньше. Одна из статей Б. Эфрона в сборнике [16] называется так: "Бутстреп-методы: новый взгляд на метод складного ножа". Упомянутый "метод складного ножа" (jackknife) предложен М. Кенуем еще в 1949 г., за 30 лет до появления статьи Б.Эфрона. "Размножение выборок" при этом осуществляется путем исключения одного наблюдения. Таким образом для выборки объема n получаем n "похожих" на нее выборок объема $(n - 1)$ каждая. Если же исключать по 2 наблюдения, то число "похожих" выборок возрастает до $n (n - 1) / 2$ объема $(n - 2)$ каждая.

Преимущества и недостатки бутстрепа как статистического метода обсуждались в главе 11 выше. Там же приводится информация о ряде аналогичных методов. Необходимо подчеркнуть, что бутстреп по Эфрону - лишь один из вариантов методов "размножения выборки" (resampling), и, на наш взгляд, не самый удачный. Метод "складного ножа" представляется более полезным. На его основе можно сформулировать следующую простую практическую рекомендацию.

Предположим, что Вы по выборке делаете какие-либо статистические выводы. Вы хотите узнать также, насколько эти выводы устойчивы. Если у Вас есть другие (контрольные) выборки, описывающие то же явление, то Вы можете применить к ним ту же статистическую процедуру и сравнить результаты. А если таких выборок нет? Тогда Вы можете их построить искусственно. Берете исходную выборку и исключаете один элемент. Получаете похожую выборку (она взята из того же распределения, только объем на единицу меньше). Затем возвращаете этот элемент выборки и исключаете другой. Получаете вторую похожую выборку. Поступив таким образом со всеми элементами исходной выборки, получаете столько выборок, похожих на исходную, каков ее объем. Остается обработать их тем же способом, что и исходную, и изучить устойчивость получаемых выводов - разброс оценок параметров, частоты принятия или отклонения гипотез и т.д.

Можно изменять не выборку, а сами данные. Поскольку всегда имеются погрешности измерения, то реальные данные - это не числа, а интервалы (результат измерения плюс-минус погрешность). Нужна статистическая теория анализа таких данных.

Статистика интервальных данных (см. также главу 9). Перспективное и быстро развивающееся направление последних лет - прикладная математическая статистика интервальных данных. Речь идет о развитии методов математической статистики в ситуации, когда статистические данные - не числа, а интервалы, в частности, порожденные наложением ошибок измерения на значения случайных величин.

Статистика интервальных данных идейно связана с интервальной математикой, в которой в роли чисел выступают интервалы. Это направление математики является дальнейшим развитием всем известных правил приближенных вычислений, посвященных выражению погрешностей суммы, разности, произведения, частного через погрешности тех чисел, над которыми осуществляются перечисленные операции. К настоящему времени удалось решить, в частности, ряд задач теории интервальных дифференциальных уравнений, в которых коэффициенты, начальные условия и решения описываются с помощью интервалов.

Одна из ведущих научных школ в области статистики интервальных данных - это школа проф. А.П. Вощинина, активно работающая с конца 70-х годов. В частности, изучены проблемы регрессионного анализа, планирования эксперимента, сравнения альтернатив и принятия решений в условиях интервальной неопределенности.

Рассмотрим другое направление в статистике интервальных данных, которое также представляется перспективным. В нем развиваются асимптотические методы статистического анализа интервальных данных при больших объемах выборок и малых погрешностях измерений. В отличие от классической математической статистики, сначала устремляется к бесконечности объем выборки и только потом - уменьшаются до нуля погрешности. В частности, с помощью такой асимптотики были сформулированы правила выбора метода оценивания параметров гамма-распределения в ГОСТ 11.011-83 [12].

В рамках рассматриваемого научного направления, разработана общая схема исследования, включающая расчет нотны (максимально возможного отклонения статистики, вызванного интервальностью исходных данных) и рационального объема выборки (превышение которого не дает существенного повышения точности оценивания). Она применена к оцениванию математического ожидания, дисперсии, коэффициента вариации, параметров гамма-распределения и характеристик аддитивных статистик, при проверке гипотез о параметрах нормального распределения, в т.ч. с помощью критерия Стьюдента, а также гипотезы однородности с помощью критерия Смирнова. Разработаны подходы к рассмотрению интервальных данных в основных постановках регрессионного, дискриминантного и кластерного анализов. В частности, изучено влияние погрешностей измерений и наблюдений на свойства алгоритмов регрессионного анализа, разработаны способы расчета нотн и рациональных объемов выборок, введены и исследованы новые понятия многомерных и асимптотических нотн, доказаны соответствующие предельные теоремы. Начата разработка интервального дискриминантного анализа, в частности, рассмотрено влияние интервальности данных на введенный в главе 5 показатель качества классификации. Изучено асимптотическое поведение оценок метода моментов и оценок максимального правдоподобия (а также более общих - оценок минимального контраста), проведено асимптотическое сравнение этих методов в случае интервальных данных. Найдены общие условия, при которых, в отличие от классической математической статистики, метод моментов дает более точные оценки, чем метод максимального правдоподобия.

В области асимптотической математической статистики интервальных данных российская наука имеет мировой приоритет. Развертывание работ по рассматриваемой тематике позволит закрепить этот приоритет, получить теоретические результаты, основополагающие в новой области математической статистики и необходимые для обоснованного статистического анализа почти всех типов данных. Со временем во все виды статистического программного обеспечения должны быть включены алгоритмы интервальной статистики, "параллельные" обычно используемым алгоритмам прикладной математической статистики. Это позволит в явном виде учесть наличие погрешностей у результатов наблюдений, сблизить позиции метрологов и статистиков.

Статистика объектов нечисловой природы как часть прикладной статистики. Согласно общепринятой в настоящее время классификации статистических методов прикладная статистика делится на следующие четыре области:

статистика (числовых) случайных величин (см. главу 4),

многомерный статистический анализ (см. главу 5),

статистика временных рядов и случайных процессов (см. главу 6),

статистика объектов нечисловой природы (см. главу 8),.

Первые три из этих областей являются классическими. Они были хорошо известны еще в первой половине XX в. Остановимся на четвертой, сравнительно недавно вошедшей в массовое сознание специалистов. Ее именуют также статистикой нечисловых данных или попросту нечисловой статистикой. Анализ динамики развития эконометрики и прикладной статистики приводит к выводу, что в XXI в. она станет центральной областью прикладной статистики, поскольку содержит наиболее общие подходы и результаты.

Исходный объект в прикладной математической статистике - это выборка. В вероятностной теории статистики выборка - это совокупность независимых одинаково распределенных случайных элементов. Какова природа этих элементов? В классической математической статистике элементы выборки - это числа. В многомерном статистическом анализе - вектора. А в нечисловой статистике элементы выборки - это объекты нечисловой природы, которые нельзя складывать и умножать на числа. Другими словами, объекты нечисловой природы лежат в пространствах, не имеющих векторной структуры.

Примерами объектов нечисловой природы являются (подробнее см. главу 8):

значения качественных признаков, т.е. результаты кодировки объектов с помощью заданного перечня категорий (градаций);

упорядочения (ранжировки) экспертами образцов продукции (при оценке её технического уровня и конкурентоспособности)) или заявок на проведение научных работ (при проведении конкурсов на выделение грантов);

классификации, т.е. разбиения объектов на группы сходных между собой (кластеры);

толерантности, т.е. бинарные отношения, описывающие сходство объектов между собой, например, сходства тематики научных работ, оцениваемого экспертами с целью рационального формирования экспертных советов внутри определенной области науки;

результаты парных сравнений или контроля качества продукции по альтернативному признаку ("годен" - "брак"), т.е. последовательности из 0 и 1;

множества (обычные или нечеткие), например, зоны, пораженные коррозией, или перечни возможных причин аварии, составленные экспертами независимо друг от друга;

слова, предложения, тексты;

вектора, координаты которых - совокупность значений разнотипных признаков, например, результат составления статистического отчета о научно-технической деятельности (т.н. форма № 1-наука) или заполненная компьютеризированная история болезни, в которой часть признаков носит качественный характер, а часть - количественный; ответы на вопросы экспертной, маркетинговой или социологической анкеты, часть из которых носит количественный характер (возможно, интервальный), часть сводится к выбору одной из нескольких подсказок, а часть представляет собой тексты; и т.д.

Интервальные данные (см. выше) тоже можно рассматривать как пример объектов нечисловой природы, а именно, как частный случай нечетких множеств.

С начала 70-х годов под влиянием запросов прикладных исследований в социально-экономических, технических, медицинских науках в России активно развивается статистика объектов нечисловой природы, известная также как статистика нечисловых данных или нечисловая статистика. В создании этой сравнительно новой области эконометрики и прикладной математической статистики приоритет принадлежит российским ученым.

Большую роль сыграл основанный в 1973 г. научный семинар "Экспертные оценки и анализ данных". В 60-е годы советское научное сообщество стало интересоваться методами экспертных оценок (об их истории и современном состоянии см. главу 12). Как следствие, началось знакомство с конкретными математизированными теориями, связанными с этими методами. Речь идет о репрезентативной теории измерений, ставшей известной в нашей стране по статье П.Суппеса и Дж.Зинеса в сборнике [17] и книге И.Пфанцагля [18], о теории нечеткости, современный этап которой начался с работ Л.А.Заде [19], теории парных сравнений, описанной в монографии Г.Дэвида [20]. К этому кругу идей примыкают теория случайных множеств (см., например, книгу Ж.Матерона [21]) и методы многомерного шкалирования (описаны, в частности, в монографиях А.Ю.Терехиной [22] и В.Т.Перекреста [23]). Но наибольшее влияние оказали идеи Дж.Кемени, который аксиоматически ввел расстояние между ранжировками (теперь оно именуется в литературе расстоянием Кемени) и предложил использовать в качестве средней величины решение оптимизационной задачи (теперь - медиана Кемени). Его скромная книжка [24], написанная в соавторстве с Дж.Снеллом, породила большой поток исследований.

В течение 70-х годов на основе запросов теории экспертных оценок (а также социологии, экономики, техники и медицины) развивались конкретные направления статистики объектов нечисловой природы. Были установлены связи между конкретными видами таких объектов, разработаны для них вероятностные модели (см. главу 8). Научные итоги этого периода подведены в монографиях [14,25,26]).

Следующий этап - выделение статистики объектов нечисловой природы в качестве самостоятельного направления в эконометрике и прикладной статистике, ядром которого являются методы статистического анализа данных произвольной природы. Программа развития этого нового научного направления впервые была сформулирована в статье [27]. Реализация этой программы была осуществлена в 80-е годы. Для работ этого периода характерна сосредоточенность на внутренних проблемах нечисловой статистики. Ссылки на конкретные монографии, сборники, статьи и иные публикации нескольких десятков авторов приведены в главе 8. Отметим лишь сборник научных статей [28], полностью посвященный нечисловой статистике.

К 90-м годам статистика объектов нечисловой природы с теоретической точки зрения была достаточно хорошо развита, основные идеи, подходы и методы были разработаны и изучены математически, в частности, доказано достаточно много теорем. Однако она оставалась недостаточно апробированной на практике. Это было связано как с ее сравнительной молодостью, так и с общеизвестными особенностями организации науки в

80-е годы, когда отсутствовали достаточные стимулы к тому, чтобы теоретики занялись широким внедрением своих результатов. И в 90-е годы наступило время от математико-статистических исследований перейти к применению полученных результатов на практике.

Следует отметить, что в статистике объектов нечисловой природы, как и в других областях эконометрики, прикладной математической статистики и прикладной математики вообще, одна и та же математическая схема может с успехом применяться и в технических исследованиях, и в менеджменте, и в экономике, и в геологии, и в медицине, и в социологии, и для анализа экспертных оценок, и во многих иных областях, а потому ее лучше всего формулировать и изучать в наиболее общем виде, для объектов произвольной природы.

Основные идеи статистики объектов нечисловой природы. В чем принципиальная новизна нечисловой статистики? Для классической математической статистики характерна операция сложения. При расчете выборочных характеристик распределения (выборочное среднее арифметическое, выборочная дисперсия и др.), в регрессионном анализе и других областях этой научной дисциплины постоянно используются суммы. Математический аппарат - законы больших чисел, Центральная предельная теорема и другие теоремы - нацелены на изучение сумм. В нечисловой же статистике нельзя использовать операцию сложения, поскольку элементы выборки лежат в пространствах, где нет операции сложения. Методы обработки нечисловых данных основаны на принципиально ином математическом аппарате - на применении различных расстояний в пространствах объектов нечисловой природы.

Кратко рассмотрим несколько идей, развиваемых в статистике объектов нечисловой природы для данных, лежащих в пространствах произвольного вида. Решаются классические задачи описания данных, оценивания, проверки гипотез - но для неклассических данных, а потому неклассическими методами.

Первой обсудим проблему определения средних величин. В рамках репрезентативной теории измерений удается указать вид средних величин, соответствующих тем или иным шкалам измерения (см. главу 3). В классической математической статистике средние величины вводят с помощью операций сложения (выборочное среднее арифметическое, математическое ожидание) или упорядочения (выборочная и теоретическая медианы). В пространствах произвольной природы средние значения нельзя определить с помощью операций сложения или упорядочения. Теоретические и эмпирические средние приходится вводить как решения экстремальных задач. Для теоретического среднего это - задача минимизации математического ожидания (в классическом смысле) расстояния от случайного элемента со значениями в рассматриваемом пространстве до фиксированной точки этого пространства (минимизируется указанная функция от этой точки). Для эмпирического среднего математическое ожидание берется по эмпирическому распределению, т.е. берется сумма расстояний от некоторой точки до элементов выборки и затем минимизируется по этой точке. При этом как эмпирическое, так и теоретическое средние как решения экстремальных задач могут быть не единственным элементом пространства, а состоять из множества таких элементов, которое может оказаться и пустым. Тем не менее удалось сформулировать и доказать законы больших чисел для средних величин, определенных указанным образом, т.е. установить сходимость эмпирических средних к теоретическим .

Оказалось, что методы доказательства законов больших чисел допускают существенно более широкую область применения, чем та, для которой они были разработаны. А именно, удалось изучить асимптотику решений экстремальных статистических задач, к которым, как известно, сводится большинство постановок прикладной статистики. В частности, кроме законов больших чисел установлена и состоятельность оценок минимального контраста, в том числе оценок максимального правдоподобия и робастных оценок. К настоящему времени подобные оценки изучены также и в интервальной статистике.

В статистике в пространствах произвольной природы большую роль играют непараметрические оценки плотности, используемые, в частности, в различных алгоритмах регрессионного, дискриминантного, кластерного анализов. В нечисловой статистике предложен и изучен ряд типов непараметрических оценок плотности в пространствах произвольной природы, в частности, доказана их состоятельность, изучена скорость сходимости и установлен примечательный факт совпадения наилучшей скорости сходимости в произвольном случае с той, которая имеет быть в классической теории для числовых случайных величин.

Дискриминантный, кластерный, регрессионный анализы в пространствах произвольной природы основаны либо на параметрической теории - и тогда применяется подход, связанный с асимптотикой решения экстремальных статистических задач - либо на непараметрической теории - и тогда используются алгоритмы на основе непараметрических оценок плотности.

Для проверки гипотез могут быть использованы статистики интегрального типа, в частности, типа омега-квадрат. Любопытно, что предельная теория таких статистик, построенная первоначально в классической постановке [29], приобрела естественный (завершенный, изящный) вид именно для пространств произвольного вида [30], поскольку при этом удалось провести рассуждения, опираясь на базовые математические соотношения, а не на те частные (с общей точки зрения), что были связаны с конечномерным пространством.

Представляют практический интерес результаты, связанные с конкретными областями статистики объектов нечисловой природы, в частности, со статистикой нечетких множеств, развитой в книге [31], и со статистикой случайных множеств [14] (следует отметить, что теория нечетких множеств в определенном смысле сводится к теории случайных множеств [14,31]), с непараметрической теорией парных сравнений, с аксиоматическим введением метрик в конкретных пространствах объектов нечисловой природы [28], и с рядом других конкретных постановок (см. главу 8).

Для анализа нечисловых, в частности, экспертных данных весьма важны методы классификации. С другой стороны, наиболее естественно ставить и решать задачи классификации, основанные на использовании расстояний или показателей различия, в рамках статистики объектов нечисловой природы. Это касается как распознавания образов с учителем (другими словами, дискриминантного анализа), так и распознавания образов без учителя (т.е. кластерного анализа). Современное состояние дискриминантного и кластерного анализа с точки зрения статистики объектов нечисловой природы отражено в главе 5.

Статистические методы анализа нечисловых данных особенно хорошо приспособлены для применения в экономике, социологии и экспертных оценках, поскольку в этих областях от 50% до 90% данных являются нечисловыми.

Другие точки роста. Выше рассмотрены пять "точек роста" эконометрики и прикладной статистики. Разумеется, они не исчерпывают все многообразие фронта научных исследований в рассматриваемых областях. Кроме того, в настоящей главе почти не затронуты разнообразные применения эконометрических и статистических методов в конкретных прикладных исследованиях и разработках. Много интересных проблем есть в планировании экспериментов, особенно кинетических (см., например, статью [31]), при анализе проблем надежности, в новых статистических методах управления качеством продукции (см. главу 13), в том числе в связи с идеями Г. Тагути, при анализе рисков (см. главу 14), в вопросах экологии и безопасности и др.

В течение последних более чем 60 лет в России наблюдается огромный разрыв между государственной статистикой и научным сообществом специалистов по статистическим методам (подробнее об этом см. статью [7]). В учебнике по истории статистики [32] даже не упоминаются имена членов-корреспондентов АН СССР Н.В.Смирнова и Л.Н. Большева! А ведь они – единственные представители именно математической статистики как таковой в Академии наук в XX в. (еще ряд членов Академии наук имели математическую статистику среди своих интересов, но Н.В. Смирнов и Л.Н. Большев занимались практически только ею). Поэтому нет ничего удивительного в том, что тенденции развития современной эконометрики и прикладной математической статистики столь же мало обсуждаются отечественными авторами, как и ее история.

15.3. О некоторых нерешенных вопросах эконометрики и прикладной статистики

За последние 30 лет выявился целый ряд нерешенных вопросов эконометрики и прикладной статистики , как чисто научных, так и научно-организационных. Обсудим пять из них:

влияние отклонений от традиционных предпосылок (вероятностно-статистических моделей) на свойства эконометрических и статистических процедур;

оправданность использования асимптотических теоретических результатов эконометрики и прикладной математической статистики при конечных объемах выборок;

формулировки и обоснования правил выбора одного из многих критериев для проверки конкретной гипотезы;

конкретные способы организации теоретических работ в области эконометрики и прикладной математической статистики;

организация и проведение прикладных работ с использованием методов эконометрики и прикладной математической статистики.

Настоящий раздел отнюдь не претендует на решение перечисленных вопросов. Его цель гораздо скромнее - обратить внимание на существование ряда нерешенных вопросов в надежде, что коллективными усилиями удастся продвинуться в их решении.

Влияние отклонений от традиционных предпосылок. В вероятностной теории статистических методов выборка обычно моделируется как конечная последовательность независимых одинаково распределенных случайных величин или векторов. Часто предполагается, что эти величины (вектора) имеют нормальное распределение.

На основе сформулированных классических предпосылок построено огромное здание классической математической статистики с большим числом теорем. Оно за последние 100 лет обросло горой учебников и программных продуктов.

Однако при внимательном взгляде совершенно ясна нереалистичность классических предпосылок. Независимость результатов измерений обычно принимается "из общих предположений", между тем во многих случаях очевидна их коррелированность [33].

Одинаковая распределенность также вызывает сомнения из-за изменения во времени свойств измеряемых образцов, средств измерения и психофизического состояния специалистов, проводящих измерения (наблюдения, испытания, анализы, опыты). Даже обоснованность самой возможности применения вероятностных моделей также часто вызывает сомнения, например, при моделировании уникальных измерений (теорию вероятностей обычно привлекают при изучении массовых явлений). И уж совсем редко распределения результатов измерений можно считать нормальными (см. главу 4).

Итак, методы классической математической статистики обычно используют вне сферы их обоснованной применимости. Каково влияние отклонений от традиционных предпосылок на статистические выводы? В настоящее время об этом имеются лишь отрывочные сведения. Приведем три примера.

Пример 1. Построение доверительного интервала для математического ожидания обычно проводят с использованием распределения Стьюдента (при справедливости гипотезы нормальности). Как следует из Центральной Предельной Теоремы (ЦПТ) теории вероятностей, в асимптотике (при большом объеме выборки) такие расчетные методы дают правильные результаты. А именно, из ЦПТ вытекает использование квантилей нормального распределения, а из классической теории - квантилей распределения Стьюдента, но при росте объема выборки квантили распределения Стьюдента стремятся к соответствующим квантилям нормального распределения.

Пример 2. Для проверки однородности двух независимых выборок (на самом деле - для проверки равенства математических ожиданий) обычно рекомендуют использовать двухвыборочный критерий Стьюдента. Что будет при отклонении от нормальности распределений, из которых взяты выборки? Если объемы выборок равны или если дисперсии результатов наблюдений в выборках совпадают, то в асимптотике (когда объемы выборок безгранично возрастают) классический метод является корректным. Если же объемы выборок существенно отличаются и их дисперсии различны, то двухвыборочную статистику Стьюдента применять нельзя. Поскольку проверка равенства дисперсий - более сложная задача, чем проверка равенства математических ожиданий, то для выборок разного объема использовать двухвыборочную статистику Стьюдента не следует, лучше применять критерий Крамера- Уэлча, как это подробно обосновано в главе 4.

Пример 3. В задаче отбраковки (исключения) резко выделяющихся наблюдений (выбросов) расчетные методы, основанные на нормальности, являются крайне неустойчивыми по отношению к отклонениям от нормальности, что полностью лишает эти методы научной обоснованности (подробнее см. главу 4).

Примеры 1-3 показывают весь спектр возможных свойств классических расчетных методов в случае отклонения от нормальности. Методы примера 1 оказываются вполне пригодными при таких отклонениях, примера 2 - пригодными в некоторых случаях, примера 3 - полностью непригодными.

Итак, имеется **необходимость изучения свойств расчетных методов классической математической статистики, опирающихся на предположение нормальности, в ситуациях, когда это предположение не выполнено.** Аппаратом для такого изучения наряду с методом Монте-Карло (статистических испытаний) могут послужить предельные теоремы теории вероятностей (и опирающиеся на них асимптотические методы математической статистики), прежде всего ЦПТ, поскольку интересующие нас расчетные методы обычно используют разнообразные суммы.

Пока подобное изучение не проведено, остается неясной научная ценность, например, применения факторного анализа к векторам из переменных, принимающих небольшое число градаций и к тому же измеренных в порядковой шкале. Этот пример показывает важность еще одного направления исследований - изучения свойств алгоритмов, предназначенных для анализа числовых данных, в случаях, когда данные измерены в шкалах, отличных от абсолютной, в частности, в порядковой шкале. Подробнее это направление рассмотрено в главе 3.

Из большого числа возможных постановок, относящихся к изучению влияния отклонений от традиционных предпосылок, укажем лишь на то, что реальные данные имеют небольшое число значащих цифр (обычно от 2 до 5), в то время как в классической математической статистике используются непрерывные случайные величины, для которых вероятность получения подобного результата наблюдения равна 0. Действительно, вероятность того, что хотя бы один элемент выборки из распределения с непрерывной функцией распределение попадет в заданное счетное множество, в частности, в множество рациональных чисел, равна 0 (согласно классическим свойствам вероятностной меры). Событиями, имеющими вероятность 0, принято пренебрегать. Следовательно, с точки зрения классической математической статистики любыми реальными данными нужно пренебречь! Выходов из этого парадокса несколько. Один из них - бурно развивающаяся в настоящее время статистика интервальных данных (см. главу 9), другой - использование классических поправок Шеппарда для сгруппированных данных [34,35]. Здесь еще много работы. Так, даже для такого широко используемого статистического показателя, как коэффициент корреляции, поправки на группировку (поправки Шеппарда) были получены сравнительно недавно - лишь в 1980 г. [35].

Почему на первый план выдвинуто изучение классических алгоритмов, а не построение новых, специально предназначенных для работы в условиях отклонения от классических предпосылок? Во-первых, потому, что классические алгоритмы в настоящее время наиболее распространены (благодаря сложившейся системе образования как прикладников, так и математиков). Во-вторых, более новые подходы зачастую методологически уязвимы. Так, известная робастная модель засорения Тьюки-Хубера (см. главу 10) нацелена на борьбу с большими выбросами, которые зачастую физически невозможны из-за ограниченности интервала возможных значений измеряемой характеристики, в котором работает конкретное средство измерения. Следовательно, модель Тьюки-Хубера имеет скорее теоретическое значение, чем практическое. Сказанное, конечно, не означает, что следует прекратить разработку, изучение и внедрение непараметрических и устойчивых методов, выделенных выше как "точки роста" современных эконометрики и прикладной статистики.

Использование асимптотических результатов при конечных объемах выборок. Как отмечено выше, изучение классических алгоритмов во многих случаях может быть проведено с помощью асимптотических методов математической статистики, в частности, с помощью ЦПТ и методов наследования сходимости [14, п.2.4]. Отрыв классической математической статистики от нужд прикладных исследований проявился, в частности, в том, что в распространенных монографиях недостает математического аппарата, необходимого, в частности, для изучения двухвыборочных статистик. Суть в том, что переходить к пределу приходится не по одному параметру, а по двум – объемам двух выборок. Пришлось разработать соответствующую теорию – теорию наследования сходимости, изложенную в монографии [14, п.2.4].

Однако применять результаты подобного изучения придется при конечных объемах выборок. Возникает целый букет проблем, связанных с таким переходом. Часть из них обсуждалась в статье [37] в связи с изучением свойств статистик, построенных по выборкам из конкретных распределений.

Однако при обсуждении влияния отклонений от исходных предположений на свойства статистических процедур возникают дополнительные проблемы. Какие отклонения считать типичными? Ориентироваться ли на наиболее "вредные" отклонения, в наибольшей степени искажающие свойства алгоритмов, или же сосредоточить внимание на "типичных" отклонениях?

При первом подходе получаем гарантированный результат, но "цена" этого результата может быть излишне высокой. В качестве примера укажем на универсальное неравенство Берри-Эссеена для погрешности в ЦПТ [38,39]. Совершенно справедливо подчеркивает академик РАН А.А. Боровков [39, с,172], что "скорость сходимости в реальных задачах, как правило, оказывается лучше."

При втором подходе возникает вопрос, какие отклонения считать "типичными". Попытаться ответить на этот вопрос можно, анализируя большие массивы реальных данных. Вполне естественно, что ответы различных исследовательских групп будут различаться.

Одна из ложных идей - использование при анализе возможных отклонений только какого-либо конкретного параметрического семейства – распределений Вейбулла-Гнеденко, трехпараметрического семейства гамма - распределений и др. Как уже отмечалось выше, еще в 1927 г. акад. АН СССР С.Н. Бернштейн обсуждал методологическую ошибку, состоящую в сведении всех эмпирических распределений к четырехпараметрическому семейству Пирсона [5]. Однако и до сих пор параметрические методы статистики весьма популярны, особенно среди прикладников, и вина за это заблуждение лежит прежде всего на преподавателях статистических методов.

Выбор одного из многих критериев для проверки конкретной гипотезы. Во многих случаях для решения конкретной практической задачи разработано много методов, и специалист по математическим методам исследования стоит перед проблемой: какой из них предложить прикладнику для анализа конкретных данных?

В качестве примера рассмотрим задачу проверки однородности двух независимых выборок. Как известно [13], для ее решения можно предложить массу критериев: Стьюдента, Крамера-Уэлча, Лорда, хи - квадрат, Вилкоксона (Манна-Уитни), Ван – дер - Вардена, Сэвиджа, Н.В.Смирнова, типа омега-квадрат (Лемана-Розенблатта), Г.В. Мартынова и др. Какой выбрать?

Естественным образом приходит в голову идея "голосования": провести проверку по многим критериям, а затем принять решение "по большинству голосов". С точки зрения статистической теории такая процедура приводит попросту к построению еще одного критерия, который априори ничем не лучше прежних (но и не хуже), но более труден для изучения. С другой стороны, если совпадают решения по всем рассмотренным статистическим критериям, исходящим из различных принципов, то в соответствии с концепцией устойчивости, развитой в монографии [14], это повышает доверие к полученному общему решению.

Распространено, особенно среди математиков, ложное и вредное мнение о необходимости поиска оптимальных методов, решений и т.д. Дело в том, что оптимальность обычно исчезает при отклонении от исходных предпосылок. Так, среднее арифметическое в качестве оценки математического ожидания является оптимальной оценкой только тогда, когда исходное распределение - нормальное (см., например, монографию [40]), в то время как состоятельной оценкой - всегда, лишь бы математическое ожидание существовало. С другой стороны, для любого произвольно взятого метода оценивания или проверки гипотез обычно можно так сформулировать понятие оптимальности, чтобы рассматриваемый метод стал оптимальным – с этой специально выбранной точки зрения. Возьмем, например, выборочную медиану как оценку математического ожидания. Она, разумеется, оптимальна, хотя и в другом смысле, чем среднее арифметическое (оптимальное для нормального распределения). А именно, для распределения Лапласа выборочная медиана является оценкой максимального правдоподобия, а потому оптимальной - в том смысле, в каком оптимальной является любая оценка максимального правдоподобия. Соответствующее понятие оптимальности требует аккуратных формулировок, оно строго изложено в монографии [41]. Как известно, оценки максимального правдоподобия удобны при теоретических рассмотрениях, а при анализе конкретных экономических, технических и иных данных следует применять одношаговые оценки (см. об этом статью [42]).

Критерии однородности были проанализированы в монографии проф. Я.Ю. Никитина [43]. Естественных подходов к сравнению критериев несколько - на основе асимптотической относительной эффективности по Бахадуру, Ходжесу - Леману, Питмену. И выяснилось, что каждый критерий является оптимальным при соответствующей альтернативе или подходящем распределении на множестве альтернатив. При этом математические выкладки обычно используют альтернативу сдвига, сравнительно редко встречающуюся в практике анализа реальных статистических данных (в связи с критерием Вилкоксона эта альтернатива обсуждалась в главе 4). Итог печален - блестящая математическая техника, продемонстрированная в монографии [43], не позволяет дать рекомендации для выбора критерия проверки однородности при анализе реальных данных. Другими словами, с точки зрения работы прикладника, т.е. анализа конкретных данных, монография [43] бесполезна. Блестящее владение математикой и огромное трудолюбие, продемонстрированные автором этой монографии, увы, ничего не принесли практике.

Конечно, каждый практически работающий статистик так или иначе решает для себя проблему выбора статистического критерия. На основе ряда методологических соображений в главе 4 мы остановили свой выбор на состоятельном против любой альтернативы критерии типа омега-квадрат (Лемана-Розенблатта). Однако остается чувство неудовлетворенности в связи с недостаточной теоретической обоснованностью этого выбора.

Организация теоретических работ в области эконометрики и прикладной статистики. Выше продемонстрирована необходимость большой теоретической работы по развитию нацеленных на практическое использование математических методов исследования. В статье [6] 1992 г. обоснован вывод о необходимости создания сети научно-исследовательских организаций, которая выполняла бы такую работу. Как известно, количество научных работников к настоящему времени сократилось по крайней мере в 3 раза по сравнению с началом 1990-х годов, так что на осуществление в ближайшие годы сформулированной в [6] научно-организационной программы надеяться не приходится.

Приходится с сожалением констатировать, что в рамках научной специальности "теория вероятностей и математическая статистика" наблюдается четко выраженное игнорирование проблем статистического анализа реальных данных и уход в глубь узкоматематических исследований, которые ничего не могут дать практике. Причины этого явления, типичного для математических дисциплин, обсуждались выше. Поэтому нет оснований ожидать, что при "естественном ходе событий" будут получены существенные продвижения в рассмотренных выше нерешенных проблемах эконометрики и прикладной математической статистики.

Помочь может выделение государственными структурами системы грантов, направленных на поддержку работ в области нерешенных эконометрики и прикладной математической статистики. Принципиальным шагом явилось бы выделение эконометрики и прикладной математической статистики как самостоятельных научных направлений, отличных как от чисто математических дисциплин типа "теории вероятностей и математической статистики", так и от, например, ветви экономической теории, известной в официальных кругах под названием "статистика".

О прикладных работах с использованием методов прикладной статистики. Проблемы организации теоретических работ в области эконометрики и прикладной математической статистики лишь в перспективе важны для практической работы. Как правило, те, кто обрабатывает реальные данные, недостаточно знакомы с теоретическими основами алгоритмов и тем более не следят за событиями "на переднем крае" обсуждаемой научно-методической дисциплины. Это вполне естественно, поскольку основная специальность у таких специалистов - иная.

Несколько огрубляя, можно сказать, что реально используется только то, что имеется в учебниках и справочниках, в широко распространенных программных продуктах, а научные публикации с точки зрения прикладника представляют собой "информационный шум". Ситуация усугубляется традиционным ненормальным положением в отечественной статистике [7], наличием ошибок во многих изданиях.

К сожалению, учебная и научная литература на русском языке (как, впрочем, и на иных языках) по эконометрике и прикладной статистике в целом далека от совершенства, переполнена устаревшими методологическими подходами и прямыми ошибками. До сих пор наилучшим изданием остаются "Таблицы математической статистики" Л.Н. Большева и Н.В.Смирнова [13], созданные в 60-х годах.

Хотя студенты почти всех специальностей изучают в конце курса высшей математики раздел "теория вероятностей и математическая статистика", реально они знакомятся лишь с некоторыми основными понятиями и результатами, которых явно не достаточно для практической работы. С некоторыми математическими методами исследования студенты встречаются в специальных курсах (например, таких, как "Прогнозирование и технико-экономическое планирование", "Технико-экономический анализ", "Контроль качества продукции", "Маркетинг", "Контроллинг", "Математические методы прогнозирования" и др.), однако изложение в большинстве случаев носит весьма сокращенный и рецептурный характер. В результате подавляющую часть специалистов по эконометрике, прикладной математической статистике и их применению следует считать самоучками.

Поэтому большое значение имеет введение в технических вузах курса "Прикладная математическая статистика", а на экономических факультетах таких вузов – курса «Эконометрика», поскольку эконометрика – это, как известно, статистический анализ конкретных экономических данных (см. главу 1). Это естественно делать, например, в рамках подпрограммы "Технологии подготовки кадров для национальной технологической базы" федеральной целевой программы "Национальная технологическая база". Естественно, что курсы "Прикладная математическая статистика" и «Эконометрика» должны быть обеспечены соответствующими учебниками и учебными пособиями, методическими материалами и обучающими компьютерными системами.

Только через систему образования можно поднять уровень массового применения эконометрики и прикладной статистики и сократить отставание от "переднего края" теории. А это отставание в настоящее время составляет не менее 20 (но и не более 100) лет.

15.4. Высокие статистические технологии и эконометрика

В настоящем пункте подробно обсуждается ранее введенное понятие "высокие статистические технологии". Рассматриваются причины широкого распространения устаревших и частично ошибочных "низких" статистических технологий. Показано, что из всех путей повышения качества прикладных статистических исследований наиболее эффективным является расширение обучения "высоким статистическим технологиям", в том числе под именем эконометрики. Описан опыт работы Института высоких статистических технологий и эконометрики МГТУ им. Н.Э. Баумана.

Термин "высокие технологии" популярен в современной научно-технической литературе. Он используется для обозначения наиболее передовых технологий, опирающихся на последние достижения научно-технического прогресса. Есть такие технологии и среди технологий статистического анализа данных - как в любой интенсивно развивающейся научно-практической области.

Примеры высоких статистических технологий и входящих в них алгоритмов анализа данных, подробный анализ современного состояния и перспектив развития даны выше при обсуждении “точек роста” эконометрики как научно-практической дисциплины. В качестве "высоких статистических технологий" были выделены технологии непараметрического анализа данных; устойчивые (робастные) технологии; технологии, основанные на размножении выборок, на использовании достижений статистики нечисловых данных и статистики интервальных данных.

Термин "высокие статистические технологии". Обсудим пока не вполне привычный термин "высокие статистические технологии". Каждое из трех слов несет свою смысловую нагрузку.

"Высокие", как и в других областях, означает, что статистическая технология опирается на современные достижения статистической теории и практики, в частности, теории вероятностей и прикладной математической статистики. При этом "опирается на современные научные достижения" означает, во-первых, что математическая основа технологии получена сравнительно недавно в рамках соответствующей научной дисциплины, во-вторых, что алгоритмы расчетов разработаны и обоснованы в соответствии в нею (а не являются т.н. "эвристическими"). Со временем, если новые подходы и результаты не заставляют пересмотреть оценку применимости и возможностей технологии, заменить ее на более современную, "высокие статистические технологии" переходят в "классические статистические технологии", такие, как метод наименьших квадратов. Итак, высокие статистические технологии - плоды недавних серьезных научных исследований. Здесь два ключевых понятия - "молодость" технологии (во всяком случае, не старше 50 лет, а лучше - не старше 10 или 30 лет) и опора на "высокую науку".

Термин "статистические" привычен, но разъяснить его нелегко. Во всяком случае, к деятельности Государственного комитета РФ по статистике высокие статистические технологии отношения не имеют. Как известно, сотрудники проф. В.В. Налимова собрали более 200 определений термина "статистика" [44]. Полемика вокруг терминологии иногда принимает весьма острые формы (см., например, редакционные замечания к статье [1], написанные в стиле известных высказываний о генетике и кибернетике конца 1940-х годов). Современное представление о терминологии в области теории вероятностей и математической статистики отражено в Приложении 1 к настоящей книге, подготовленном в противовес распространенным ошибкам и неточностям в этой области. В частности, с точки зрения эконометрики статистические данные – это результаты измерений, наблюдений, испытаний, анализов, опытов, а "статистические технологии" - это технологии анализа статистических данных.

Наконец, редко используемый применительно к статистике термин "технологии". Статистический анализ данных, как правило, включает в себя целый ряд процедур и алгоритмов, выполняемых последовательно, параллельно или по более сложной схеме. В частности, можно выделить следующие этапы:

- планирование статистического исследования;
- организация сбора необходимых статистических данных по оптимальной или хотя бы рациональной программе (планирование выборки, создание организационной структуры и подбор команды эконометриков или статистиков, подготовка кадров, которые будут заниматься сбором данных, а также контролеров данных и т.п.);
- непосредственный сбор данных и их фиксация на тех или иных носителях (с контролем качества сбора и отбраковкой ошибочных данных по соображениям предметной области);
- первичное описание данных (расчет различных выборочных характеристик, функций распределения, непараметрических оценок плотности, построение гистограмм, корреляционных полей, различных таблиц и диаграмм и т.д.),
- оценивание тех или иных числовых или нечисловых характеристик и параметров распределений (например, непараметрическое интервальное оценивание коэффициента вариации или восстановление зависимости между откликом и факторами, т.е. оценивание функции),
- проверка статистических гипотез (иногда их цепочек - после проверки предыдущей гипотезы принимается решение о проверке той или иной последующей гипотезы),
- более углубленное изучение, т.е. применение различных алгоритмов многомерного статистического анализа, алгоритмов диагностики и построения классификации, статистики нечисловых и интервальных данных, анализа временных рядов и др.;
- проверка устойчивости полученных оценок и выводов относительно допустимых отклонений исходных данных и предпосылок используемых вероятностно-статистических моделей, допустимых преобразований шкал измерения, в частности, изучение свойств оценок методом размножения выборок;
- применение полученных статистических результатов в прикладных целях (например, для диагностики конкретных материалов, построения прогнозов, выбора инвестиционного проекта из предложенных вариантов, нахождения оптимальных режима осуществления технологического процесса, подведения итогов испытаний образцов технических устройств и др.),
- составление итоговых отчетов, в частности, предназначенных для тех, кто не является специалистами в эконометрических и статистических методах анализа данных, в том числе для руководства - "лиц, принимающих решения".

Возможны и иные структуризации статистических технологий. Важно подчеркнуть, что квалифицированное и результативное применение статистических методов - это отнюдь не проверка одной отдельно взятой статистической гипотезы или оценка параметров одного заданного распределения из фиксированного семейства. Подобного рода операции - только отдельные кирпичики, из которых складывается здание статистической технологии. Между тем учебники и монографии по статистике обычно рассказывают об отдельных кирпичиках, но не обсуждают проблемы их организации в технологию, предназначенную для прикладного использования.

Итак, процедура эконометрического или статистического анализа данных – это информационный технологический процесс, другими словами, та или иная информационная технология. Статистическая информация подвергается разнообразным операциям (последовательно, параллельно или по более сложным схемам). В настоящее время об автоматизации всего процесса статистического анализа данных говорить было бы несерьезно, поскольку имеется слишком много нерешенных проблем, вызывающих дискуссии среди статистиков. "Экспертные системы" в области статистического анализа данных пока не стали рабочим инструментом статистиков. Ясно, что и не могли стать. Можно сказать и жестче - это пока научная фантастика или даже вредная утопия.

В литературе статистические технологии рассматриваются явно недостаточно. В частности, обычно все внимание сосредотачивается на том или ином элементе технологической цепочки, а переход от одного элемента к другому остается в тени. Между тем проблема "стыковки" статистических алгоритмов, как известно, требует специального рассмотрения, поскольку в результате использования предыдущего алгоритма зачастую нарушаются условия применимости последующего. В частности, результаты наблюдений могут перестать быть независимыми, может измениться их распределение и т.п. (см. обсуждение этой проблемы в статье [45]).

Например, при проверке статистических гипотез большое значение имеют такие хорошо известные характеристики статистических критериев, как уровень значимости и мощность. Методы их расчета и использования при проверке одной гипотезы обычно хорошо известны. Если же сначала проверяется одна гипотеза, а потом с учетом результатов ее проверки - вторая, то итоговая процедура, которую также можно рассматривать как проверку некоторой (более сложной) статистической гипотезы, имеет характеристики (уровень значимости и мощность), которые, как правило, нельзя просто выразить через характеристики двух составляющих гипотез, а потому они обычно неизвестны. В результате итоговую процедуру нельзя рассматривать как научно обоснованную, она относится к эвристическим алгоритмам. Конечно, после соответствующего изучения, например, методом Монте-Карло, она может войти в число научно обоснованных процедур прикладной статистики.

Почему живучи "низкие статистические технологии"? "Высоким статистическим технологиям" противостоят, естественно, "низкие статистические технологии". Это те технологии, которые не соответствуют современному уровню науки и техники. Обычно они одновременно и устарели, и не адекватны сути решаемых эконометрических и статистических задач.

Примеры таких технологий неоднократно критически рассматривались на страницах различных изданий. В главе 4 рассматривались примеры неправильного использования критерия Вилкоксона для проверки совпадения теоретических медиан или функций распределения двух выборок. Можно также вспомнить критику использования классических процентных точек критериев Колмогорова и омега-квадрат в ситуациях, когда параметры оцениваются по выборке и эти оценки подставляются в "теоретическую" функцию распределения [46]. Приходилось констатировать широкое распространение таких порочных технологий и конкретных алгоритмов, в том числе в государственных и международных стандартах (перечень ошибочных стандартов дан в статье [47]), учебниках и распространенных пособиях. Тиражирование ошибок происходит обычно в процессе обучения в вузах или путем самообразования при использовании недоброкачественной литературы.

На первый взгляд вызывает удивление устойчивость "низких статистических технологий", их постоянное возрождение во все новых статьях, монографиях, учебниках. Поэтому, как ни странно, наиболее "долгоживущими" оказываются не работы, посвященные новым научным результатам, а публикации, разоблачающие ошибки, типа статьи [46]. Прошло больше 15 лет с момента ее публикации, но она по-прежнему актуальна, поскольку ошибочное применение критериев Колмогорова и омега-квадрат по-прежнему распространено.

Целесообразно рассмотреть здесь по крайней мере три обстоятельства, которые определяют эту устойчивость ошибок.

Во-первых, прочно закрепившаяся традиция. Учебники по т.н. «Общей теории статистики», написанные экономистами (поскольку учебная дисциплина "статистика" официально относится к экономике), если беспристрастно проанализировать их содержание, состоят в основном из введения в прикладную статистику, изложенного в стиле «низких статистических технологий», на уровне 1950-х годов. К "низкой" прикладной статистике добавлена некоторая информация о деятельности органов Госкомстата РФ. Примерно таково же положение со статистическими методами в медицине - одни и те же "низкие статистические технологии" переписываются из книги в книгу. Кратко говоря, «профессора-невежды порождают новых невежд» [7]. Так мы писали в 1990 г., но никто из указанных невежд даже не поинтересовался, какие ошибки имеются в виду. Новое поколение, обучившись ошибочным алгоритмам, их использует, а с течением времени и достижением должностей, ученых званий и степеней– пишет новые учебники со старыми ошибками.

Руководство Госкомстата РФ, воспользовавшись катаклизмами начала 1990-х годов, сделало вид, что ему неизвестно о создании в 1990 г. Всесоюзной статистической ассоциации и секции статистических методов в ее составе. Госкомстат РФ по-прежнему закрыт от "высоких статистических технологий" и работает на уровне позапрошлого века. Защита стала надежнее, поскольку в соответствии с современным стилем аппаратной работы на письма и обращения можно не отвечать.

Второе обстоятельство связано с большими трудностями при оценке экономической эффективности применения статистических методов вообще и при оценке вреда от применения ошибочных методов в частности. (А без такой оценки как докажешь, что "высокие статистические технологии" лучше "низких"?) Некоторые соображения по первому из этих вопросов приведены в статье [1], содержащей оценки экономической эффективности ряда работ по применению статистических методов. При оценке вреда от применения ошибочных методов приходится учитывать, что общий успех в конкретной инженерной или научной работе вполне мог быть достигнут вопреки их применению, за счет "запаса прочности" других составляющих общей работы. Например, преимущество одного технологического приема над другим можно продемонстрировать как с помощью критерия Крамера-Уэлча проверки равенства математических ожиданий (что правильно), так и с помощью двухвыборочного критерия Стьюдента (что, вообще говоря, неверно, т.к.

обычно не выполняются условия применимости этого критерия - нет ни нормальности распределения, ни равенства дисперсий). Кроме того, приходится выдерживать натиск невежд, защищающих свои ошибочные работы, например, государственные стандарты. Вместо исправления ошибок применяются самые разные приемы бюрократической борьбы с теми, кто разоблачает ошибки (подробнее см. статью [47]).

Третье существенное обстоятельство – трудности со знакомством с высокими статистическими технологиями. В течение последних 10 лет только журнал "Заводская лаборатория" систематически предоставлял такие возможности. К сожалению, поток современных отечественных и переводных статистических книг, выпускавшихся ранее, в частности, издательством “Финансы и статистика”, практически превратился в узкий ручеек… Возможно, более существенным является влияние естественной задержки во времени между созданием "новых статистических технологий" и написанием полноценной и объемной учебной и методической литературы. Она должна позволять знакомиться с новой методологией, новыми методами, теоремами, алгоритмами, технологиями не по кратким оригинальным статьям, а при обычном обучении в высшей школе.

Как ускорить внедрение "высоких статистических технологий"? Таким образом, весь арсенал используемых эконометрических и статистических методов можно распределить по трем потокам:

- высокие статистические технологии;
- классические статистические технологии,
- низкие статистические технологии.

Основная современная проблема статистических технологий состоит в обеспечении того, чтобы в конкретных эконометрических и статистических исследованиях использовались только технологии первых двух типов. При этом под классическими статистическими технологиями понимаем технологии почтенного возраста, сохранившие свое значение для современной статистической практики. Таковы метод наименьших квадратов, статистики Колмогорова, Смирнова, омега-квадрат, непараметрические коэффициенты корреляции Спирмена и Кендалла и многие другие статистические процедуры.

Каковы возможные пути решения основной современной проблемы в области статистических технологий?

Бороться с конкретными невеждами - дело почти безнадежное. Отстаивая свое положение и должности, они либо нагло игнорируют информацию о своих ошибках, как это делают авторы ряда учебников по "Общей теории статистики", либо с помощью различных бюрократических приемов уходят и от ответственности, и от исправления ошибок по существу (как это было со стандартами по статистическим методам - см. статью [6]). Третий вариант - признание и исправление ошибок - встречается, увы, редко. Но встречается.

Конечно, необходима демонстрация квалифицированного применения высоких статистических технологий. В 1960-70-х годах этим занималась лаборатория акад. А.Н. Колмогорова в МГУ им. М.В. Ломоносова. Секция "Математические методы исследования" журнала 2Заводская лаборатория" опубликовала за последние 40 лет более 1000 статей в стиле "высоких статистических технологий". В настоящее время действует Институт высоких статистических технологий и эконометрики МГТУ им. Н.Э.Баумана. Есть, конечно, целый ряд других научных коллективов, работающих на уровне "высоких статистических технологий".

Но самое основное - обучение. Какие бы новые научные результаты ни были получены, если они остаются неизвестными студентам, то новое поколение исследователей и инженеров вынуждено осваивать их по одиночке, а то и переоткрывать. Т.е. практически новые научные результаты почти исчезают, едва появившись. Как уже от меячалось, избыток публикаций превратился в тормоз развития. По нашим данным, к настоящему времени по статистическим технологиям опубликовано не менее миллиона статей и книг, из них не менее 100 тысяч являются актуальными для современного специалиста. Реальное число публикаций, которые способен освоить исследователь, по нашей оценке, не превышает 2-3 тысяч. Во всяком случае, в наиболее "толстом" (на русском языке) трехтомнике по статистике М. Дж. Кендалла и А. Стьюарта [8-10] приведено около 2 тысяч литературных ссылок. Итак, каждый исследователь знаком не более чем с 2-3% актуальных литературных источников. Поскольку существенная часть публикаций заражена "низкими статистическими технологиями", то исследователь самоучка имеет мало шансов выйти на уровень "высоких статистических технологий". Одновременно приходится констатировать, что масса полезных результатов погребена в изданиях прошлых десятилетий и имеет мало шансов встать в ряды "высоких статистических технологий" без специально организованных усилий современных специалистов.

Итак, еще и еще раз: основное - обучение. Несколько огрубляя, можно сказать: что то, что попало в учебные курсы и соответствующие учебные пособия - то сохраняется, что не попало - то пропадает. Подробнее об обучении - несколько позже. Сейчас - об упомянутом выше Институте высоких статистических технологий и эконометрики МГТУ им. Н.Э.Баумана.

Институт высоких статистических технологий и эконометрики. Организованный нами в 1989 г. Институт высоких статистических технологий и эконометрики (ИВСТЭ) действует на базе кафедры ИБМ-2 “Экономика и организация производства” Московского государственного технического университета им. Н.Э.Баумана. Институт на хоздоговорных и госбюджетных началах занимается развитием, изучением и внедрением эконометрики и "высоких статистических технологий", т.е. наиболее современных технологий анализа экономических, технических, социологических, медицинских данных, ориентированных на использование в условиях современного производства и экономики. Основной интерес представляют применения "высоких статистических технологий" для анализа конкретных экономических данных, т.е. в эконометрике. Наиболее перспективным представляется применение "высоких статистических технологий" для поддержки принятия управленческих решений, прежде всего в таком новом (для России) современном направлении экономической науки и практики, как контроллинг (см., например, монографию [48]).

Вначале Институт действовал как Всесоюзный центр статистических методов и информатики Центрального правления Всесоюзного экономического общества. В 1990-1992 гг. было выполнено более 100 хоздоговорных работ, в том числе для НИЦентра по безопасности атомной энергетики, ВНИИ нефтепереработки, ПО “Пластик”, ЦНИИ черной металлургии им. Бардина, НИИ стали, ВНИИ эластомерных материалов и изделий, НИИ прикладной химии, ЦНИИ химии и механики, НПО “Орион”, ВНИИ экономических проблем развития науки и техники, ПО “Уралмаш”, “АвтоВАЗ”, МИИТ, Казахского политехнического института, Донецкого государственного университета и многих других.

Затем Институт в качестве Лаборатории эконометрических исследований разрабатывал эконометрические методы анализа нечисловых данных, а также процедуры расчета и прогнозирования индекса инфляции и валового внутреннего продукта. Институт высоких статистических технологий и эконометрики развивал методологию построения и использования математических моделей процессов налогообложения (для Министерства налогов и сборов РФ), методологию оценки рисков реализации инновационных проектов высшей школы (для Министерства промышленности, науки и технологий РФ). Институт оценивал влияние различных факторов на формирование налогооблагаемой базы ряда налогов (для Минфина РФ), прорабатывал перспективы применения современных статистических и экспертных методов для анализа данных о научном потенциале (для Министерства промышленности, науки и технологий РФ). Важное направление связано с эколого-экономической тематикой - разработка методологического, программного и информационного обеспечения анализа рисков химико-технологических объектов (для Международного научно-технического центра), методов использования экспертных оценок в задачах экологического страхования (совместно с Институтом проблем рынка РАН). Институт проводил маркетинговые исследования (в частности, для Institute for Market Research GfK MR, Промрадтехбанка, фирм, торгующих растворимым кофе, программным обеспечением, оказывающих образовательные услуги). Интерес вызывали работы Института по прогнозированию социально-экономического развития России методом сценариев, по экономико-математическому моделированию развития малых предприятий и созданию современных систем информационной поддержки принятия решений для таких организаций.

Институт ведет фундаментальные исследования в области высоких статистических технологий и эконометрики, в частности, в рамках НИЧ МГТУ им. Н.Э. Баумана и Российского фонда фундаментальных исследований. Информация об Институте представлена на сайте в ИНТЕРНЕТе (http://antorlov.nm.ru, зеркала http://antorlov.euro.ru, http://www.newtech.ru/~orlov), который в 2000 г. посетили более 10000 пользователей. Институтом издается еженедельная компьютерная газета «Эконометрика» (около 1000 подписчиков). Архив выпусков газеты "Эконометрика" можно рассматривать как хрестоматию по различным разделам эконометрики, а также по высоким статистическим технологиям.

Зачем нужны высокие статистические технологии, разве недостаточно обычных статистических методов? Это вполне естественный вопрос. Мы считаем и доказываем своими теоретическими и прикладными работами, что совершенно недостаточно. Так, многие данные в информационных системах имеют нечисловой характер, например, являются словами или принимают значения из конечных множеств. Нечисловой характер имеют и упорядочения, которые дают эксперты или менеджеры, например, выбирая главную цель, следующую по важности и т.д. Значит, нужна статистика нечисловых данных. Она построена (см. главу 8). Далее, многие величины известны не абсолютно точно, а с некоторой погрешностью - от и до. Другими словами, исходные данные - не числа, а интервалы. Нужна статистика интервальных данных. Она развита (см. главу 9). В монографии [48] по контроллингу на с.138 хорошо сказано: "Нечеткая логика - мощный элегантный инструмент современной науки, который на Западе (и на Востоке - в Японии, Китае - А.О.) можно встретить в десятках изделий - от бытовых видеокамер до систем управления сооружениями, - у нас до самого последнего времени был практически неизвестен". Напомним, первая монография российского автора по теории нечеткости была выпущена в 1980 г. [49]. Ни статистики нечисловых данных, ни статистики интервальных данных, ни статистики нечетких данных нет и не могло быть в классической статистике. Все это - высокие статистические технологии. Они разработаны за последние 10-30-50 лет. А обычные вузовские курсы по общей теории статистики и по математической статистике разбирают научные результаты, полученные в первой половине XX века.

Важная часть эконометрики - применение высоких статистических технологий к анализу конкретных экономических данных, что зачастую требует дополнительной теоретической работы по доработке статистических технологий применительно к конкретной ситуации. Большое значение имеют конкретные эконометрические модели, например, модели экспертных оценок (глава 12) или экономики качества (глава 13). И конечно, такие конкретные применения, как расчет и прогнозирование индекса инфляции (глава 7).. Сейчас уже многим ясно, что годовой бухгалтерский баланс предприятия может быть использован для оценки его финансово-хозяйственной деятельности только с привлечением данных об инфляции.

Термин "эконометрика" пока мало известен в России. А между тем в мировой науке эконометрика занимает достойное место. Напомним, что Нобелевские премии по экономике получили эконометрики Ян Тильберген, Рагнар Фриш, Лоуренс Клейн, Трюгве Хаавельмо. В 2000 г. к ним добавились еще двое Джеймс Хекман и Дэниель Мак-Фадден. Выпускается ряд научных журналов, полностью посвященных эконометрике, в том числе: Journal of Econometrics (Швеция), Econometric Reviews (США), Econometrica (США), Sankhya (Indian Journal of Statistics. Ser.D. Quantitative Economics. Индия), Publications Econometriques (Франция).

Применение эконометрики дает заметный экономический эффект. Например, в США - не менее 20 миллиардов долларов ежегодно только в области статистического контроля качества. А что у нас? Повторим, что в секции "Математические методы исследования" журнала "Заводская лаборатория" за последние 40 лет напечатано более 1000 статей по высоким статистическим технологиям и их применениям. Однако в нашей стране по ряду причин эконометрика не была сформирована как самостоятельное направление научной и практической деятельности, в отличие, например, от Польши, не говоря уже об англосаксонских странах. В результате специалистов - эконометриков у нас на порядок меньше, чем в США и Великобритании (Американская статистическая ассоциация включает более 20000 членов).

Преподавание высоких статистических технологий и эконометрики. Приходится с сожалением констатировать, что в России практически отсутствует подготовка специалистов по высоким статистическим технологиям. В курсах по теории вероятностей и математической статистике обычно даются лишь классические основы этих дисциплин, разработанные в первой половине XX в., а преподаватели свою научную деятельность предпочитают посвящать доказательству теорем, имеющих лишь внутриматематический интерес, а не высоким статистическим технологиям.

В настоящее время появилась надежда на эконометрику. В России начинают развертываться эконометрические исследования и преподавание эконометрики, в том числе не только Институтом высоких статистических технологий и эконометрики. Преподавание этой дисциплины ведется в Московском государственном университете экономики, статистики и информатики (МЭСИ), на экономическом факультете МГУ им. М.В. Ломоносова, в Высшей школе экономики и еще в нескольких экономических учебных заведениях. Среди технических вузов факультет "Инженерный бизнес и менеджмент" МГТУ им. Н.Э.Баумана имеет в настоящее время приоритет в преподавания эконометрики.

Мы полагаем, что экономисты, менеджеры и инженеры, прежде всего специалисты по контроллингу [48], должны быть вооружены современными средствами информационной поддержки, в том числе высокими статистическими технологиями и эконометрикой.

Очевидно, преподавание должно идти впереди практического применения. Ведь как применять то, чего не знаешь?

Один раз - в 1990-1992 гг. мы уже обожглись на недооценке необходимости предварительной подготовки тех, для кого предназначены современные компьютерные средства. Наш коллектив (Всесоюзный центр статистических методов и информатики Центрального правления Всесоюзного экономического общества) разработал систему диалоговых программных систем обеспечения качества продукции. Их созданием руководили ведущие специалисты страны. Но распространение программных продуктов шло на 1-2 порядка медленнее, чем ожидалось. Причина стала ясна не сразу. Как оказалось, работники предприятий просто не понимали возможностей разработанных систем, не знали, какие задачи можно решать с их помощью, какой экономический эффект они дадут. А не понимали и не знали потому, что в вузах никто их не учил статистическим методам управления качеством. Без такого систематического обучения нельзя обойтись - сложные концепции "на пальцах" за пять минут не объяснишь.

Есть и противоположный пример - положительный. В середине 1980-х годов в советской средней школе ввели новый предмет "Информатика". И сейчас молодое поколение превосходно владеет компьютерами, мгновенно осваивая быстро появляющиеся новинки, и этим заметно отличается от тех, кому за 30-40 лет. Если бы удалось ввести в средней школе курс вероятности и статистики - а такой курс есть в Японии и США, Швейцарии, Кении и Ботсване, почти во всех странах (см. подготовленный ЮНЕСКО сборник докладов [50]) - то ситуация могла бы быть резко улучшена. Надо, конечно, добиться, чтобы такой курс был построен на высоких статистических технологиях, а не на низких. Другими словами, он должен отражать современные достижения, а не концепции пятидесятилетней или столетней давности.

Необходимо активизировать деятельность Российской ассоциации статистических методов. Но не стоит ограничиваться только внутренними проблемами сообщества специалистов по статистическим методам. Например, в созданном в России профессиональном экономическом обществе - Ассоциации контроллеров России - необходимо, на наш взгляд, выделить направление, посвященное применению высоких статистических технологий и эконометрики в контроллинге, а также учесть необходимость обучения основам этого направления при формировании мощной образовательной базы контроллинга.

Цитированная литература

1. Орлов А.И. Что дает прикладная статистика народному хозяйству? / Вестник статистики. 1986. № 8. С.52 - 56

2. Комаров Д;М., Орлов А.И. Роль методологических исследований в разработке методоориентированных экспертных систем (на примере оптимизационных и статистических методов) - В сб.: Вопросы применения экспертных систем. - Минск: Центросистем, 1988. С.151-160.

3. Ленин В.И. Развитие капитализма в России. Процесс образования внутреннего рынка для крупной промышленности. - М.: Политиздат, 1986. - XII+610 с.

4. Гнеденко Б.В. Курс теории вероятностей: Учебник. - Изд. 6-е, переработанное и дополненное. - М.: Наука, Гл. ред. физ. - мат. лит., 1988. - 448 с.

5. Бернштейн С.Н. Современное состояние теории вероятностей и ее приложений. - В сб.: Труды Всероссийского съезда математиков в Москве 27 апреля - 4 мая 1927 г. - М.-Л.: ГИЗ, 1928. С.50-63.
6. Орлов А.И. О современных проблемах внедрения прикладной статистики и других статистических методов. / Заводская лаборатория. 1992. Т.58. № 1. С.67-74.
7. Орлов А.И. О перестройке статистической науки и её применений. / Вестник статистики. 1990. № 1. С.65 - 71.
8. Кендалл М., Стьюарт А. Теория распределений. - М.: Наука, 1966. - 566 с.
9. Кендалл М., Стьюарт А. Статистические выводы и связи. - М.: Наука, 1973. - 899 с.
10. Кендалл М., Стьюарт А. Многомерный статистический анализ и временные ряды. - М.: Наука, 1976. - 736 с.
11. Налимов В.В., Мульченко З.М. Наукометрия. Изучение развития науки как информационного процесса. - М.: Наука, 1969. - 192 с.
12. ГОСТ 11.011-83. Прикладная статистика. Правила определения оценок и доверительных границ для параметров гамма-распределения. - М.: Изд-во стандартов. 1984. - 53 с.
13. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. - М.: Наука, 1965 (1-е изд.), 1968 (2-е изд.), 1983 (3-е изд.).
14. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука,1979. - 296 с.
15. Смоляк С.А., Титаренко Б.П. Устойчивые методы оценивания: Статистическая обработка неоднородных совокупностей. - М;: Статистика, 1980. - 208 с.
16. Эфрон Б. Нетрадиционные методы многомерного статистического анализа. - М.: Финансы и статистика, 1988. - 263 с.
17. Суппес П., Зинес Дж. Основы теории измерений. - В сб.: Психологические измерения. - М: Мир,1967. С. 9-110.
18. Пфанцагль И. Теория измерений. - М.: Мир, 1976. - 166 с.
19. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений. - М.: Мир, 1976. - 168 с.
20. Дэвид Г. Метод парных сравнений. - М.: Статистика, 1978. - 144 с.
21. Матерон Ж. Случайные множества и интегральная геометрия. - М.: Мир, 1978. - 318 с.
22. Терехина А.Ю. Анализ данных методами многомерного шкалирования. - М.: Наука, 1986. - 168 с.
23. Перекрест В.Т. Нелинейный типологический анализ социально-экономической информации: Математические и вычислительные методы. - Л.: Наука, 1983. - 176 с.
24. Кемени Дж., Снелл Дж. Кибернетическое моделирование: Некоторые приложения. - М.: Советское радио, 1972. - 192 с.
25. Тюрин Ю.Н., Литвак Б.Г., Орлов А.И., Сатаров Г.А., Шмерлинг Д.С. Анализ нечисловой информации. - М.: Научный Совет АН СССР по комплексной проблеме "Кибернетика", 1981. - 80 с.
26. Литвак Б.Г. Экспертная информация: Методы получения и анализа. - М.: Радио и связь, 1982. - 184 с.
27. Орлов А.И. Статистика объектов нечисловой природы и экспертные оценки. - В сб.: Экспертные оценки. Вопросы кибернетики. Вып.58. - М.: Научный Совет АН СССР по комплексной проблеме "Кибернетика", 1979. С.17-33.
28. Анализ нечисловой информации в социологических исследованиях. / Под ред. В.Г. Андреенкова, А.И.Орлова, Ю.Н. Толстовой. - М.: Наука, 1985. - 220 с.

29. Орлов А.И. Асимптотическое поведение статистик интегрального типа. / Доклады АН СССР. 1974. Т.219. № 4. С.808-811.
30. Орлов А.И. Асимптотическое поведение статистик интегрального типа. - В сб.: Вероятностные процессы и их приложения. Межвузовский сборник. - М.: МИЭМ, 1989. С.118-123.
31. Горский В.Г. Современные статистические методы обработки и планирования экспериментов в химической технологии. - В сб.: Инженерно-химическая наука для передовых технологий. Международная школа повышения квалификации Труды третьей сессии. 26-30 мая 1997, Казань, Россия / Под ред. В.А. Махлина. - М.: Научно-исследовательский физико-химический институт им. Карпова, 1997. С.261-293.
32. Плошко Б.Г., Елисеева И.И. История статистики: Учебное пособие. - М.: Финансы и статистика. 1990. - 295 с.
33. Эльясберг П.Е. Измерительная информация. Сколько ее нужно, как ее обрабатывать? - М.: Наука, 1983. - 208 с.
34. Крамер Г. Математические методы статистики. - М.: Мир, 1975. - 648 с.
35. Орлов А.И., Орловский И.В. О поправках на группировку. - В сб.: Прикладной многомерный статистический анализ. - М.: Наука, 1978. - С.339-342.
36. Орлов А.И. Поправка на группировку для коэффициента корреляции. / Экономика и математические методы. - 1980. - Т.XVI. - №4. - С.800-801.
37. Орлов А.И. Методы оценки близости допредельных и предельных распределений статистик. / Заводская лаборатория. - 1998. - Т.64. - № 5. - С.64-67.
38. Феллер В. Введение в теорию вероятностей и ее приложения. Т.2. - М.: Мир, 1984. - 751 с.
39. Боровков А.А. Теория вероятностей. - М.: Наука, 1976. - 352 с.
40. Каган А.М., Линник Ю.В., Рао С.Р. Характеризационные задачи математической статистики. - М.: Наука, 1972. - 656 с.
41. Ибрагимов И.А., Хасьминский Р.З. Асимптотическая теория оценивания. - М.: Наука, 1979. - 528 с.
42. Орлов А.И. О нецелесообразности использования итеративных процедур нахождения оценок максимального правдоподобия. / "Заводская лаборатория", 1986. Т.52. No.5. С.67-69.
43. Никитин Я.Ю. Асимптотическая эффективность непараметрических критериев. - М.: Наука, 1995. - 240 с.
44. Никитина Е.П., Фрейдлина В.Д., Ярхо А.В. Коллекция определений термина "статистика" / Межфакультетская лаборатория статистических методов. Вып.37. - М.: Изд-во Московского государственного университета им. М.В. Ломоносова, 1972. - 46 с.
45. Орлов А.И. Проблема множественных проверок статистических гипотез. / Заводская лаборатория. 1996. Т.62. No.5. С.51-54.
46. Орлов А.И. Распространенная ошибка при использовании критериев Колмогорова и омега-квадрат. / Заводская лаборатория. - 1985. - Т.51. - No.1. - С.60-62.
47. Орлов А.И. Сертификация и статистические методы. / Заводская лаборатория. 1997. Т.63. No.3. С.55-62.
48. Контроллинг в бизнесе. Методологические и практические основы построения контроллинга в организациях / А.М. Карминский, Н.И. Оленев, А.Г. Примак, С.Г.Фалько. - М.: Финансы и статистика, 1998. - 256 с.
49. Орлов А. И. Задачи оптимизации и нечеткие переменные. - М.: Знание, 1980.- 64 с.

50. The teaching of statistics / Studies in mathematics education. Vol.7. - Paris, UNESCO, 1989. - 258 pp.

Приложение 1

Вероятностно-статистические основы эконометрики

Эконометрика опирается на твердый научный фундамент - теорию вероятностей и статистику. В области теории вероятностей наша страна является признанным мировым лидером. Практически все специалисты в этой области исходят в своей работе из аксиоматики теории вероятностей, предложенной академиком А.Н. Колмогоровым в 1933 г. [1].

Однако в отечественной и зарубежной литературе присутствуют различные интерпретации терминов и разделов эконометрики, теории вероятностей, статистики. Одна из причин состоит в том, что используют в своей работе эти научные области специалисты разных профессий - экономисты, инженеры, математики… Поэтому мы приводим основную терминологию и краткое описание математической статистики и ее новых разделов.

П1-1. Определения терминов теории вероятностей и прикладной статистики

Определения практически всех используемых в литературе понятий теории вероятностей и математической статистики и основные сведения о соответствующих математических объектах собраны в Энциклопедии [2]. Ниже приведены определения и обозначения (в стиле [2]) лишь для основных понятий теории вероятностей и прикладной статистики, используемых в настоящем учебном пособии. Как показали предыдущие публикации (см., например, [3]), эта сводка позволяет осознанно изучать и применять эконометрические методы для анализа конкретных экономических данных. Однако она, очевидно, не заменяет систематических курсов теории вероятностей и прикладной математической статистики, знакомство с которыми - необходимая предпосылка для изучения эконометрики.

Споры по поводу терминов весьма распространены. Весьма популярно желание добиться единства терминологии. Однако практика терминологических дискуссий показывает, что придти к единому мнению обычно не удается. Не помогают достижению единства и административные меры, например, принятие государственных стандартов, "несоблюдение которых карается по закону". Зачастую такие стандарты содержат в себе много спорного, а то и ошибочного (подробнее об этом см. [3]).

Почти в каждой области знания параллельно существуют различные терминологические системы. Большого вреда это обычно не приносит. Так, операция умножения двух чисел a и b может быть обозначена четырьмя способами - крестиком (т.е. a x b), точкой ($a\cdot b$), отсутствием знака между сомножителями (ab) или звездочкой, как при программировании ($a*b$). Случайные величины обозначают либо латинскими буквами, либо греческими. Для математического ожидания используют либо символ M, либо символ E, и т.п.. Обычно можно без труда понять, о чем идет речь.

Однако при изучении настоящего курса эконометрики необходимо пользоваться вполне определенной терминологической системой. Она и приводится ниже. При этом мы отнюдь не отрицаем пригодности других систем терминов и определений в тех или иных случаях.

№№ пп.	Термины	Определения	Примечания
		1. Теория вероятностей	
		1.1. Общие понятия	
1.1.1.	Пространство элементарных событий	Множество, элементы которого, называемые элементарными событиями, соответствуют возможным результатам наблюдения, измерения, анализа, проверки, исходам опыта, эксперимента, испытания.	Пространство элементарных событий $\Omega = \{\omega\}$ лежит в основе вероятностных моделей явлений (процессов). Вместо явного описания пространства элементарных событий часто используют косвенное или частичное описание, например, с помощью распределений случайных величин.
1.1.2.	Случайное событие	Измеримое подмножество пространства элементарных событий.	Термин "измеримое" понимают в смысле теории измеримых множеств. Случайные события образуют σ-алгебру G.
1.1.3.	Вероятностная мера	Сигма-аддитивная мера P, определенная на всех случайных событиях и такая, что $P(\Omega) = 1$, где Ω - пространство элементарных событий	Вероятностная мера P - функция, ставящая в соответствие каждому случайному событию A его вероятность P(A). Термин "мера" понимают в смысле математической теории меры. Синонимы: вероятностное распределение, распределение вероятностей, распределение, вероятность на пространстве элементарных событий.
1.1.4.	Вероятностное	Совокупность {Ω, G, P} пространства элемен-	Вероятностное пространство (синоним: поле вероят-

	пространство	тарных событий Ω, класса случайных событий G и вероятностной меры P.	ностей) - основной исходный объект теории вероятностей и вероятностных моделей реальных явлений (процессов).
1.1.5.	Вероятность события A	Значение P(A) вероятностной меры P на случайном событии A.	В силу закона больших чисел частота реализации события A при неограниченном увеличении числа независимых повторений одного и того же комплекса условий, описываемого вероятностным пространством {Ω, G, P}, стремится к вероятности этого события P(A), т.е. для любого $\varepsilon > 0$ $\lim_{n\to\infty} P\{\mid m/n - p \mid \le \varepsilon\} = 1$, где m/n - частота, p - вероятность события A, n - число повторений. Это свойство нельзя принимать за определение вероятности события в математической теории вероятностей. Оно указывает способ оценивания вероятности по опытным данным.
1.1.6.	Независимость случайных событий	Случайные события A и B являются независимыми, если $P(AB) = P(A)P(B)$, где AB - пересечение множеств A и B (произведение событий A и B). Случайные события A_1, A_2,..., A_n называются независимыми (в совокупности), если $P(A_1A_2...A_n) = P(A_1)P(A_2)...P(A_n)$ и аналогичные равенства справедливы для всех	Общематематическое понятие пересечения множеств $A \cap B$ в теории вероятностей по традиции эквивалентно понятию произведения событий AB.

		поднаборов этих событий $A_{(1)}$, $A_{(2)}$,..., $A_{(k)}$, $2 \le k \le n-1$.	
1.1.7.	Случайный элемент	Измеримая функция, определенная на вероятностном пространстве.	Случайный элемент X принимает значения в измеримом пространстве (Z,J), где Z - пространство значений X, а J - класс измеримых подмножеств Z; при этом для любого Q∈J множество $X^{-1}(Q)$ является случайным событием. Если Z - множество действительных чисел R^1, то случайный элемент X называют случайной величиной. Если $Z = R^k$ - конечномерное векторное пространство размерности k=2,3,...., то случайный элемент X называют случайным вектором.
1.1.8.	Распределение случайного элемента	Функция множества, задающая вероятность принадлежности случайного элемента измеримому подмножеству его области значений.	Для случайного элемента X, определенного на вероятностном пространстве {Ω, G, P} со значениями в измеримом пространстве (Z,J), его распределение P_1:J $\to$ [0,1] задается формулой $P_1(Q) = P(X^{-1}(Q))$, Q∈J.
1.1.9.	Дискретный случайный элемент	Случайный элемент, область значений которого состоит из конечного или счетного множества точек.	Распределение случайного элемента X, принимающего только значения x_1, x_2,..., полностью описывается числами $p_i = P(X=x_i)$, i = 1,2,..., причем $p_1 + p_2 + ... = 1$.
1.1.10.	Параметрическое семейство	Функция, определенная на параметрическом пространстве (подмножестве конечномерного	Параметр может быть одномерным или конечномерным. Вместо "зависимость от k-мерного параметра"

	распределений	векторного пространства), которая каждому значению параметра (числу или вектору, входящему в параметрическое пространство) ставит в соответствие распределение случайного элемента.	часто говорят "зависимость от k параметров".
1.1.11.	Независимость случайных элементов	Определенные на одном и том же вероятностном пространстве случайные элементы X_1, $X_2,...,X_k$ со значениями в измеримых пространствах (Z_1, J_1), (Z_2, J_2),..., (Z_k, J_k) соответственно называются независимыми, если для любых $Q_1 \in J_1$, $Q_2 \in J_2$,..., $Q_k \in J_k$ имеем $P(X_1 \in Q_1, X_2 \in Q_2,..., X_k \in Q_k) = P(X_1 \in Q_1)P(X_2 \in Q_2)... P(X_k \in Q_k)$.	Для случайных величин и векторов, имеющих плотности вероятности, независимость эквивалентна тому, что плотность вероятности вектора $(X_1, X_2,..., X_k)$ равна произведению плотностей вероятностей случайных величин X_i, т.е. $f(x_1, x_2,..., x_k) = f(x_1)f(x_2)...f(x_k)$. Результаты экспериментов, которые проведены независимо друг от друга, как правило, моделируются с помощью независимых случайных величин.
1.1.12	Вероятностная модель явления (процесса)	Математическая модель явления (процесса), в которой использованы понятия теории вероятностей и математической статистики.	Установление (формулировка) исходной вероятностной модели - необходимый первый этап для применения методов прикладной статистики.

		1.2. Случайная величина	
1.2.1.	Случайная величина	Однозначная действительная измеримая функция на вероятностном пространстве.	Однозначная действительная функция $X:\Omega \to R^1$ является случайной величиной, если для любого $x \in R^1$ множество $\{\omega: X(\omega) \le x\}$ является случайным событием. Случайная величина - это случайный элемент со значениями в R^1. (Здесь R^1 - множество действительных чисел.)
1.2.2.	Функция распределения	Функция, определяющая для всех действительных чисел x вероятность того, что случайная величина X принимает значения, меньшие x.	Функция распределения $F(x) = P(X < x) = P\{\omega: X(\omega) < x\}$. Функция распределения непрерывна слева. *Примечание.* Иногда функцию распределения определяют как $F(x) = P(X < x) = P\{\omega: X(\omega) < x\}$. Тогда она непрерывна справа.
1.2.3.	Плотность вероятности	Функция p(t) такая, что $$F(x) = \int_{-\infty}^{x} p(t)\,dt$$ при всех x, где F(x) - функция распределения рассматриваемой случайной величины.	Сокращенная форма: плотность.
1.2.4.	Непрерывная случайная величина	Случайная величина, функция распределения которой при всех действительных x непрерывна.	

1.2.5.	Квантиль порядка p	Значение случайной величины, для которого функция распределения принимает значение p или имеет место "скачок" со значения меньше p до значения больше p.	Число x_p - квантиль порядка p для случайной величины с функцией распределения F(x) тогда и только тогда, когда $\lim_{x \to x_p+0} F(x) \geq p,\ F(x_p) \leq p.$ Может случиться, что вышеуказанное условие выполняется для всех значений x, принадлежащих некоторому интервалу. Тогда каждое такое значение называется квантилью порядка p. Примечание. Одни авторы употребляют термин "квантиль" в мужском роде, другие - в женском.
1.2.6.	Медиана	Квантиль порядка p = 1/2.	
1.2.7.	Мода непрерывной случайной величины	Значение случайной величины, соответствующее локальному максимуму ее плотности вероятности.	Мод у непрерывной случайной величины может быть несколько (конечное число или бесконечно много). Краткая форма термина: мода.
1.2.8.	Математическое ожидание	Среднее взвешенное по вероятностям значение случайной величины X(ω), т.е. $\int_{\Omega} X(\omega)\, P\,(d\omega)$	Математическое ожидание обозначают M(X), E(X), MX, EX и др. Рекомендуемое обозначение: M(X). При этом $M(X) = \int_{\Omega} X(\omega)\, P\,(d\omega) =$ $\int_{-\infty}^{+\infty} x\, dF(x) = \int_{-\infty}^{+\infty} t\, p(t)\, dt$

			где $F(x)$ - функция распределения, а $p(t)$ - плотность вероятности случайной величины $X = X(\omega)$. Математическое ожидание существует не для всех случайных величин X. Для существования математического ожидания необходимо и достаточно абсолютной сходимости соответствующего интеграла.
1.2.9.	Дисперсия (случайной величины X)	Математическое ожидание квадрата разности между случайной величиной и ее математическим ожиданием.	Для случайной величины X дисперсия $D(X) = \sigma^2=\sigma^2(X)=M(X-M(X))^2$. Дисперсия равна 0 тогда и только тогда когда $P(X=a)=1$ для некоторого a.
1.2.10.	Среднее квадратическое отклонение	Неотрицательный квадратный корень из дисперсии.	
1.2.11.	Коэффициент вариации	Отношение среднего квадратического отклонения к математическому ожиданию.	Применяется для положительных случайных величин как показатель разброса.
1.2.12.	Момент порядка q (случайной величины X)	Математическое ожидание случайной величины X^q.	
1.2.13.	Центральный момент порядка q (случайной величины X)	Математическое ожидание случайной величины $(X-M(X))^q$, где $M(X)$ - математическое ожидание X.	Дисперсия - центральный момент порядка 2.
1.2.14.	Характеристи-	Функция от $t \in R^1$, при каждом t равная мате-	$M(e^{itX}) = M(\cos(tX) + i\sin(tX)) = M(\cos(tX)) +$

	ческая функция (случайной величины X)	матическому ожиданию случайной величины e^{itX}, где i - мнимая единица, е - основание натуральных логарифмов.	iM(sin(tX)).
		1.3. Случайный вектор	
1.3.1.	Случайный вектор	Однозначная измеримая функция на вероятностном пространстве со значениями в конечномерном евклидовом пространстве R^k.	Случайный вектор X - это случайный элемент со значениями в R^k, т.е. $X = X(\omega) = (X_1(\omega), X_2(\omega),...., X_k(\omega))$, где $X_i(\omega)$, i = 1,2,...,k, - случайные величины, заданные на одном и том же вероятностном пространстве.
1.3.2.	Функция распределения (случайного вектора)	Функция распределения $F(x_1, x_2,...., x_k)$ случайного вектора $X(\omega) = (X_1(\omega), X_2(\omega),...., X_k(\omega))$ удовлетворяет равенству $F(x_1, x_2,...., x_k) =$ $P(X_1<x_1, X_2<x_2,..., X_k<x_k) = P\{\omega: X_1(\omega)< x_1, X_2(\omega)< x_2,..., X_k(\omega)< x_k)$.	
1.3.3.	Плотность вероятности (случайного вектора)	Функция p(x) такая, что $P(X \in A) = \int_A p(x)\, dx$ для случайного вектора X = X(ω) и любого борелевского подмножества A конечномерного евклидова пространства R^k.	

1.3.4.	Математическое ожидание случайного вектора	Вектор, компоненты которого - математические ожидания компонент случайного вектора.	Математическое ожидание случайного вектора X = $(X_1, X_2,...., X_k)$ есть $(M(X_1), M(X_2),...., M(X_k))$, где $M(X_i)$ - математическое ожидание случайной величины X_i, являющейся i - ой компонентой случайного вектора X, i = 1,2,...,k.
1.3.5.	Ковариация (для двумерного вектора)	Ковариацией вектора (X,Y) называется математическое ожидание случайной величины (X - MX))(Y - M(Y)), где M(X) и M(Y) - математические ожидания случайных величин X и Y.	cov(X,Y) = M (X - M(X))(Y - M(Y)) ; если X = Y, то cov(X,Y) = D(X) - дисперсия X.
1.3.6.	Ковариационная матрица случайного вектора	Квадратная матрица $\|c_{ij}\|$ порядка k, в которой c_{ij} - ковариация двумерного вектора (X_i, X_j), где X_i и X_j - компоненты случайного вектора $X = (X_1, X_2,...., X_k)$, i,j = 1,2,...,k.	Ковариационная матрица симметрична, на главной диагонали стоят дисперсии X_i - компонент X, i = 1,2,...,k.
1.3.7.	Коэффициент корреляции (для двумерного вектора)	Отношение ковариации вектора (X,Y) к произведению средних квадратических отклонений σ(X) и σ(У) случайных величин X и У.	$$r(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sigma(X)\sigma(Y)}$$ Если Y = aX+b, то $\|r(X,Y)\| = 1$. Верно и обратное: если $\|r(X,Y)\| = 1$, то Y = aX+b..
1.3.8.	Корреляционная матрица	Квадратная матрица $\|r_{ij}\|$ порядка k, в которой r_{ij} - коэффициент корреляции двумерного	Корреляционная матрица симметрична, на главной диагонали стоят единицы.

	случайного вектора	вектора (X_i, X_j), где X_i и X_j - компоненты случайного вектора $X = (X_1, X_2,...., X_k)$, $i,j = 1,2,...,k$.	
		2. Прикладная статистика	
		2.1. Общие понятия	
2.1.1.	Признак	Свойство (характеристика) объекта наблюдения.	Частными видами наблюдения являются измерение, испытание, анализ, опыт, проверка и т.д.
2.1.2.	Результат наблюдения	Значение признака объекта наблюдения.	Результат наблюдения может быть числом, вектором, элементом конечного множества или математическим объектом иной природы.
2.1.3.	Выборка	Совокупность значений одного и того же признака у подвергнутых наблюдению объектов.	Выборка - совокупность чисел или векторов, или математических объектов иной природы, соответствующих изучаемым реальным объектам наблюдения.
2.1.4.	Объем выборки	Число результатов наблюдений, включенных в выборку.	Объем выборки обычно обозначают n.
2.1.5.	Вероятностная модель выборки	Вероятностная модель получения результатов наблюдений, включаемых в выборку.	Примерами вероятностных моделей выборок являются простая случайная выборка и случайная выборка из конечной совокупности.
2.1.6.	Простая случайная выборка	Выборка, в которой результаты наблюдений моделируются как совокупность независимых одинаково распределенных случайных эле-	Если результаты наблюдений имеют распределение F, то говорят, что "выборка извлечена из распределения F".

		ментов.	
2.1.7.	Случайная выборка из конечной совокупности	Выборка объема n, в которую включены результаты наблюдений над объектами, отбираемыми из конечной совокупности так, что любой набор n объектов имеет одинаковую вероятность быть отобранным.	Если N - число объектов конечной совокупности, то для получения случайной выборки объема n из этой совокупности, $n < N$, отбор объектов для проведения наблюдений должен проводиться так, чтобы любой набор из n объектов имел одну и ту же вероятность быть отобранным, равную $n!(N-n)!/ N!$, т.е. обратной величине к числу сочетаний из N элементов по n.
2.1.8.	Статистика	Измеримая функция результатов наблюдений, включенных в выборку, используемая для получения статистических выводов.	Статистики используются для описания данных, оценивания, проверки гипотез. Статистика, как функция случайного элемента, является случайным элементом. Статистика принимает значения в некотором измеримом пространстве (Z,J), своем для каждой статистики.
		2.2. Описание данных	
2.2.1.	Частота события	Отношение числа наблюдений, в которых осуществилось событие, к объему выборки.	
2.2.2.	Эмпирическое распределение	Распределение случайного элемента, в котором каждому результату наблюдения, включенному в выборку, соответствует одна и та же вероятность, равная обратной величине объема выборки.	Если в выборку включены результаты наблюдений x_1, x_2,...., x_n, то эмпирическое распределение - это распределение случайной величины X такой, что $P(X= x_i) = 1/n$, $i = 1,2,..., n$. Если несколько результатов наблюдений совпадают: $x_1 = x_2 =.... = x_k = a$, то полагают $P(X=a) = k/n$.

2.2.3.	Эмпирическая функция распределения	Функция эмпирического распределения.	Определена, когда результаты наблюдений - числа или вектора (функции распределения по пп.1.2.2 и 1.3.2 соответственно).
2.2.4.	Выборочное среднее арифметическое	Сумма результатов наблюдений, включенных в выборку, деленная на ее объем.	Выборочное среднее арифметическое равно математическому ожиданию случайной величины, имеющей эмпирическое распределение.
2.2.5.	Выборочная дисперсия	Сумма квадратов отклонений результатов наблюдений, включенных в выборку, от их выборочного среднего арифметического, деленная на объем выборки.	Выборочная дисперсия $s^2 = 1/n \sum_{i=1}^{n} (x_i - x_{cp})^2$-, где $x_1, x_2, \ldots, x_n$ - результаты наблюдений, включенные в выборку; x_{cp} - выборочное среднее арифметическое, $x_{cp} = 1/n \sum_{i=1}^{n} x_i$. Выборочная дисперсия равна дисперсии случайной величины, имеющей эмпирическое распределение.
2.2.6.	Выборочное среднее квадратическое отклонение	Неотрицательный квадратный корень из выборочной дисперсии.	

2.2.7.	Выборочный момент порядка q	Момент порядка q случайной величины, имеющей эмпирическое распределение.	$m_q = 1/n \sum_{i=1}^{n} x_i^q$, где x_i по п.2.2.5.
2.2.8.	Выборочный центральный момент порядка q	Центральный момент порядка q случайной величины, имеющей эмпирическое распределение.	$m_q = 1/n \sum_{i=1}^{n} (x_i - x_{ср})^q$, где x_i и $x_{ср}$ по п.2.2.5.
2.2.9.	k-я порядковая статистика	k-й элемент $x_{(k)}$ в вариационном ряду, полученном из выборки объема n, элементы которой $x_1, x_2, ..., x_n$ расположены в порядке неубывания: $x_{(1)} \le x_{(2)} \le ... \le x_{(k)} \le ... \le x_{(n)}$.	
2.2.10.	Размах выборки	Разность между наибольшим и наименьшим значениями результатов наблюдений в выборке.	Если $x_{(1)}$ и $x_{(n)}$ - первая и n-ая порядковые статистики в выборке объема n, то размах $R = x_{(n)} - x_{(1)}$.
2.2.11.	Выборочная ковариация	Ковариация двумерного случайного вектора, имеющего эмпирическое распределение.	Если (x_i, y_i), i=1,2,....,n, - результаты наблюдений, включенные в выборку, то выборочная ковариация равна $1/n \sum_{i=1}^{n} (x_i - x_{ср})(y_i - y_{ср})$, где x_i и $x_{ср}$ по п.2.2.5, $y_{ср} = 1/n \sum_{i=1}^{n} y_i$.

2.2.12.	Выборочная ковариационная матрица	Ковариационная матрица случайного вектора, имеющего эмпирическое распределение.	На главной диагонали выборочной ковариационной матрицы стоят выборочные дисперсии по п.2.2.5, а вне главной диагонали - выборочные ковариации по п.2.2.11.
2.2.13.	Выборочный коэффициент корреляции	Коэффициент корреляции двумерного случайного вектора, имеющего эмпирическое распределение.	Выборочный коэффициент корреляции равен $$r_n = \frac{\sum_{1\le i\le n}(x_i - x_{cp})(y_i - y_{cp})}{\{\sum_{1\le i\le n}(x_i - x_{cp})^2 \sum_{1\le i\le n}(y_i - y_{cp})^2\}^{1/2}}$$ где x_i и x_{cp} по п.2.2.5, y_i и y_{cp} по п.2.2.11.
2.2.14.	Выборочная корреляционная матрица	Корреляционная матрица случайного вектора, имеющего эмпирическое распределение.	На главной диагонали выборочной корреляционной матрицы стоят 1, а вне главной диагонали - выборочные коэффициенты корреляции по п.2.2.13.
2.2.15	Выборочный коэффициент вариации	Отношение выборочного среднего квадратического отклонения к выборочному среднему арифметическому.	Выборочный коэффициент вариации используют, когда результаты наблюдений положительны.
		2.3. Оценивание	
2.3.1.	Оценивание	Приближенное определение интересующей специалиста составляющей вероятностной	Составляющими вероятностных моделей могут быть: значение параметра распределения; характеристика

		модели явления (процесса) по выборке.	распределения (математическое ожидание, коэффициент вариации и др.); функция распределения; плотность вероятности; регрессионная зависимость, и т.д.
2.3.2.	Оценка	Результат оценивания по конкретной выборке.	Оценка является статистикой, а потому случайным элементом, в частных случаях - случайной величиной или случайным вектором.
2.3.3.	Точечное оценивание	Вид оценивания, при котором для оценивания используется одно определенное значение.	
2.3.4.	Доверительное оценивание	Вид оценивания, при котором для оценивания используется множество.	Рассматриваемое множество лежит в пространстве возможных состояний оцениваемой составляющей вероятностной модели явления (процесса).
2.3.5.	Доверительное множество	Определяемое по выборке множество в пространстве возможных состояний оцениваемой составляющей, используемое при доверительном оценивании.	Доверительное множество является случайным множеством.
2.3.6.	Доверительная вероятность	Вероятность того, что доверительное множество содержит действительное значение оцениваемой составляющей.	В конкретных задачах оценивания для фиксированных доверительных вероятностей строят соответствующие доверительные множества.
2.3.7.	Доверительный интервал	Доверительное множество, являющееся интервалом.	Интервалы могут быть как ограниченными, так и неограниченными (лучами).
2.3.8.	Доверительные границы	Концы (границы) доверительного интервала.	

2.3.9.	Верхняя доверительная граница	Граница доверительного интервала, являющегося лучом, не ограниченным снизу.	Для доверительного интервала (-∞; а) верхней доверительной границей является число а.
2.3.10.	Нижняя доверительная граница	Граница доверительного интервала, являющегося лучом, не ограниченным сверху.	Различие верхних, нижних и двусторонних доверительных границ необходимо учитывать при проведении конкретных расчетов, т.к. часто все виды границ определяются с помощью одних и тех же таблиц.
2.3.11.	Двусторонние доверительные границы	Границы ограниченного (и сверху, и снизу) доверительного интервала	Для двусторонних границ (T_1;T_2) с вероятностью 1 справедливо неравенство $T_1 \leq T_2$.
		2.4. Проверка статистических гипотез	
2.4.1.	Статистическая гипотеза	Определенное предположение о свойствах распределений случайных элементов, лежащих в основе наблюдаемых случайных явлений (процессов).	
2.4.2.	Нулевая гипотеза	Статистическая гипотеза, подлежащая проверке по статистическим данным (результатам наблюдений, вошедшим в выборку).	Из возможных статистических гипотез в качестве нулевой выбирают ту, прннятие справедливости которой наиболее важно для дальнейших выводов.
2.4.3.	Альтернативная гипотеза	Статистическая гипотеза, которая считается справедливой, если нулевая гипотеза неверна.	Сокращенная форма - альтернатива.
2.4.4.	Статистический критерий	Правило, по которому на основе результатов наблюдений принимается решение о приня-	Принимаемое решение может однозначно определяться по результатам наблюдений (нерандомизированный

		тии или отклонении нулевой гипотезы.	критерий) или в некоторой степени зависеть от случая (рандомизированный критерий).
2.4.5.	Статистика критерия	Статистика, на основе которой сформулировано решающее правило.	Как правило, нерандомизированный статистический критерий основан на статистике критерия, принимающей числовые значения.
2.4.6.	Критическая область статистического критерия	Область в пространстве возможных выборок со следующими свойствами: если наблюдаемая выборка принадлежит данной области, то отвергают нулевую гипотезу (и принимают альтернативную), в противном случае ее принимают (и отвергают альтернативную).	Если статистический критерий основан на статистике критерия, то критическая область статистического критерия однозначно определяется по критической области статистики критерия. Краткая форма: критическая область.
2.4.7.	Критическая область статистики критерия	Множество чисел такое, что при попадании в него статистики критерия нулевую гипотезу отвергают, в противном случае принимают.	Краткая форма: критическая область.
2.4.8.	Критические значения	Границы (концы) одного или двух интервалов, составляющих критическую область статистики критерия.	Критическими значениями являются одно или два из чисел t_1, t_2 в случае, если критическая область имеет вид $\{T_n<t_1\}$, $\{T_n>t_1\}$ или $\{T_n<t_1\}\cup\{T_n>t_2\}$, где T_n - статистика критерия.
2.4.9.	Ошибка первого рода	Ошибка, заключающаяся в том, что нулевую гипотезу отвергают, в то время как в действительности эта гипотеза верна.	
2.4.10.	Уровень значи-	Вероятность ошибки первого рода или точная	Если нулевая гипотеза является сложной (например,

	мости	верхняя грань таких вероятностей.	задается с помощью множества параметров $\Theta_{0)}$, то вероятность ошибки первого рода может быть не числом (α), а функцией ($\alpha(\theta_0)$, $\theta_0 \in \Theta_0$). В качестве уровня значимости берут точную верхнюю грань значений указанной функции: $$\alpha = \sup_{\theta o \in \Theta o} \alpha(\theta o).$$
2.4.11.	Ошибка второго рода	Ошибка, заключающаяся в том, что нулевую гипотезу принимают, в то время как в действительности эта гипотеза неверна (а верна альтернативная гипотеза).	
2.4.12.	Мощность критерия	Вероятность того, что нулевая гипотеза будет отвергнута, если альтернативная гипотеза верна.	Мощность критерия является однозначной действительной функцией, определенной на составляющем альтернативу множестве гипотез, заданном в конкретной задаче статистической проверки гипотез, в частности, на параметрическом множестве, соответствующем альтернативным гипотезам.
2.4.13.	Функция мощности статистического критерия	Функция, определяющая вероятность того, что нулевая гипотеза будет отклонена.	Функция мощности критерия задана на множестве всех гипотез, используемых в конкретной задаче статистической проверки гипотез. Сужением ее на нулевую гипотезу является функция, задающая вероят-

			ность ошибки первого рода. Сужением ее на альтернативу является мощность критерия.
2.4.14.	Оперативная характеристика статистического критерия	Функция, определяющая вероятность того, что нулевая гипотеза будет принята.	Оперативная характеристика - дополнение до единицы функции мощности статистического критерия.
2.4.15.	Критерий согласия	Критерий проверки гипотезы согласия, т.е. того, что функция распределения результатов наблюдения, включенных в простую случайную выборку, совпадает с заданной или входит в заданное параметрическое семейство.	
2.4.16.	Критерий однородности	Критерий для проверки гипотезы о том, что функции распределений результатов наблюдений из двух или нескольких независимых простых случайных выборок совпадают (абсолютная однородность) или отдельные их характеристики совпадают (однородность в смысле математических ожиданий, коэффициентов вариации и т.д.).	Рассматривают также критерии независимости, симметрии, случайности, отбраковки и др.
2.4.17.	Номинальный (заданный) уро-	Число, используемое в статистических таблицах, с помощью которого выбирают критиче-	Номинальный (заданный) уровень значимости обычно берут равным 0,1; 0,05; 0,01.

	вень значимости	ское значение статистики критерия при проверке статистической гипотезы.	
2.4.18.	Реальный (истинный) уровень значимости	Уровень значимости статистического критерия, выбранного по номинальному уровню значимости.	Из-за дискретности распределения статистики критерия реальный уровень значимости может быть в несколько раз меньше номинального.
2.4.19.	Достигаемый уровень значимости	Случайная величина, равная вероятности попадания статистики критерия в критическую область, заданную рассчитанным по выборке значением статистики критерия.	Для критической области вида {x:x>a} достигаемый уровень значимости есть $F(X_n)$, где X_n - рассчитанное по выборке значение статистики критерия X, а $F(a) = P(X>a)$ - дополнение до 1 функции распределения статистики критерия X. Достигаемый уровень значимости - это вероятность того, что статистика критерия X в новом независимом эксперименте примет значение большее, чем при расчете по конкретной выборке, т.е. большее, чем X_n.
2.4.20.	Независимые выборки	Выборки, объединение элементов которых моделируется набором независимых (в совокупности) случайных элементов.	См. п.1.1.11.

П1-2. Математическая статистика и ее новые разделы

Приведем краткие описания (типа статей в энциклопедических изданиях) математической статистики и ее наиболее важных для эконометрики сравнительно новых разделов, разработанных в основном после 1970 г., а именно, статистики объектов нечисловой природы и статистики интервальных данных.

Статистика математическая - наука о математических методах анализа данных, полученных при проведении массовых наблюдений (измерений, опытов). В зависимости от математической природы конкретных результатов наблюдений *статистика математическая* делится на статистику чисел, многомерный статистический анализ, анализ функций (процессов) и временных рядов, статистику объектов нечисловой природы. Существенная часть *статистики математической* основана на вероятностных моделях.

Выделяют общие задачи описания данных, оценивания и проверки гипотез. Рассматривают и более частные задачи, связанные с проведением выборочных обследований, восстановлением зависимостей, построением и использованием классификаций (типологий) и др.

Для описания данных строят таблицы, диаграммы, иные наглядные представления, например, корреляционные поля. Вероятностные модели обычно не применяются. Некоторые методы описания данных опираются на продвинутую теорию и возможности современных компьютеров. К ним относятся, в частности, кластер-анализ, нацеленный на выделение групп объектов, похожих друг на друга, и многомерное шкалирование, позволяющее наглядно представить объекты на плоскости, в наименьшей степени исказив расстояния между ними.

Методы оценивания и проверки гипотез опираются на вероятностные модели порождения данных. Эти модели делятся на параметрические и непараметрические. В параметрических моделях предполагается, что изучаемые объекты описываются функциями распределения, зависящими от небольшого числа (1-4) числовых параметров. В непараметрических моделях функции распределения предполагаются произвольными непрерывными. В *статистике математической* оценивают параметры и характеристики распределения (математическое ожидание, медиану, дисперсию, квантили и др.), плотности и функции распределения, зависимости между переменными (на основе линейных и непараметрических коэффициентов корреляции, а также параметрических или непараметрических оценок функций, выражающих зависимости) и др. Используют точечные и интервальные (дающие границы для истинных значений) оценки.

В *статистике математической* есть общая теория проверки гипотез и большое число методов, посвященных проверке конкретных гипотез. Рассматривают гипотезы о значениях параметров и характеристик, о проверке однородности (т.е. о совпадении характеристик или функций распределения в двух выборках), о согласии эмпирической функции распределения с заданной функцией распределения или с параметрическим семейством таких функций, о симметрии распределения и др.

Большое значение для эконометрики имеет раздел *статистики математической,* связанный с проведением выборочных обследований, со свойствами различных схем организации выборок и построением адекватных методов оценивания и проверки гипотез.

Задачи восстановления зависимостей активно изучаются более 200 лет, с момента разработки К. Гауссом в 1794 г. метода наименьших квадратов. В настоящее время наиболее актуальны методы поиска информативного подмножества переменных и непараметрические методы.

Различные методы построения (кластер-анализ), анализа и использования (дискриминантный анализ) классификаций (типологий) именуют также методами распознавания образов (с учителем и без), автоматической классификации и др.

Математические методы в статистике основаны либо на использовании сумм (на основе Центральной Предельной Теоремы теории вероятностей) или показателей различия (расстояний, метрик), как в статистике объектов нечисловой природы. Строго обоснованы обычно лишь асимптотические результаты. В настоящее время компьютеры играют большую роль в *статистике математической.* Они используются как для расчетов, так и для имитационного моделирования (в частности, в методах размножения выборок и при изучении пригодности асимптотических результатов).

Классическая *статистика математическая* лучше всего представлена в [2,4]. По историческим причинам основные российские работы публикуются в [3]. Обзор современного состояния *статистики математической* дан в [6].

Статистика объектов нечисловой природы - раздел математической статистики, в котором статистическими данными являются объекты нечисловой природы, т.е. элементы множеств, не являющихся линейными пространствами. Объекты нечисловой природы нельзя складывать и умножать на число. Примерами являются результаты измерений в шкалах наименований, порядка, интервалов; ранжировки, разбиения, толерантности и другие бинарные отношения; результаты парных и множественных сравнений; люсианы, т.е. конечные последовательности из 0 и1; множества; нечеткие множества. Необходимость применения объектов нечисловой природы возникает во многих областях научной и практической деятельности, в том числе и в социологии. Примерами являются ответы на "закрытые" вопросы в эконометрических, маркетинговых, социологических анкетах, в которых респондент должен выбрать одну или несколько из фиксированного числа подсказок, мили измерение мнений о привлекательности (товаров, услуг, профессий, политиков и др.), проводимое по порядковой шкале. Наряду со специальными теориями для каждого отдельного вида объектов нечисловой природы в с*татистике объектов нечисловой природы* имеется и теория обработки данных, лежащих в пространстве общей природы, результаты которой применимы во всех специальных теориях.

В *статистике объектов нечисловой природы* классические задачи математической статистики - описание данных, оценивание, проверку гипотез - рассматривают для данных неклассического типа, что приводит к своеобразию постановок задач и методов их решения. Например, из-за отсутствия линейной структуры в пространстве, в котором лежат статистические данные, в с*татистике объектов нечисловой природы* математическое ожидание определяют не через сумму или интеграл, как в классическом случае, а как решение задачи минимизации некоторой функции. Эта функция представляет собой математическое ожидание (в классическом смысле) показателя различия между значением случайного объекта нечисловой природы и фиксированным элементом пространства. Эмпирическое среднее определяют как результат минимизации суммы расстояний от нечисловых результатов наблюдений до фиксированного элемента пространства. Справедлив закон больших чисел: эмпирическое среднее сходится при увеличении объема выборки к математическому ожиданию, если результаты наблюдений являются независимыми одинаково распределенными случайными объектами нечисловой природы и выполнены некоторые математические "условия регулярности".

Аналогичным образом определяют условное математическое ожидание и регрессионную зависимость. Из доказанной в с*татистике объектов нечисловой природы* сходимости решений экстремальных статистических задач к решениям соответствующих предельных задач вытекает состоятельность оценок в параметрических задачах оценивания параметров и аппроксимации, а также ряд результатов в многомерном статистическом анализе. Большую роль в с*татистике объектов нечисловой природы* играют непараметрические методы, в частности, методы непараметрической оценки плотности и регрессионной зависимости в пространствах общей природы, в том числе и в дискретных пространствах.

Для решения многих задач *статистики объектов нечисловой природы* - нахождения эмпирического среднего, оценки регрессионной зависимости, классификации наблюдений и др. - используют показатели различия (меры близости, расстояния, метрики) между элементами рассматриваемых пространств, вводимые аксиоматически. Так, в монографии [7] аксиоматически введено расстояние между множествами. Принятое в теории измерений как части *статистики объектов нечисловой природы* условие адекватности (инвариантности) алгоритмов анализа данных позволяет указать вид средних величин, расстояний, показателей связи и т.д., соответствующих измерениям в тех или иных шкалах. Методы построения, анализа и использования классификаций и многомерного шкалирования дают возможность сжать информацию и дать ей наглядное представление. К *статистике объектов нечисловой природы* относятся методы ранговой корреляции, статистического анализа бинарных отношений (ранжировок, разбиений, толерантностей), параметрические и непараметрические методы обработки результатов парных и множественных сравнений. Теория люсианов (последовательностей независимых испытаний Бернулли) развита в асимптотике растущей размерности.

Статистика объектов нечисловой природы как самостоятельный раздел прикладной математической статистики выделена в монографии [7]. Обзору ее основных направлений посвящен, например, сборник [8]. Ей посвящен раздел в энциклопедии [2].

Статистика интервальных данных (СИД) - раздел *статистики объектов нечисловой природы*, в котором элементами выборки являются интервалы в R, в частности, порожденные наложением ошибок измерения на значения случайных величин. СИД входит в теорию устойчивости (робастности) статистических процедур (см. [7]) и примыкает к интервальной математике (см. [9]). В СИД изучены проблемы регрессионного анализа, планирования эксперимента, сравнения альтернатив и принятия решений в условиях интервальной неопределенности и др. (см.[10-13]).

Развиты асимптотические методы статистического анализа интервальных данных при больших объемах выборок и малых погрешностях измерений. В отличие от классической математической статистики, сначала устремляется к бесконечности объем выборки и только потом - уменьшаются до нуля погрешности. Разработана общая схема исследования (см. [14]), включающая расчет двух основных характеристик СИД - н о т н ы (максимально возможного отклонения статистики, вызванного интервальностью исходных данных) и р а ц и о н а л ь н о г о о б ъ е м а в ы б о р к и (превышение которого не дает существенного повышения точности оценивания и статистических выводов, связанных с проверкой гипотез). Она применена к оцениванию математического ожидания и дисперсии, медианы и коэффициента вариации, параметров гамма-распределения в ГОСТ 11.011-83 [15] и характеристик аддитивных статистик, для проверки гипотез о параметрах нормального распределения, в т.ч. с помощью критерия Стьюдента, а также гипотезы однородности двух выборок по критерию Смирнова, и т.д.. Разработаны подходы СИД в основных постановках регрессионного, дискриминантного и кластерного анализов (см. [16]).

Многие утверждения СИД отличаются от аналогов из классической математической статистики. В частности, не существует состоятельных оценок: средний квадрат ошибки оценки, как правило, асимптотически равен сумме дисперсии этой оценки, рассчитанной согласно классической теории, и квадрата нотны. Метод моментов иногда оказывается точнее метода максимального правдоподобия (см. [15, 17]). Нецелесообразно с целью повышения точности выводов увеличивать объем выборки сверх некоторого предела. В СИД классические доверительные интервалы должны быть расширены вправо и влево на величину нотны, и длина их не стремится к 0 при росте объема выборки.

Многим задачам классической математической статистики могут быть поставлены в соответствие задачи СИД, в которых элементы выборок - действительные числа заменены на интервалы. В статистическое программное обеспечение включают алгоритмы СИД, "параллельные" их аналогам из классической математической статистики. Это позволяет учесть наличие погрешностей у результатов наблюдений.

Цитированная литература

1. Колмогоров А.Н. Основные понятия теории вероятностей. 2-е изд. - М.: Наука, 1974. - 120 с.
2. Вероятность и математическая статистика. Энциклопедия / Гл. ред. Ю. В. Прохоров. – М.: Изд-во «Большая Российская Энциклопедия», 1999. – 910 с.
3. Орлов А.И. Термины и определения в области вероятностно-статистических методов. – Журнал «Заводская лаборатория». 1999. Т.65. No.7. С.46-54.
4. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. - М.: Наука, 1983.
5. Секция "Математические методы исследования" журнала "Заводская лаборатория. Диагностика материалов".
6. Орлов А.И. Современная прикладная статистика. - Журнал "Заводская лаборатория". 1998. Т.64. No.3. С. 52-60.
7. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979. - 296 с.
8. Анализ нечисловой информации в социологических исследованиях. - М.: Наука, 1985. - 220 с.
9. Шокин Ю.И. Интервальный анализ. - Новосибирск: Наука, 1981. - 112 с.
10. Вощинин А.П. Метод оптимизации объектов по интервальным моделям целевой функции. - М.: МЭИ, 1987. - 109 с.
11. Вощинин А.П., Сотиров Г.Р. Оптимизация в условиях неопределенности. - М.: МЭИ - София: Техника, 1989. - 224 с.
12. Кузнецов В.П. Интервальные статистические модели. - М.: Радио и связь, 1991. - 352 с.
13. Сборник трудов Международной конференции по интервальным и стохастическим методам в науке и технике (ИНТЕРВАЛ-92). Тт. 1,2. - М.: МЭИ, 1992. - 216 с., 152 с.
14. Орлов А.И. О развитии реалистической статистики. - В сб.: Статистические методы оценивания и проверки гипотез. Межвузовский сборник научных трудов. Пермь: Изд-во Пермского государственного университета, 1990, с..89-99.
15. ГОСТ 11.011-83. Прикладная статистика. Правила определения оценок и доверительных границ для параметров гамма-распределения. - М.: Изд-во стандартов, 1984. - 53 с.
16. Орлов А.И. Интервальный статистический анализ. - В сб.: Статистические методы оценивания и проверки гипотез. Межвузовский сборник научных трудов. Пермь: Пермский государственный университет, 1993, с.149-158.
17. Орлов А.И. Интервальная статистика: метод максимального правдоподобия и метод моментов. - В сб.: Статистические методы оценивания и проверки гипотез. Межвузовский сборник научных трудов. - Пермь: Изд-во Пермского государственного университета, 1995, с.114-124.

Приложение 2

Нечеткие и случайные множества

В главе 8 рассматривались такие виды объектов нечисловой природы, как нечеткие и случайные множества. Цель настоящего приложения - глубже изучить свойства нечетких множеств и показать, что теория нечетких множеств в определенном смысле сводится к теории случайных множеств. Для достижения поставленной цели формулируется и доказывается цепь теорем.

В дальнейшем считается, что все рассматриваемые нечеткие множества являются подмножествами одного и того же множества Y.

П2-1. Законы де Моргана для нечетких множеств

Как известно, законами же Моргана называются следующие тождества алгебры множеств

$$\overline{A \bigcup B} = \overline{A} \bigcap \overline{B}, \quad \overline{A \bigcap B} = \overline{A} \bigcup \overline{B}. \qquad (1)$$

Теорема 1. *Для нечетких множеств справедливы тождества*

$$\overline{A \bigcup B} = \overline{A} \bigcap \overline{B}, \quad \overline{A \bigcap B} = \overline{A} \bigcup \overline{B}, \qquad (2)$$

$$\overline{A + B} = \overline{A}\,\overline{B}, \quad \overline{AB} = \overline{A} + \overline{B}. \qquad (3)$$

Доказательство теоремы 1 состоит в непосредственной проверке справедливости соотношений (2) и (3) путем вычисления значений функций принадлежности участвующих в этих соотношениях нечетких множеств на основе определений, данных в главе 8.

Тождества (2) и (3) назовем *законами де Моргана для нечетких множеств*. В отличие от классического случая соотношений (1), они состоят из четырех тождеств, одна пара которых относится к операциям объединения и пересечения, а вторая - к операциям произведения и суммы. Как и соотношение (1) в алгебре множеств, законы де Моргана в алгебре нечетких множеств позволяют преобразовывать выражения и формулы, в состав которых входят операции отрицания.

П2-2. Дистрибутивный закон для нечетких множеств

Некоторые свойства операций над множествами не выполнены для нечетких множеств. Так, $A + A \neq A$, за исключением случая, когда A - "четкое" множество (т.е. функция принадлежности принимает только значения 0 и 1).

Верен ли дистрибутивный закон для нечетких множеств? В литературе иногда расплывчато утверждается, что "не всегда". Внесем полную ясность.

Теорема 2. *Для любых нечетких множеств A, B и C*

$$A \bigcap (B \bigcup C) = (A \bigcap B) \bigcup (A \bigcap C). \qquad (4)$$

В то же время равенство

$$A(B + C) = AB + AC \qquad (5)$$

справедливо тогда и только тогда, когда при всех $y \in Y$

$$(\mu_A^2(y) - \mu_A(y))\mu_B(y)\mu_C(y) = 0.$$

Доказательство. Фиксируем произвольный элемент $y \in Y$. Для сокращения записи обозначим $a = \mu_A(y), b = \mu_B(y), c = \mu_C(y)$. Для доказательства тождества (4) необходимо показать, что

$$\min(a, \max(b, c)) = \max(\min(a, b), \min(a, c)). \qquad (6)$$

Рассмотрим различные упорядочения трех чисел *a, b, c*. Пусть сначала $a \le b \le c$. Тогда левая часть соотношения (6) есть $\min(a,c) = a$, а правая $\max(a,a) = a$, т.е. равенство (6) справедливо.

Пусть $b \le a \le c$. Тогда в соотношении (6) слева стоит $\min(a,c) = a$, а справа $\max(b,a) = a$, т.е. соотношение (6) опять является равенством.

Если $b \le c \le a$, то в соотношении (6) слева стоит $\min(a,c) = c$, а справа $\max(b,c) = c$, т.е. обе части снова совпадают.

Три остальные упорядочения чисел *a, b, c* разбирать нет необходимости, поскольку в соотношение (6) числа *b* и *c* входят симметрично. Тождество (4) доказано.

Второе утверждение теоремы 2 вытекает из того, что в соответствии с определениями операций над нечеткими множествами (см. главу 8)

$$\mu_{A(B+C)}(y) = a(b + c - bc) = ab + ac - abc$$

и

$$\mu_{AB+AC}(y) = ab + ac - (ab)(ac) = ab + ac - a^2bc.$$

Эти два выражения совпадают тогда и только тогда, когда, когда $a^2bc = abc$, что и требовалось доказать.

Определение 1. *Носителем нечеткого множества A называется совокупность всех точек* $y \in Y$ *, для которых* $\mu_A(y) > 0$.

Следствие теоремы 2. *Если носители нечетких множеств B и C совпадают с У, то равенство (5) имеет место тогда и только тогда, когда A - "четкое" (т.е. обычное, классическое, не нечеткое) множество.*

Доказательство. По условию $\mu_B(y)\mu_C(y) \ne 0$ при всех $y \in Y$. Тогда из теоремы 2 следует, что $\mu_A^2(y) - \mu_A(y) = 0$, т.е. $\mu_A(y) = 1$ или $\mu_A(y) = 0$, что и означает, что *A* - четкое множество.

П2-3. Нечеткие множества как проекции случайных множеств

С самого начала появления современной теории нечеткости в 1960-е годы началось обсуждение ее взаимоотношений с теорией вероятностей. Дело в том, что функция принадлежности нечеткого множества напоминает распределение вероятностей. Отличие только в том, что сумма вероятностей по всем возможным значениям случайной величины (или интеграл, если множество возможных значений несчетно) всегда равна 1, а сумма *S* значений функции принадлежности (в непрерывном случае - интеграл от функции принадлежности) может быть любым неотрицательным числом. Возникает искушение пронормировать функцию принадлежности, т.е. разделить все ее значения на *S* (при *S* $\ne 0$), чтобы свести ее к распределению вероятностей (или к плотности вероятности). Однако специалисты по нечеткости справедливо возражают против такого "примитивного" сведения", поскольку оно проводится отдельно для каждой размытости (нечеткого множества), и определения обычных операций над нечеткими множествами с ним согласовать нельзя. Последнее утверждение означает следующее. Пусть указанным образом преобразованы функции принадлежности нечетких множеств *A* и *B*. Как при этом преобразуются функции принадлежности $A \cap B, A \cup B, A + B, AB$? Установить это *невозможно в принципе.* Последнее утверждение становится совершенно ясным после рассмотрения нескольких примеров пар нечетких множеств с одними и теми же суммами значений функций принадлежности, но различными результатами теоретико-множественных операций над ними, причем и суммы значений соответствующих функций принадлежности для этих результатов теоретико-множественных операций, например, для пересечений множеств, также различны.

В работах по нечетким множествам довольно часто утверждается, что теория нечеткости является самостоятельным разделом прикладной математики и не имеет отношения к теории вероятностей (см., например, обзор литературы в монографиях [1,2]). Авторы, сравнивавшие теорию нечеткости и теорию вероятностей, обычно подчеркивали различие между этими областями теоретических и прикладных исследований. Обычно сравнивают аксиоматику и сравнивают области приложений. Надо сразу отметить, что аргументы при втором типе сравнений не имеют доказательной силы, поскольку по поводу границ применимости даже такой давно выделившейся научной области, как вероятностно-статистические методы, имеются различные мнения. Напомним, что итог рассуждений одного из наиболее известных французских математиков Анри Лебега по поводу границ применимости арифметики таков: "Арифметика применима тогда, когда она применима" (см. его монографию [3, с.21-22]).

При сравнении различных аксиоматик теории нечеткости и теории вероятностей нетрудно увидеть, что списки аксиом различаются. Из этого, однако, отнюдь не следует, что между указанными теориями нельзя установить связь, типа известного сведения евклидовой геометрии на плоскости к арифметике (точнее к теории числовой системы R^2 - см., например, монографию [4]). Напомним, что эти две аксиоматики - евклидовой геометрии и арифметики - на первый взгляд весьма сильно различаются.

Можно понять желание энтузиастов нового направления подчеркнуть принципиальную новизну своего научного аппарата. Однако не менее важно установить связи нового подхода с ранее известными.

Как оказалось, теория нечетких множеств тесно связана с теорией случайных множеств. Еще в 1974 г. в работе [5] было показано, что нечеткие множества естественно рассматривать как "проекции" случайных множеств. Рассмотрим этот метод сведения теории нечетких множеств к теории случайных множеств.

Определение 2. *Пусть* $A = A(\omega)$ *- случайное подмножество конечного множества У. Нечеткое множество В, определенное на У, называется проекцией А и обозначается Proj A, если*

$$\mu_B(y) = P(y \in A) \qquad (7)$$

при всех $y \in Y$.

Очевидно, каждому случайному множеству *A* можно поставить в соответствие с помощью формулы (7) нечеткое множество *B = Proj A*. Оказывается, верно и обратное.

Теорема 3. *Для любого нечеткого подмножества В конечного множества У существует случайное подмножество А множества У такое, что В = Proj A.*

Доказательство. Достаточно задать распределение случайного множества *А*. Пусть $У_1$ - носитель *В* (см. определение 1 выше). Без ограничения общности можно считать, что $Y_1 = \{y_1, y_2, ..., y_m\}$ при некотором *m* и элементы $У_1$ занумерованы в таком порядке, что

$$0 < \mu_B(y_1) \leq \mu_B(y_2) \leq ... \leq \mu_B(y_m).$$

Введем множества

$$Y(1) = Y_1, Y(2) = \{y_2, ..., y_m\}, ..., Y(t) = \{y_t, ..., y_m\}, ..., Y(m) = \{y_m\}.$$

Положим

$$P(A = Y(1)) = \mu_B(y_1), \quad P(A = Y(2)) = \mu_B(y_2) - \mu_B(y_1), ...,$$

$$P(A = Y(t)) = \mu_B(y_t) - \mu_B(y_{t-1}), ..., P(A = Y(m)) = \mu_B(y_m) - \mu_B(y_{m-1}),$$

$$P(A = \varnothing) = 1 - \mu_B(y_m).$$

Для всех остальных подмножеств *X* множества *У* положим *P(A=X)=0*. Поскольку элемент y_t входит во множества *Y(1), Y(2),…, Y(t)* и не входит во множества *Y(t+1),…, Y(m),* то из приведенных выше формул следует, что $P(y_t \in A) = \mu_B(y_t)$. Если $y \notin Y_1$, то, очевидно, $P(y \in A) = 0$. Теорема 3 доказана.

Распределение случайного множества с независимыми элементами, как следует из рассмотрений главы 8, полностью определяется его проекцией. Для конечного случайного множества общего вида это не так. Для уточнения сказанного понадобится следующая теорема.

Теорема 4. *Для случайного подмножества А множества У из конечного числа элементов наборы чисел* $P(A=X), X\subseteq Y$, *и* $P(X\subseteq A), X\subseteq Y$, *выражаются один через другой.*

Доказательство. Второй набор выражается через первый следующим образом:

$$P(X\subseteq A)=\sum_{X':X\subseteq X'} P(A=X').$$

Элементы первого набора выразить через второй можно с помощью формулы включений и исключений из формальной логики, в соответствии с которой $P(A=X)=P(X\subseteq A)-\sum P(X\bigcup\{y\}\subseteq A)+\sum P(X\bigcup\{y_1,y_2\}\subseteq A)-...\pm P(Y\subseteq A)$. В этой формуле в первой сумме *y* пробегает все элементы множества *Y\X,* во второй сумме переменные суммирования y_1 и y_2 не совпадают и также пробегают это множество, и т.д. Ссылка на формулу включений и исключений завершает доказательство теоремы 4.

В соответствии с теоремой 4 случайное множество А можно характеризовать не только распределением, но и набором чисел $P(X\subseteq A), X\subseteq Y$. В этом наборе $P(\varnothing\subseteq A)=1$, а других связей типа равенств нет. В этот набор входят числа $P(\{y\}\subseteq A)=P(y\in A)$, следовательно, фиксация проекции случайного множества эквивалентна фиксации *k = Card(Y)* параметров из (2^k-1) параметров, задающих распределение случайного множества *A* в общем случае.

Будет полезна следующая теорема.

Теорема 5. *Если Proj A = B, то* $\mathrm{Pr}oj\overline{A}=\overline{B}$.

Для доказательства достаточно воспользоваться тождеством из теории случайных множеств $P(\overline{A}=X)=P(A=\overline{X})$, формулой для вероятности накрытия $P(y\in A)$ из главы 8, определением отрицания нечеткого множества и тем, что сумма всех P(A=X) равна 1.

П2-4. Пересечения и произведения нечетких и случайных множеств

Выясним, как операции над случайными множествами соотносятся с операциями над их проекциями. В силу законов де Моргана (теорема 1) и теоремы 5 достаточно рассмотреть операцию пересечения случайных множеств.

Теорема 6. *Если случайные подмножества* A_1 *и* A_2 *конечного множества* У *независимы, то нечеткое множество* $\mathrm{Pr}oj(A_1\cap A_2)$ *является произведением нечетких множеств Proj* A_1 *и Proj* A_2.

Доказательство. Надо показать, что для любого $y\in Y$

$$P(y\in A_1\cap A_2)=P(y\in A_1)P(y\in A_2). \qquad (8)$$

По формуле для вероятности накрытия точки случайным множеством (глава 8)

$$P(y\in A_1\cap A_2)=\sum_{X:y\in X} P((A_1\cap A_2)=X). \qquad (9)$$

Как известно, распределение пересечения случайных множеств $A_1\cap A_2$ можно выразить через их совместное распределение следующим образом:

$$P(A_1\cap A_2=X)=\sum_{X_1,X_2:X_1\cap X_2=X} P(A_1=X_1,A_2=X_2). \qquad (10)$$

Из соотношений (9) и (10) следует, что вероятность накрытия для пересечения случайных множеств можно представить в виде двойной суммы

$$P(y\in A_1\cap A_2)=\sum_{X:y\in X}\ \sum_{X_1,X_2:X_1\cap X_2=X} P(A_1=X_1,A_2=X_2). \qquad (11)$$

Заметим теперь, что правую часть формулы (11) можно переписать следующим образом:

$$\sum_{X_1, X_2 : e \in X_1, e \in X_2} P(A_1 = X_1, A_2 = X_2). \quad (12)$$

Действительно, формула (11) отличается от формулы (12) лишь тем, что в ней сгруппированы члены, в которых пересечение переменных суммирования $X_1 \cap X_2$ принимает постоянное значение. Воспользовавшись определением независимости случайных множеств и правилом перемножения сумм, получаем, что из (11) и (12) вытекает равенство

$$P(y \in A_1 \cap A_2) = \left(\sum_{X_1 : y \in X_1} P(A_1 = X_1) \right) \left(\sum_{X_2 : y \in X_2} P(A_2 = X_2) \right).$$

Для завершения доказательства теоремы 6 достаточно еще раз сослаться на формулу для вероятности накрытия точки случайным множеством (глава 8).

Определение 3. *Носителем случайного множества С называется совокупность всех тех элементов* $y \in Y$, *для которых* $P(y \in C) > 0$.

Теорема 7. *Равенство*

$$\Pr oj(A_1 \cap A_2) = (\Pr ojA_1) \cap (\Pr ojA_2)$$

верно тогда и только тогда, когда пересечение носителей случайных множеств $\overline{A_1} \cap A_2$ *и* $A_1 \cap \overline{A_2}$ *пусто.*

Доказательство. Необходимо выяснить условия, при которых

$$P(y \in A_1 \cap A_2) = \min(P(y \in A_1), P(y \in A_2)). \quad (13)$$

Положим

$$p_1 = P(y \in A_1 \cap A_2), p_2 = P(y \in \overline{A_1} \cap A_2), p_3 = P(y \in A_1 \cap \overline{A_2}).$$

Тогда равенство (13) сводится к условию

$$p_1 = \min(p_1 + p_2, p_1 + p_3). \quad (14)$$

Ясно, что соотношение (14) выполнено тогда и только тогда, когда $p_2 p_3 = 0$ при всех $y \in Y$, т.е. не существует ни одного элемента $y_0 \in Y$ такого, что одновременно $P(y_0 \in \overline{A_1} \cap A_2) > 0$ и $P(y_0 \in A_1 \cap \overline{A_2}) > 0$, а это эквивалентно пустоте пересечения носителей случайных множеств $\overline{A_1} \cap A_2$ и $A_1 \cap \overline{A_2}$. Теорема 7 доказана.

П2-5. Сведение последовательности операций над нечеткими множествами к последовательности операций над случайными множествами

Выше получены некоторые связи между нечеткими и случайными множествами. Стоит отметить, что изучение этих связей в работе [5] (эта работа выполнена в 1974 г. и доложена на семинаре "Многомерный статистический анализ и вероятностное моделирование реальных процессов" 18 декабря 1974 г. - см. [5, с.169]) началось с введения случайных множеств с целью развития и обобщения аппарата нечетких множеств Л. Заде. Дело в том, что математический аппарат нечетких множеств не позволяет в должной мере учитывать различные варианты зависимости между понятиями (объектами), моделируемыми с его помощью, не является достаточно гибким. Так, для описания "общей части" двух нечетких множеств есть лишь две операции - произведение и пересечение. Если применяется первая из них, то фактически предполагается, что множества ведут себя как проекции независимых случайных множеств (см. выше теорему 6). Операция пересечения также накладывает вполне определенные ограничения на вид зависимости между множествами (см. выше теорему 7), причем в этом случае найдены даже необходимые и достаточные условия. Желательно иметь более широкие возможности для моделирования зависимости между множествами (понятиями, объектами). Использование математического аппарата случайных множеств предоставляет такие возможности.

Цель сведения нечетких множеств к случайным состоит в том, чтобы за любой конструкцией из нечетких множеств видеть конструкцию из случайных множеств, определяющую свойства первой, аналогично тому, как плотностью распределения вероятностей мы видим случайную величину. В настоящем пункте приводим результаты по сведению алгебры нечетких множеств к алгебре случайных множеств.

Определение 4. *Вероятностное пространство {Ω, G, P} назовем делимым, если для любого измеримого множества* $X \in G$ *и любого положительного числа* α, *меньшего P(X), можно указать измеримое множество* $Y \subset X$ *такое, что* $P(Y) = \alpha$.

Пример. Пусть Ω - единичный куб конечномерного линейного пространства, *G* есть сигма-алгебра борелевских множеств, а *P* - мера Лебега. Тогда *{Ω, G, P}* - делимое вероятностное пространство.

Таким образом, делимое вероятностное пространство - это не экзотика. Обычный куб является примером такого пространства.

Доказательство сформулированного в примере утверждения проводится стандартными математическими приемами, основанными на том, что измеримое множество можно сколь угодно точно приблизить открытыми множествами, последние представляются в виде суммы не более чем счетного числа открытых шаров, а для шаров делимость проверяется непосредственно (от шара X тело объема $\alpha < P(X)$ отделяется соответствующей плоскостью).

Теорема 8. *Пусть даны случайное множество A на делимом вероятностном пространстве {Ω, G, P} со значениями во множестве всех подмножеств множества У из конечного числа элементов, и нечеткое множество D на У. Тогда существуют случайные множества* C_1, C_2, C_3, C_4 *на том же вероятностном пространстве такие, что*

$$\Pr oj(A \cap C_1) = B \cap D, \quad \Pr oj(A \cap C_2) = BD, \quad \Pr oj(A \cup C_3) = B \cup D,$$

$$\Pr oj(A \cup C_4) = B + D, \quad \Pr ojC_i = D, \; i = 1,2,3,4,$$

где B = Proj A.

Доказательство. В силу справедливости законов де Моргана для нечетких (см. теорему 1 выше) и для случайных множеств, а также теоремы 5 выше (об отрицаниях) достаточно доказать существование случайных множеств C_1 и C_2.

Рассмотрим распределение вероятностей во множестве всех подмножеств множества *У*, соответствующее случайному множеству *C* такому, что *Proj C = D* (оно существует в силу теоремы 3). Построим случайное множество C_2 с указанным распределением, независимое от *A*. Тогда $\Pr oj(A \cap C_2) = BD$ по теореме 6.

Перейдем к построению случайного множества C_1. По теореме 7 необходимо и достаточно определить случайное множество $C_1(\omega)$ так, чтобы $ProjC_1 = D$ и пересечение носителей случайных множеств $A \cap \overline{C_1}$ и $\overline{A} \cap C_1$ было пусто, т.е.

$$p_3 = P(y \in A \cap \overline{C_1}) = 0$$

для $y \in Y_1 = \{y : \mu_B(y) \le \mu_D(y)\}$ и

$$p_2 = P(y \in \overline{A} \cap C_1) = 0$$

для $y \in Y_2 = \{y : \mu_B(y) \ge \mu_D(y)\}$.

Построим $C_1(\omega)$, исходя из заданного случайного множества $A(\omega)$. Пусть $y_1 \in Y_2$. Исключим элемент y_1 из $A(\omega)$ для стольких элементарных событий ω, чтобы для полученного случайного множества $A_1(\omega)$ было справедливо равенство

$$P(y_1 \in A_1) = \mu_D(y_1)$$

(именно здесь используется делимость вероятностного пространства, на котором задано случайное множество $A(\omega)$). Для $y \neq y_1$, очевидно,

$$P(y \in A_1) = P(y \in A).$$

Аналогичным образом последовательно исключаем *y* из $A(\omega)$ для всех $y \in Y_2$ и добавляем *y* в $A(\omega)$ для всех $y \in Y_1$, меняя на каждом шагу $P(y \in A_i)$ только для $y = y_i$ так, чтобы

$$P(y_i \in A_i) = \mu_D(y_i)$$

(ясно, что при рассмотрении $y_i \in Y_1 \cap Y_2$ случайное множество $A_i(\omega)$ не меняется). Перебрав все элементы *У*, получим случайное множество $A_k(\omega) = C_1(\omega)$, для которого выполнено требуемое. Теорема 8 доказана.

Основной результат о сведении теории нечетких множеств к теории случайных множеств дается следующей теоремой.

Теорема 9. *Пусть* $B_1, B_2, B_3, ..., B_t$ *- некоторые нечеткие подмножества множества У из конечного числа элементов. Рассмотрим результаты последовательного выполнения теоретико-множественных операций*

$$B^m = ((...((B_1 \circ B_2) \circ B_3) \circ ...) \circ B_{m-1}) \circ B_m, \quad m = 1,2,...,t,$$

где $\circ$ *- символ одной из следующих теоретико-множественных операций над нечеткими множествами: пересечение, произведение, объединение, сумма (на разных местах могут стоять разные символы). Тогда существуют случайные подмножества* $A_1, A_2, A_3, ..., A_t$ *того же множества У такие, что*

$$\mathrm{Pr}\,oj A_i = B_i, \ \ i = 1,2,...,t,$$

и, кроме того, результаты теоретико-множественных операций связаны аналогичными соотношениями

$$\mathrm{Pr}\,oj\{((...((A_1 \otimes A_2) \otimes A_3) \otimes ...) \otimes A_{m-1}) \otimes A_m\} = B^m, \ \ m = 1,2,...,t,$$

где знак $\otimes$ *означает, что на рассматриваемом месте стоит символ пересечения* $\cap$ *случайных множеств, если в определении* B^m *стоит символ пересечения или символ произведения нечетких множеств, и соответственно символ объединения* $\cup$ *случайных множеств, если в* B^m *стоит символ объединения или символ суммы нечетких множеств.*

Комментарий. Поясним содержание теоремы. Например, если

$$B^5 = (((B_1 + B_2) \cap B_3) B_4) \cup B_5,$$

то

$$(((A_1 \otimes A_2) \otimes A_3) \otimes A_4) \otimes A_5 = (((A_1 \cup A_2) \cap A_3) \cap A_4) \cup A_5.$$

Как совместить справедливость дистрибутивного закона для случайных множеств (вытекающего из его справедливости для обычных множеств) с теоремой 2 выше, в которой показано, что для нечетких множеств, вообще говоря, $(B_1 + B_2)B_3 \neq B_1 B_3 + B_2 B_3$? Дело в том, что хотя в соответствии с теоремой 9 для любых трех нечетких множеств B_1, B_2 и B_3 можно указать три случайных множества A_1, A_2 и A_3 такие, что

$$\mathrm{Pr}\,oj(A_i) = B_i, \ \ i = 1,2,3, \ \ \mathrm{Pr}\,oj(A_1 \cup A_2) = B_1 + B_2, \ \ \mathrm{Pr}\,oj((A_1 \cup A_2) \cap A_3) = B^3,$$

где

$$B^3 = (B_1 + B_2) B_3,$$

но при этом, вообще говоря,

$$\mathrm{Pr}\,oj(A_1 \cap A_3) \neq B_1 B_3$$

и, кроме случаев, указанных в теореме 2,

$$\mathrm{Pr}\,oj((A_1 \cup A_2) \cap A_3) \neq B_1 B_3 + B_2 B_3.$$

Доказательство теоремы 9 проводится по индукции. При *t*=1 распределение случайного множества строится с помощью теоремы 3. Затем конструируется само случайное множество A_1, определенное на делимом вероятностном пространстве (нетрудно проверить, что на делимом вероятностном пространстве можно построить случайное подмножество конечного множества с любым заданным распределением именно в силу делимости пространства). Далее случайные множества A_2, A_3, ..., A_t строим по индукции с помощью теоремы 8. Теорема 9 доказана.

Замечание. Проведенное доказательство теоремы 9 проходит и в случае, когда при определении B^m используются отрицания, точнее, кроме B^m ранее введенного вида используются также последовательности результатов теоретико-множественных операций, очередной шаг в которых имеет вид

$$B_1^m = \overline{B^{m-1}} \circ B_m, \quad B_2^m = B^{m-1} \circ \overline{B_m}, \quad B_3^m = \overline{B^{m-1}} \circ \overline{B_m}.$$

А именно, сначала при помощи законов де Моргана (теорема 1 выше) проводится преобразование, в результате которого в последовательности B^m остаются только отрицания отдельных подмножеств из совокупности $B_1, B_2, B_3, ..., B_t$, а затем с помощью теоремы 5 вообще удается избавиться от отрицаний и вернуться к условиям теоремы 9.

Итак, в настоящем приложении описаны связи между такими объектами нечисловой природы, как нечеткие и случайные множества, установленные в нашей стране в первой половине 1970-х годов. Через несколько лет, а именно, в начале 1980-х годов, близкие подходы стали развиваться и за рубежом. Одна из работ [6] носит примечательное название "Нечеткие множества как классы эквивалентности случайных множеств".

В эконометрике разработан ряд методов статистического анализа нечетких данных, в том числе методы классификации, регрессии, проверки гипотез о совпадении функций принадлежности по опытным данным и т.д., при этом оказались полезными общие подходы статистики объектов нечисловой природы (см. главу 8 и работы [1,2,5]). Методологические и прикладные вопросы теории нечеткости мы обсуждали в работах [1,2,7].

Цитированная литература

1. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука,1979.- 296 с.
2. Орлов А.И. Задачи оптимизации и нечеткие переменные. - М.: Знание, 1980. - 64 с.
3. Лебег А. Об измерении величин. - М.: Учпедгиз, 1960. - 204 с.
4. Ефимов Н.В. Высшая геометрия. - М.: ГИФМЛ, 1961. - 580 с.
5. Орлов А.И. Основания теории нечетких множеств (обобщение аппарата Заде). Случайные толерантности. – В сб.: Алгоритмы многомерного статистического анализа и их применения. - М.: Изд-во ЦЭМИ АН СССР, 1975. - С.169-175.
6. Goodman I.R. Fuzzy sets as eguivalence classes of random sets // Fuzzy Set and Possibility Theory: Recent Developments. - New York-Oxford-Toronto-Sydney-Paris-Frankfurt, Pergamon Press, 1982. - P.327-343. (Перевод: Гудмэн И. Нечеткие множества как классы эквивалентности случайных множеств. - В сб.: Нечеткие множества и теория возможностей. Последние достижения. - М.: Радио и связь, 1986. - С. 241-264.)
7. Орлов А.И. Математика нечеткости. - Наука и жизнь. 1982. No.7. С.60-67.

Приложение 3

Методика сравнительного анализа родственных эконометрических моделей

В методике введено понятие родственных эконометрических моделей. Выделены теоретические и эмпирические единичные показатели качества эконометрических моделей с целью сравнения родственных моделей. Рассмотрены методы получения ранжировок родственных математических моделей по тем или иным показателям их качества и указаны методы согласования таких ранжировок. Рассмотрены методы проверки согласованности, кластеризации и усреднения ранжировок. Разобран пример сравнения родственных математических моделей на основе эмпирических единичных показателей качества. Приведены математические основы методов согласования ранжировок и классификаций, включая соответствующие теоремы с доказательствами. Дан обзор теоретических основ методов проверки согласованности, кластеризации и усреднения ранжировок.

П3-1. Общие положения

1.1. Методика имеет целью:

- по единой схеме оценивать качество эконометрических моделей;
- проводить сравнение однотипных эконометрических моделей;
- осуществлять выбор эконометрических моделей среди однотипных с целью практического использования или углубленной доработки.

1.2. Методика основана на выделении теоретических и эмпирических единичных показателей качества эконометрической модели, построении на их основе групповых и обобщенных показателей качества, их согласования и использовании для решения задач, указанных в п.1.1

1.3. Методика предусматривает использование как методов, основанных на анализе результатов наблюдений или специально поставленных экспериментов, так и методов, использующих экспертные оценки специалистов.

П3-2. Родственные эконометрические модели

2.1. Под эконометрической моделью в настоящей методике понимается функция, отображающая набор входных переменных в набор выходных переменных. Входные и выходные переменные могут иметь как числовую, так и нечисловую природу, быть измеренными в различных шкалах, сами быть функциями. Модель может задаваться уравнением, системой уравнений, алгоритмом, таблицей, графиком, словесно (при использовании нечисловых переменных).

2.2. Входные переменные делятся на:

- экономические переменные (цены, стоимости, предпочтения потребителей и т.п.);
- переменные управления (менеджмента);
- социальные переменные (состав персонала, демографический профиль);
- технические переменные, описывающие технологические процессы и характеристики изделий;
- физические переменные (пространственные координаты, время, температура,..., например, высота источника загрязнения, диаметр пролива, скорость ветра), описывающие условия воздействия;
- химические переменные, описывающие свойства воздействующих химических веществ;
- переменные, описывающие объекты воздействия (реципиентов);

- переменные, описывающие окружающую среду;

- эмпирические и иные константы, рассчитанные авторами моделей по результатам экспериментов или теоретически,

- иные виды переменных.

Входные переменные, например, цены или скорость ветра, могут зависеть от времени.

2.3. Под выходными переменными понимают те, которые используются при формировании окончательных суждений об объектах и процессах или являются входными в смежных моделях. Они так же, как входные переменные, могут быть функциями.

2.4. Эконометрические модели называются родственными по выходу, если наборы их выходных переменных совпадают. Модели называются частично родственными по выходу, если наборы их выходных переменных частично совпадают. Эконометрические модели называются родственными, если наборы их входных и выходных переменных совпадают. Частично родственные по выходу модели становятся родственным по выходу, если отказываемся от рассмотрения всех выходных переменных, кроме совпадающих. При формальном расширении множества входных переменных путем объединения таковых для нескольких родственных по выходу моделей получаем родственные модели.

Поэтому без ограничения общности можно считать, что на множестве эконометрических моделей задано отношение толерантности "быть родственными моделями". Пара моделей входит в это отношение тогда и только тогда, когда пересечение множеств их выходных переменных не пусто. В таком случае их можно преобразовать в пару родственных моделей (в смысле определения, данного в предыдущем абзаце).

2.5. Для сравнения родственных моделей используются как объективные (теоретические и экспериментальные) методы, так и субъективные экспертные оценки. Для сравнения и оценки моделей строится иерархическая система показателей качества: на основе единичных показателей формируются групповые, а с их помощью - обобщенные. Теоретические единичные показатели качества рассматриваются в пункте П3-3. Основанные на обработке результатов экспериментов единичные показатели качества рассматриваются в П3-4.

2.6. Результаты оценивания показателей качества, полученные экспертным или объективным путем, в литературе выражаются различными способами. Согласно методологии настоящей методики их рекомендуется выражать, как правило, в виде ранжировок или упорядоченных классификаций (нестрогих линейных порядков). Допускается использование метода парных сравнений, нечетких и интервальных оценок или иного способа получения информации от экспертов (главу 12 выше).

2.7. Для получения агрегированных (групповых или обобщенных) оценок применяют согласование ранжировок. При этом выявляются противоречия между различными ранжировками и привлекается дополнительная информация для упорядочения пар объектов, являющихся противоречивыми (которые по-разному упорядочены в исходных ранжировках). Методы согласования ранжировок рассмотрены в пункте П3-5.

2.8. При невозможности согласования ранжировок (из-за большого числа противоречий) проводится усреднение ранжировок в соответствии с правила расчета эмпирических средних в статистике объектов нечисловой природы (см. главу 8). При этом предварительно проводится проверка согласованности экспертов или методов оценки и при необходимости - разбиение их на кластеры, т.е. группы экспертов, имеющих сходные между собой мнения, резко отличные от мнений иных групп. Итоговые ранжировки рассчитываются отдельно для каждого кластера и передаются лицу, принимающему решения. Методы проверки согласованности экспертов и - в случае необходимости - разбиения их на группы-кластеры, а также усреднения ранжировок рассматриваются в пункте П3-6.

П3-3. Теоретические единичные показатели качества

3.1. К теоретическим единичным показателям качества эконометрической модели относятся показатели, не связанные с непосредственным использованием при оценивании и сравнении моделей данных реальных наблюдений, а именно, группы показателей:

- адекватности (обоснованности);
- внутренней согласованности;
- устойчивости;
- полноты;
- эффективности использования.

3.2. К показателям адекватности (обоснованности) относятся показатели:

- соответствия модели экономической (управленческой, технической, физической, химической, биологической, экологической и др.) сути явления, ее связи с научными результатами соответствующих областей;
- наличия надежной информации об экспериментальном подтверждении модели, о точности и практическом применении сделанных с ее помощью выводов и прогнозов.

3.3. К показателям внутренней согласованности относятся показатели:

- логической непротиворечивости модели;
- соблюдения различных законов сохранения (в частности, балансовых соотношений);
- соблюдения соотношений теории размерностей;
- наличия или отсутствия произвольных (не обоснованных научными методами) предположений;
- выполнения естественных граничных и предельных соотношений.

3.4. К показателям устойчивости относятся показатели:

- непрерывности расчетных величин эконометрической модели в случае непрерывности входных переменных;
- устойчивости выводов относительно допустимых преобразований шкал измерения;
- малой вариабельности выходных переменных при незначительных изменениях входных переменных (включая эмпирические константы);
- устойчивости к аномальным значениям отдельных экспериментальных результатов (в случае, если применение модели предполагает проведение экспериментов);
- устойчивости к вычислительным погрешностям.

3.5. К показателям полноты относятся показатели, показывающие:

- насколько исчерпывающе набор выходных переменных отражает потребности конечного пользователя;
- насколько исчерпывающе набор выходных переменных отражает потребности смежных моделей;
- насколько исчерпывающе отражены в модели главные черты описываемого объекта (с точки зрения имеющейся в объектной области теории).

3.6. К показателям эффективности использования относятся показатели:

- простоты и наглядности модели;
- удобства пользования моделью;
- доступности информации, необходимой для применения модели;
- стоимости получения информации, необходимой для применения модели.

3.7. Перечень теоретических единичных показателей качества родственных моделей может быть дополнен в соответствии со спецификой моделируемого явления или процесса.

3.8. Оценивание теоретических единичных показателей качества родственных моделей проводится экспертным путем. Допускается использование одного или нескольких экспертов - в случае доступности исходной информации и отсутствии причин для возможного расхождения мнений экспертов. В противном случае необходимо проведение экспертного опроса в полном объеме.

ПЗ 4. Эмпирические единичные показатели качества

4.1. Эмпирические единичные показатели качества применяются, когда имеются надежные результаты экспериментов, позволяющие сравнивать результаты измерений, наблюдений, анализов, проб с расчетными значениями, полученными на основе родственных моделей.

4.2. Для проведения такого сравнения на основе того или иного эмпирического единичного показателя качества могут быть использованы различные методы эконометрики, прикладной статистики и планирования экспериментов.

4.3. В случае одной числовой выходной переменной в соответствии с принятым в настоящей методике подходом используют два основных метода ранжировки родственных математических моделей (использующих два основных эмпирических единичных показателя качества):

- по сумме относительных отклонений результатов измерений от расчетных значений (п.4.4);

- по сумме рангов, присвоенных моделям в каждой экспериментальной точке в зависимости от близости к измеренным значениям (п.4.5);

а также дополнительные - по числу экспериментальных точек, в которых модель оказалась наилучшей (п.4.6).

4.4. При использовании метода ранжировки по сумме относительных отклонений для каждой модели подсчитывают относительную погрешность в каждой экспериментальной точке (относительно наблюденного значения), складывают эти относительные погрешности (без учета знаков) и ранжируют модели в порядке возрастания указанных сумм.

4.5. При использовании метода ранжировки по сумме рангов в каждой экспериментальной точке строят ранжировку (упорядочение) моделей по степени близости соответствующих расчетных значений к результату наблюдений (без учета знака отклонения) - самая точная модель получает ранг 1, вторая по точности - ранг 2, и т.д. Затем ранги складываются по всем экспериментальным точкам и модели ранжируются в порядке возрастания суммы рангов.

4.6. Вариантом метода ранжировки по числу экспериментальных точек, в которых модель оказалась наилучшей (без учета знака отклонения), является метод разбиения рассматриваемой совокупности родственных моделей на два класса - тех, которые оказались наилучшими хотя бы для одной экспериментальной точки (т.е. оптимальных по Парето), и остальных, никогда не бывших наилучшими.

Другой вариант предполагается учет числа точек, в которых та или иная из рассматриваемой совокупности родственных моделей оказалась наилучшей (наиболее точной). Чем в большем числе точек модель оказалась точнее, тем выше она оценивается. Число классов в этом варианте не фиксировано, оно может меняться от 1 (когда каждая модель оптимальна ровно в одной точке) до общего числа экспериментальных точек.

4.7. При наличии достаточной эмпирической информации целесообразно построить формально-статистические модели, например, линейного регрессионного анализа, примененного к входным переменным или функциям от них (например, логарифмам, если исходные модели описываются степенными функциями), рассчитать оценки параметров по экспериментальным данным (т.е. идентифицировать модели), добавить вновь построенные модели в перечень родственных моделей и сравнить их с остальными согласно настоящей методике.

4.8. Перечень эмпирических единичных показателей качества родственных математических моделей может быть дополнен в соответствии со спецификой моделируемого явления или процесса. В частности, могут быть использованы такие показатели, как:

- сумма квадратов отклонений расчетных значений от результатов измерений, наблюдений, анализов, проб;

- сумма модулей таких отклонений;

- сумма квадратов относительных погрешностей;

- величина максимально возможного отклонения (с учетом или без учета знака);

- расстояние (показатель близости) иного вида между вектором экспериментальных значений и вектором расчетных значений, соответствующих определенной модели, например, расстояние Махаланобиса при той или иной корреляционной (или весовой) матрице;

- в случае нескольких выходных параметров - различные характеристики качества планов эксперимента, разработанные в теории планирования эксперимента (в соответствии с перечнем, приведенным в монографии [1], глава II).

4.9. Пример сравнения (ранжировки) родственных математических моделей на основе эмпирических единичных показателей качества дан в пункте П3-7 ниже.

П3-5. Методы согласования ранжировок

5.1. Методы раздела 5 применяются в соответствии с п.2.6 для согласования ранжировок родственных моделей, полученных с помощью теоретических или эмпирических единичных показателей качества моделей, а также ранжировок, построенных на основе групповых показателей качества, в частности, показателей, рассмотренных в пп.3.2 - 3.6.

5.2. При согласовании ранжировок исходят из двух или нескольких ранжировок, вообще говоря, со связями (нестрогих линейных порядков). В каждой ранжировке модели располагаются в порядке понижения качества, причем некоторые модели могут признаваться эквивалентными (по рассматриваемому показателю качества). Ранжировки моделей могут быть получены как при их объективном сравнении по различным показателям качества, так и от экспертов, сравнивающих модели по тому или иному показателю или их набору.

5.3. На первом этапе согласования ранжировок выделяются противоречивые пары моделей. Пара моделей А и В признается противоречивой, если в одной из рассматриваемых ранжировок модель А строго лучше модели В, а в какой-то другой модель В строго лучше модели А. Тем самым определяется симметричное бинарное отношение (квазитолерантность) на множестве моделей.

5.4. В соответствии с правилами теории бинарных отношений проводится транзитивное замыкание квазитолерантности, построенной в соответствии с п. 5.3. Устанавливается порядок между классами эквивалентности, соответствующий порядкам во всех исходных ранжировках. Полученная ранжировка называется согласующей для множества исходных ранжировок. При сравнении согласующей ранжировки с любой из исходных не существует ни одной противоречивой пары (это вытекает из теорем, приведенных в пункте П3-8 ниже).

5.5. При необходимости упорядочения по качеству моделей, входящих в один класс согласующей ранжировки (т.е. эквивалентных в соответствии с ней), привлекается дополнительная информация. Эта информация может опираться на дополнительные показатели качества, на результаты дополнительных исследований, как экспериментальных, так и теоретических, в частности, проведенных методами статистики объектов нечисловой природы (см. главу 8 выше).

П3-6. Методы проверки согласованности, кластеризации и усреднения ранжировок

6.1. При необходимости упорядочения по качеству моделей, входящих в один класс согласующей ранжировки (п.5.5), применяют методы проверки (статистической) согласованности, при необходимости - кластерного анализа, а затем - усреднения ранжировок, разработанные в статистике объектов нечисловой природы.

6.2. Методы, указанные в п.6.1, предполагают использование того или иного расстояния (меры различия) в пространстве ранжировок (со связями). В соответствии с методологией настоящей методики используется расстояние Кемени-Снелла, связанное с коэффициентом ранговой корреляции Кендалла, при проверке (статистической) согласованности и - при необходимости - проведении кластерного анализа. При усреднении ранжировок используется мера различия, основанная на коэффициенте ранговой корреляции Спирмена. Допускается использование иных расстояний и мер близости (различия) в том числе:

- расстояния, основанного на понятии ближайшего соседа;
- иных расстояний и мер близости, разработанных в статистике объектов нечисловой природы.

6.3. При использовании одновременно нескольких расстояний (мер различия или близости) в пространстве ранжировок (со связями) в соответствии с методологией настоящей методики необходимо использовать выводы, устойчивые относительно выбора того или иного расстояния (меры различия) в пространстве ранжировок (со связями).

6.4. В соответствии с методологией настоящей методики сначала проверяется согласованность набора ранжировок с помощью коэффициента ранговой конкордации Кендалла и Бебингтона Смита (при небольшом числе связей) или теории люсианов (если ранжировки построены на основе парных сравнений моделей), а также согласованность оценивается экспертно.

6.5. В случае недостаточной согласованности набора ранжировок проводится их разбиение на группы схожих между собой тем или иным методом кластерного анализа. Результат разбиения должен быть достаточно устойчив относительно выбора метода кластер-анализа. Деление показателей качества на группы, по которым модели оцениваются схожим образом, или экспертов на группы с близкими мнениями используется неформально при дальнейшем сравнении родственных математических моделей.

6.6. При положительном ответе на вопрос о согласованности ранжировок результирующая (итоговая) ранжировка находится как эмпирическое среднее (медиана Кемени) согласно статистике объектов нечисловой природы (с учетом сказанного в пп. 6.2, 6.3). При отрицательном ответе на вопрос о согласованности ранжировок результирующие (итоговые) ранжировки находятся отдельно для каждого кластера.

6.7. Информация о расчетных формулах по методам раздела 6 и их теоретических основах приведены в пункте П3-9 ниже.

П3-7. Пример сравнения родственных эконометрических моделей на основе эмпирических единичных показателей качества

При решении задач экологического страхования необходимо проанализировать последствия возможных аварий на химических производствах. Другими словами, в экологическом страховании экономические проблемы переплетаются с проблемами химической безопасности биосферы. Поэтому нет ничего удивительного в том, что в качестве примера рассматриваются 8 родственных эконометрических (если угодно - математических) моделей стационарных процессов испарения жидкости с открытых поверхностей. Модели будем различать по фамилиям предложивших и изучавших их специалистов. Это модели Лебузера (в дальнейшем кратко Л), Мак-Кея (М-К), Гусева-Баранаева (Г-Б), Клячко (К), Стефана (Стеф), Братсерта (Б), Дикона (Д), Соломона (Сол). Имеются данные о 12 конкретных экспериментах. Для соответствующих 12 наборов входных переменных получены расчетные значения по упомянутым 8 моделям. В табл.1 приведены значения относительных погрешностей (в процентах и без учета знака) расчетных значений относительно реальных.

Табл.1. Относительные погрешности (в %)
для 8 родственных моделей

№ эксп.	Д	Л	М-К	Б	Г-Б	Сол	Стеф	К

1	44,3	17,0	7,6	11,2	74,8	20,7	48,8	64,5
2	36,4	15,3	6,9	0,4	103,1	4,3	42,9	52,3
3	18,0	48,6	37,7	29,4	161,7	22,1	26,4	39,1
4	38,9	14,4	7,8	7,8	109,3	4,9	50.8	18,0
5	61,7	28,3	32,4	42,1	31,3	41,1	9,0	49,3
6	27,8	30,5	19,2	15,4	12,8	9,5	33,1	51,7
7	52,1	11,6	18,7	25,5	44,6	27,4	58,3	55,5
8	43,0	0,1	9,4	6,2	70,5	18,9	43,5	75,8
9	51,6	11,4	21,8	21,3	48,0	30,9	53,3	75,8
10	39,5	5,1	2,9	1,1	78,9	14,4	40,9	74,7
11	49,2	11,9	20,2	15,6	48,9	29,1	48,3	81,7
12	8,5	106,8	95,9	59,2	268,5	85,4	17,9	129,8
Сумма	471	301	280,5	235,2	1072,4	308,7	533,2	768,2

В последней строке табл.1 в соответствии с п.4.4 методики приведены суммы относительных отклонений результатов измерений от расчетных значений. Упорядочение (ранжировка) по сумме относительных погрешностей (отклонений) имеет вид:

Б < М-К < Л < Сол < Д < Стеф < К < Г-Б . (1)

В табл.2 приведены ранги 8 моделей по точности приближения в отдельных экспериментальных точках (ранг 1 - самая точная модель, ранг 2 - вторая по точности,..., ранг 8 - самая далекая от истинного экспериментального значения модель). Они получены путем сравнения относительных погрешностей из табл.1.

Табл.2. Ранги 8 моделей по точности приближения

№ эксп.	Д	Л	М-К	Б	Г-Б	Сол	Стеф	К
1	5	3	1	2	8	4	6	7
2	5	4	3	1	8	2	6	7
3	1	7	5	4	8	2	3	6
4	6	4	2,5	2,5	8	1	7	5
5	7	1	3	5	2	4	8	6
6	5	6	4	3	2	1	7	8
7	6	1	2	3	5	4	8	7
8	5	1	3	2	7	4	6	8
9	6	1	3	2	5	4	7	8
10	5	3	2	1	8	4	6	7
11	7	1	2	2	6	4	5	8
12	1	6	5	3	8	4	2	7
Сумма	59	38	36,5	30.5	75	38	71	84
Итоговый ранг	5	3,5	2	1	7	3,5	6	8

В соответствии с п.4.5 ранги складываются по всем экспериментальным точкам (суммы приведены в предпоследней строке табл.2) и модели ранжируются в порядке возрастания суммы рангов. Итоговый ранг приведен в последней строке табл.2. Ранжировка по суммам рангов (или, что то же, по средним арифметическим рангов) имеет вид:

Б < М-К < {Л, Сол} < Д < Стеф < Г-Б < К . (2)

Поскольку модели Л и Сол получили одинаковую сумму баллов, то по этому показателю они эквивалентны, а потому объединены в группу (кластер), т.е. ранжировка (2) имеет одну связь.

Сравнивая ранжировки (1) и (2), видим, что они весьма похожи. Они отличаются только по двум позициям:

- стоящие рядом в ранжировке (1) модели Л и Сол в ранжировке (2) объединены в один кластер;
- модели К и Г-Б расположены в ранжировках (1) и (2) в противоположном порядке.

В соответствии с п.5.3. на первом этапе согласования ранжировок следует выделить противоречивые пары моделей. При сравнении ранжировок (1) и (2) только пара моделей К и Г-Б признается противоречивой. Следовательно, для ранжировок (1) и (2) согласующей является кластеризованная ранжировка

Б < М-К < Л < Сол < Д < Стеф < {К, Г-Б}, (3)

в которой модели упорядочены от лучшей к худшей.

Рассмотрим теперь дополнительные методы ранжирования, предусмотренные п. 4.6 настоящей методики. Вариантом метода ранжировки по числу экспериментальных точек, в которых модель оказалась наилучшей (без учета знака отклонения), является метод разбиения рассматриваемой совокупности родственных моделей на два класса - тех, которые оказались наилучшими хотя бы для одной экспериментальной точки (т.е. оптимальных по Парето), и остальных, никогда не бывших наилучшими. В первое множество входят модели Б, М-К, Л, Сол, Д, являющиеся оптимальными по Парето на рассматриваемом множестве экспериментальных точек, во второе - остальные модели, т.е. Стеф, К, Г-Б, и соответствующая ранжировка со связями имеет вид

{Б, М-К, Л, Сол, Д} < {Стеф, К, Г-Б} . (4)

Ранжировка (4) не имеет противоречивых пар с ранжировкой (3), поэтому можно считать, что ранжировка (3) является согласующей для всех трех ранжировок (1), (2), (4).

Другой вариант, предусмотренный п.4.6, предполагается учет числа точек, в которых та или иная из рассматриваемой совокупности родственных моделей оказалась наилучшей (наиболее точной). Чем в большем числе точек модель оказалась точнее, тем выше она оценивается. Модель Л является наилучшей в 5 экспериментах (№№ 5, 7, 8, 9, 11), модель Б - в 2 экспериментах (№№ 2, 10), как и модели Д (эксперименты №№ 3, 12) и Сол (эксперименты №№ 4, 6), модель М-К - в одном (№ 1), остальные - ни разу. Ранжировка имеет вид:

Л < {Б, Д, Сол} < М-К < {Стеф, К, Г-Б} . (5)

Сопоставим ранжировки (3) и (5). Имеем следующие четыре противоречивые пары: Л и Б, Л и М-К, Д и М-К, Сол и М-К. Значит, в один кластер с М-К надо включить Л, Д и Сол, а раз модель Л связана противоречием в Б, то и Б надо включить в этот кластер, состоящий в итоге из 5 моделей - Л, Б, Д, Сол, М-К. Итоговая ранжировка имеет вид:

{Л, Б, Д, Сол, М-К }< Стеф < {К, Г-Б} . (6)

Она является согласующей для четырех ранжировок (1), (2), (4), (5). (Напомним, что кластер {К, Г-Б} появился как следствие противоречия в упорядочении моделей К и Г-Б в ранжировках (1) и (2).)

Выше приведены результаты формального анализа семейства 8 родственных моделей по 4 критериям. Общее заключение должно быть сделано экспертным путем.

В данной ситуации по мнению экспертов итогом сравнения моделей должна быть признана ранжировка (3), являющаяся согласующей для 3 из 4 критериев:

Б < М-К < Л < Сол < Д < Стеф < {К, Г-Б}.

Ранжировка (6), согласующая для всех четырех критериев, объявляет эквивалентными 5 наиболее интересных моделей, поскольку оставшиеся 3 модели по результатам анализа экспериментальных данных можно вообще исключить из дальнейшего рассмотрения. Согласно п. 5.5 в случае необходимости упорядочения моделей, попавших в один кластер, привлекается дополнительная информация. В рассматриваемом случае дополнительная информация дает основания исключить один из четырех критериев.

В главе 12 процедура согласования ранжировок использовалась при анализе мнений экспертов. Однако в настоящем приложении 3 речь идет не о мнениях экспертов, а о сравнении эконометрических моделей. Исходные данные - табл.1 - результаты измерений, а не субъективные оценки.

П3-8. Математические основы методов согласования ранжировок и классификаций

При использовании нескольких обобщенных показателей получаются, как правило, различающиеся ранжировки объектов. Как их согласовать с целью дальнейшего использования при классификации? В настоящем пункте формулируются и обосновываются методы решения этой задачи. В отличие от главы 12 дается строгое математическое изложение с доказательствами основных утверждений.

Взвешенные агрегированные показатели. *Пусть* $X_1, X_2, \dots, X_K$ - частные (или групповые) числовые показатели. Пусть каждому из них приписан вес - $A_1, A_2, \dots, A_K$ соответственно, отражающий их относительную важность (оцененную экспертами или иным способом). Весовые коэффициенты неотрицательны и в сумме составляют 1.

Взвешенные агрегированные показатели можно определить следующим единообразным способом.

Введем (чисто формально) распределение вероятностей, приписывающее каждому значению X_M, M=1,2,...,K, вероятность A_M. Для этого распределения обычным образом определим такие характеристики, как математическое ожидание, медиана, начальные моменты, мода и т.д., которые и будем использовать в качестве взвешенных агрегированных показателей или при их расчете.

При этом математическое ожидание дает взвешенное среднее арифметическое, медиана - взвешенную медиану (в частном случае, когда одна из ступенек функции распределения приходится на высоту 0,5, целесообразно ввести понятия левой и правой медиан - т.е. левого и правого концов указанной ступеньки соответственно).

Начальный момент p-го порядка после извлечения корня p-ой степени дает взвешенное степенное. Аналогичным образом получаем обобщенное среднее по Колмогорову общего вида.

Мода указывает на значение наиболее важного показателя.

В соответствии с методологией устойчивости (см. главу 10 выше) при анализе конкретной ситуации целесообразно одновременно использовать несколько обобщенных показателей, например, взвешенную медиане и взвешенное среднее арифметическое. Такая процедура предусмотрена в настоящей методике. Хотя согласно теории измерений (см. главу 3 выше) использование среднего арифметического некорректно, но приходится учитывать традиции (проблема учета традиций подробно обсуждалась в главе 12).

Согласование упорядочений по агрегированным показателям. Сопоставим упорядочения объектов по двум видам агрегированных оценок, например, по взвешенной медиане и по взвешенному среднему арифметическому. Для этого построим "квазитолерантность расхождений (КТР)", т.е. некоторое бинарное отношение (о теории бинарных отношений см., например, книгу [2]) на множестве объектов. (Как известно, бинарное отношение на данном множестве объектов можно отождествить с подмножеством множества пар объектов, т.е. с подмножеством декартова квадрата исходного множества объектов.)

По определению два объекта связаны отношением КТР (т.е. пара объектов входит в рассматриваемое подмножество) тогда и только тогда, когда два упорядочения - по взвешенной медиане и по взвешенному среднему арифметическому - для них противоречивы. Это возможно в двух случаях. Первый - средний взвешенный арифметический показателей для первого (из двух рассматриваемых) объектов больше (или равен) такового для второго объекта, а взвешенная медиана для первого, наоборот, меньше, чем для второго. Второй - средний взвешенный арифметический показателей для первого (из двух рассматриваемых) объектов меньше такового для второго вида, а взвешенная медиана для первого, наоборот, больше (или равна), чем для второго.

Отношение КТР является симметричным (если пара (A,B) входит в него, то входит и пара (B,A)) и антирефлексивным (ни одна пара (A,A) не входит в КТР). Свойством транзитивности это бинарное отношение, вообще говоря, не обладает (если пары (A,B) и (B,C) входят в него, то пара (A,C) может входить в КТР, а может и не входить).

Формально присоединим к КТР все пары вида (A,A). Получим рефлексивное симметричное отношение, т.е. толерантность (о толерантностях много написано в монографии [2]). Будем называть ее "толерантностью расхождений (ТР)".

Построим новое бинарное отношение Зам(ТР) путем транзитивного замыкания (в смысле теории бинарных отношений, см., например, монографию [2,с.27]) "толерантности расхождений". Это означает, что подмножество пар объектов, входящих в толерантность ТР, пополняется некоторыми новыми парами. А именно, если A, B и C - три объекта такие, что пара (A,B) и пара (B,C) входят в "толерантность расхождений", то пару (A,C) включаем в замыкание этой толерантности. Для полученного множества пар повторяем описанную операцию. Продолжаем так до тех пор, пока новые пары не перестанут добавляться (процесс не может продолжаться бесконечно, поскольку общее число пар конечно).

Бинарное отношение Зам(ТР) можно описать и по-другому: пара (A,B) входит в Зам(ТР) тогда и только тогда, когда либо она входит в ТР, либо существует конечная последовательность объектов C, D, E, ..., Q такая, что пары (A,C), (C,D), (D,E), ..., (Q,B) входят в ТР, т.е. от A к B можно пройти за несколько шагов, каждый из которых - переход от первого элемента пары, входящей в ТР, ко второму.

Последнее замечание подсказывает наглядную геометрическую интерпретацию операции замыкания. Представим себе объекты точками на плоскости. Пара (A,B) входит в ТР тогда и только тогда, когда от A до B можно добраться по дороге. Тогда ясно, что пара (A,C) входит в Зам(ТР) в том и только в том случае, когда от A до C можно добраться по дороге, возможно. через несколько промежуточных пунктов (объектов).

Теорема о структуре замыкания. Описание структуры Зам(ТР) дает следующая теорема.

Теорема 1. *Замыкание "толерантности расхождений" - отношение эквивалентности (рефлексивное симметричное транзитивное отношение), задающее разбиение объектов на кластеры (группы эквивалентных в рассматриваемом смысле объектов). Кластеры между собой упорядочены: все объекты одного кластера одновременно лучше (или одновременно хуже) всех объектов другого кластера одновременно по обоим используемым агрегированным показателям. Внутри же кластеров, состоящих более чем из одного элемента, имеются противоречия: для какого-то объекта есть другой из того же кластера такой, что упорядочение по одному агрегированному показателю противоречит упорядочению по другому агрегированному показателю.*

Доказательство. Рефлексивность Зам(ТР) вытекает из рефлексивности ТР - поскольку любая пара (A,A) входит в ТР, то она входит и в Зам(Т,Р). Симметричность вытекает из симметричности ТР: если из A в B можно добраться по цепочке C, D, E, ..., Q, то из B в A - по обратной цепочке Q,...,E, D,C, каждые два соседних элемента которой образуют пару, входящую в ТР наряду с "симметричной" парой из прямой цепочки. Транзитивность вытекает из процедуры построения Зам(ТР). В теории бинарных отношений рефлексивное симметричное и транзитивное отношение, как известно, называется эквивалентностью (см., например, [2, с.54]).

Хорошо известно (см., например, теорему 2.1 в монографии [2, с.55-56]), что отношение эквивалентности задает разбиение множества объектов на кластеры (классы, группы, подмножества) такое, что пара (A,B) входит в Зам(ТР) тогда и только тогда, когда объекты A и B включены в один и тот же кластер.

Теперь введем упорядоченность кластеров.

Лемма. Пусть X = {A, B,...} и Y = {C,D,...} - два кластера. Пусть A меньше C при использовании одного из двух рассматриваемых видов агрегированных оценок (например, по взвешенной медиане или по взвешенному среднему арифметическому). Тогда A меньше

С и при сравнении по второй агрегированной оценке. Более того, любой объект из первого кластера меньше любого объекта из второго кластера в смысле любой из двух агрегированных оценок.

Докажем лемму. Если бы А было больше или равно С по второй оценке, то пара (А,С) входила бы в КТР и ТР, а потому объекты А и С входили бы в один класс разбиения, соответствующего Зам(ТР), что противоречит исходному предположению. Это рассуждение показывает также, что для любых двух объектов В и D из разных кластеров упорядоченности по двум агрегированным оценкам совпадают.

Однако совпадает ли упорядоченность В и D (или даже В и С) с упорядоченностью А и С?

Одну из упорядоченностей обозначим знаком < (т.е. "меньше"; знак > означает здесь "больше или равно"). Может ли быть так, что А<С, но В>С ? Тогда А<С<В. Вторую упорядоченность обозначим знаком //. Тогда в соответствии с рассуждениями предыдущего абзаца А//С//В, следовательно, пара (А,В) не может входить в КТР, а потому и в ТР.

Поскольку А и В лежат в одном кластере, то существует цепочка А(1)=А, А(2), А(3), ..., А(К) = В такая, что пары (А(Р), А(Р+1)) входят в КТР, Р = 1, 2, 3,..., К-1. Рассмотрим минимальное М такое, что А(М)<С, А(М+1)>С (такое М существует, поскольку А1<С, а АК>С). Тогда в рассуждениях предыдущего абзаца можно положить А=А(М), В=А(М+1). Получаем, что пара (А(М), А(М+1)) не входит в КТР, что противоречит определению Зам(ТР).

Итак, доказано, что из А<С вытекает В<С для любого В из кластера, включающего А. Аналогичным образом устанавливается, что В<D для любого D из кластера, включающего С. Лемма доказана.

Каждый из кластеров, порожденных Зам(ТР), может состоять из одного или нескольких элементов. Внутри кластера из одного элемента противоречий быть не может. Если в кластере несколько элементов, то хотя бы одна пара объектов из этого кластера входит в КТР. Однако некоторые пары могут и не содержать противоречий. Например, если упорядочения имеют вид А<В<С и С//А//В, то пары (В,С) и (А,С) входят в КТР, а пара (А,В) - нет. Если же второе упорядочение имеет вид С//В//А, то все три пары входят в квазитолерантность расхождений.

Теорема 1 доказана.

Развитие методики агрегирования. В результате описанной выше процедуры получаем ранжировку (упорядоченный ряд), элементами которой являются, вообще говоря, не отдельные объекты, а кластеры, состоящие из некоторого числа объектов (некоторые из кластеров могут состоять из одиночных объектов, для которых не оказалось рассматриваемых выше противоречий). Если построенное согласно описанной процедуре разбиение объектов на кластеры и полученный на его основе ранжировочный ряд удовлетворяет заказчика, то они и определяют итоговую ранжировку и итоговый агрегированный показатель (выражающийся, например, в номере кластера, в который входит рассматриваемый объект, в итоговой ранжировке). Если же нет (например, получился всего один класс), то требуется дополнительный анализ с привлечением экспертов. Он должен быть нацелен на уточнение предпочтений экспертов. Например, им могут быть предъявлены для сравнения пары объектов, входящих в "квазитолерантность расхождений". Это исследование может описаться на различные методики выявления предпочтений (в экономических терминах - функций полезности).

По ранжировке строится классификация путем разбиения области значений итогового агрегированного показателя на упорядоченные зоны. Границы между зонами задаются с помощью опроса экспертов с учетом процедуры дальнейшего использования этих зон.

Заметим, что описанная выше методика может применяться в различных вариантах. В облегченном варианте весовые коэффициенты не оцениваются. Например, они априори предполагаются равными или же задаются исследовательской группой, строящей агрегированный показатель.

В соответствии с общей схемой устойчивости (глава 10) целесообразно численно изучить устойчивость значений агрегированного показателя к малым отклонениям значений весовых коэффициентов, а также ответов экспертов. Развитие этой идеи ведет к разработке методики численного эксперимента, а также к применению и изучению интервальных экспертных оценок, когда ответ эксперта - интервал действительных чисел или интервал в порядковой шкале (несколько соседних градаций), и т.д. (см. главы 11 и 12).

Могут быть использованы и иные виды средних величин, кроме среднего арифметического и медианы, в частности, среднее геометрическое и другие виды средних по Колмогорову.

О согласовании классификаций. Пусть имеются две классификации H1 и H2, разбивающие множества объектов на кластеры A1, A2,..., AK и B1, B2,..., BM соответственно. Рассмотрим новую классификацию H, построенную на основе пересечений множеств A1xB1, A2xB1,..., AKxB1, A1xB2, A2xB2,..., AKxB2,..., A1xBM, A2xBM,..., AKxBM (здесь x - знак пересечения). Число кластеров в H - не более KxM, поскольку некоторые из выписанных пересечений могут оказаться пустыми. Классификация H обладает тем свойством, что любые два элемента, входящие в один из ее кластеров, входят также в один кластер и в H1, и в H2. Если же два элемента входят в разные кластеры H, то либо в H1, либо в H2, либо одновременно и в H1, и в H2 они входят в разные кластеры. Поэтому можно сказать, что классификация H согласует классификации H1 и H2.

Для классификаций с неупорядоченными кластерами сказанное в предыдущем абзаце решает проблему согласования. Для классификаций, кластеры которых строго линейно (или совершенно) упорядочены [2, с.119-120], т.е. порожденных склейкой одинаковых значений некоторого агрегирующего показателя на множестве объектов (существование такого показателя вытекает из теоремы 4.2 в [2, с.121-122]), можно продвинуться дальше.

Описанная выше *процедура согласования классификаций,* полученных различными способами на основе двух ранжировок, *является общей.* Она может быть применена для согласования любых двух классификаций, использующих строго линейно упорядоченные кластеры.

Сначала необходимо построить "квазитолерантность расхождений (КТР)", включающую те и только те пары объектов, упорядоченность которых в двух классификациях различна. Затем строим "толерантность расхождений (ТР)", добавляя к КТР все пары вида (А,А). Затем строим Зам(ТР), транзитивно замыкая ТР по правилам теории бинарных отношений [2, с.27]. Корректность этой процедуры обеспечивает следующая теорема.

Теорема 2. *Замыкание толерантности расхождений Зам(ТР) задает классификацию на упорядоченные кластеры. При этом все объекты одного кластера одновременно лучше (или одновременно хуже) всех объектов другого кластера одновременно по обоим используемым агрегированным показателям. Внутри же кластеров, состоящих более чем из одного элемента, имеются противоречия: для какого-то объекта есть другой из того же кластера такой, что упорядочение по одному агрегированному показателю противоречит упорядочению по другому агрегированному показателю.*

Доказательство. Как показано при доказательстве теоремы 1, Зам(ТР) является отношением эквивалентности, а потому задает некоторое разбиение множества объектов, т.е. классификацию.

Просматривая доказательство теоремы 1, нетрудно заметить, что в нем не используются какие-либо конкретные свойства взвешенной медианы или взвешенного среднего арифметического, а потому проведенные рассуждения верны для любых строгих совершенных (линейных) порядков. Это замечание и заканчивает доказательство теоремы 2.

Замечание. Расчет согласующей классификации как Зам (ТР) не всегда дает приемлемые с практической точки зрения результаты. Пусть например, имеется 4 объекта, описываемые точками на плоскости A = (0,0), B = (0,1), C = (1, 0), H = (1,1), первое упорядочение - по первой координате, второе - по второй (каждое из упорядочений имеет два варианта соответственно тому, как интерпретировать равенство, т.е. использовать отношение "меньше" или "меньше или равно"). Нетрудно проверить, что Зам(ТР) дает вырожденную классификацию - состоит из одного кластера. Между тем другие способы построения результирующего упорядочения, например, по сумме координат, могут оказаться более практически приемлемы.

Практический интерес представляет также задача расширения классификации по упорядоченным классам, заданной на *части* естественного множества определения, на *все* это множество. Решений, как правило, имеется несколько, и возникают *проблемы описания всех возможных расширений и выбора* из них наиболее адекватного с точки зрения рассматриваемой прикладной области, например, токсикологии как части экологического страхования.

Об алгоритмах нахождения согласующей кластеризованной ранжировки. Пусть дана конечная совокупность ранжировок моделей (возможно, со связями). Требуется построить согласующую ранжировку, возможно, кластеризованную (т.е. со связями).

Шаг 1. Находим все пары моделей, упорядочение которых хотя бы в двух исходных ранжировках противоречиво (в одной ранжировке первая модель строго лучше второй, а в другой ранжировке - наоборот, вторая модель строго лучше первой).

Шаг 2. Рассмотрим граф, вершины которого - модели из рассматриваемого семейства родственных моделей. Две вершины соединены ребром тогда и только тогда, когда они выделены на шаге 1. Выделяем связные компоненты этого графа.

Шаг 3. Устанавливаем строгий порядок между связными компонентами графа, выделенными на шаге 2 (кластерами). Получаем искомую согласующую ранжировку.

Программная реализация описанной схемы может быть осуществлена различными способами.

П3-8. Теоретические основы методов проверки согласованности, кластеризации и усреднения ранжировок

Как указано в п.6.1 настоящей методики, при необходимости упорядочения по качеству моделей, входящих в один класс согласующей кластеризованной ранжировки, применяют методы проверки (статистической) согласованности, при необходимости - кластерного анализа, а затем - усреднения ранжировок, разработанные в статистике объектов нечисловой природы. Эти методы предполагают использование того или иного расстояния (меры различия) в пространстве ранжировок (со связями). В соответствии с методологией настоящей методики используется расстояние Кемени-Снелла (см. главу 8, а также монографию [3]), связанное с коэффициентом ранговой корреляции Кендалла (см. справочник [4]), при проверке (статистической) согласованности и - при необходимости - проведении кластерного анализа. При усреднения ранжировок часто используется мера различия, основанная на коэффициенте ранговой корреляции Спирмена (см. [4]). Допускается использование иных расстояний и мер близости (различия) в том числе:

- расстояния, основанного на понятии ближайшего соседа;
- иных расстояний и мер близости, разработанных в статистике объектов нечисловой природы (см. главу 8 и монографии [5-6]).

При использовании одновременно нескольких расстояний (мер различия или близости) в пространстве ранжировок (со связями) в соответствии с методологией теории устойчивости (глава 10) необходимо использовать выводы, устойчивые относительно выбора того или иного расстояния (меры различия) в пространстве ранжировок (со связями).

Сначала проверяется согласованность набора ранжировок с помощью коэффициента ранговой конкордации Кендалла и Бебингтона Смита (при небольшом числе связей) согласно [4, табл. 6.10]. Если ранжировки построены на основе парных сравнений моделей, то используются методы теории люсианов (см., например, [7,8]; пример алгоритмов из теории люсианов описан выше в главе 13). Согласованность экспертов может также оцениваться с помощью другой группы экспертов.

В случае недостаточной согласованности набора ранжировок, т.е. отклонения гипотезы согласованности на уровне значимости 5 % или более низком, проводится их разбиение на группы схожих между собой тем или иным методом кластерного анализа (см. главу 5). Согласно методологии устойчивости (глава 10) результат разбиения должен быть достаточно устойчив относительно выбора метода кластер-анализа. Рекомендуется одновременно использовать метод ближнего соседа и метод дальнего соседа, используя в дальнейшем устойчивые ядра кластеров, выделяющиеся при одновременном применении указанных двух методов.

Деление показателей качества на группы, по которым модели оцениваются схожим образом, или экспертов на группы с близкими мнениями используется участниками проекта и пользователями банка эконометрических моделей. Это деление учитывается также и неформально при дальнейшем применении или сравнении родственных эконометрических моделей.

При положительном ответе на вопрос о согласованности ранжировок результирующая (итоговая) ранжировка находится как эмпирическое среднее, т.е. медиана Кемени, согласно методам и алгоритмам статистики объектов нечисловой природы. При отрицательном ответе на вопрос о согласованности ранжировок результирующие (итоговые) ранжировки находятся отдельно для каждого кластера. При этом, например, констатируется принципиальное различие научных школ, к которым принадлежат эксперты.

Цитированная литература

1. Налимов В.В., Голикова Т.И. Логические основания планирования эксперимента. - М.: Металлургия, 1976. 128 с.

2. Шрейдер Ю.А. Равенство, сходство, порядок. - М.: Наука, 1971. - 254 с.

3. Кемени Дж., Снелл Дж. Кибернетическое моделирование: Некоторые приложения. - М.: Советское радио, 1972. - 192 с.

4. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. - М.: Наука, 1965 (1-е изд.), 1968 (2-е изд.), 1983 (3-е изд.).

5. Тюрин Ю.Н., Литвак Б.Г., Орлов А.И., Сатаров Г.А., Шмерлинг Д.С. Анализ нечисловой информации. - М.: Научный Совет АН СССР по комплексной проблеме "Кибернетика", 1981. - 80 с.

6. Анализ нечисловой информации в социологических исследованиях. - М.: Наука, 1985. - 220 с.

7. Дэвид Г. Метод парных сравнений. - М.: Статистика, 1978. 144 с.

8. Рыданова Г.В. Некоторые вопросы статистического анализа случайных бинарных векторов. Автореф. дисс. канд. физ.-мат. наук. - М.: МГУ, 1988. 16 с.

Примеры задач по эконометрике

В настоящем приложении приведены примеры типовых задач, которые решаются на занятиях по эконометрике на факультете "Инженерный бизнес и менеджмент" МГТУ им. Н.Э.Баумана.

Проверка однородности двух независимых выборок

1. В первой случайной репрезентативной выборке объема n_1 положительный ответ дали m_1 опрошенных (респондентов), а во второй случайной репрезентативной выборке объема n_2 положительный ответ дали m_2 опрошенных. Указать доверительные границы для долей (вероятностей положительного ответа в соответствующих генеральных совокупностях) с доверительной вероятностью 0.95 и проверить гипотезу о равенстве долей (уровень значимости α=0.05):

Табл.1. Исходные данные для задачи 1.

	n_1	M_1	n_2	m_2
Вариант 1	400	300	600	500
Вариант 2	857	673	1254	856

2. Для двух независимых выборок объемов n_1 и n_2 даны выборочные средние арифметические $\bar{x}$ и $\bar{y}$ выборочные средние квадратические отклонения s_x, s_y соответственно. Указать доверительные границы для математических ожиданий (с доверительной вероятностью 0.95) и проверить гипотезу о равенстве математических ожиданий с помощью критерия Крамера-Уэлча (уровень значимости α=0.05):

Табл.2. Исходные данные для задачи 2.

	n_1	$\bar{x}$	s_x	n_2	$\bar{y}$	s_y
Вариант 1	100	13,7	7,3	200	12,1	2,5
Вариант 2	213	10,3	5,3	308	12,2	1,7

3. Проверить гипотезу об однородности функций распределения с помощью критерия Вилкоксона (на уровне значимости α=0.05):

Табл.3. Исходные данные для задачи 3.

1 выборка	33	27	12	27	39	42	47	48	50	32
2 выборка	11	20	30	31	22	18	17	25	28	29

Проверка однородности связанных выборках

4. Для каждого из N = 20 объектов даны значения X_j и Y_j , j = 1,2,…,N, результатов измерений (наблюдений, испытаний, анализов, опытов) двух признаков. Необходимо проверить, есть ли значимое различие между значениями двух признаков или же это различие может быть объяснено случайными отклонениям значений признаков. Другими словами, требуется проверить однородность (т.е. отсутствие различия) связанных выборок.

Табл.4. Исходные данные для задачи 4.

j	1	2	3	4	5	6	7	8	9	10
X_j	74	79	65	69	71	66	71	73	72	68
Y_j	73	65	71	69	70	69	78	70	60	62
j	11	12	13	14	15	16	17	18	19	20
X_j	70	69	76	74	72	69	74	72	77	75
Y_j	61	67	73	67	73	64	67	65	63	70

Проверку однородности на уровне значимости 0,05 проведите с помощью трех критериев:
1) критерия знаков (основанного на проверке гипотезы p = 0,5 для биномиального распределения с использованием теоремы Муавра-Лапласа);
2) критерия для проверки равенства 0 математического ожидания (критерий основан на асимптотической нормальности выборочного среднего арифметического, деленного на выборочное среднее квадратическое отклонение);
3) критерия Орлова (типа омега-квадрат) для проверки гипотезы симметрии функции распределения (разности результатов измерений, наблюдений, испытаний, анализов, опытов для двух признаков) относительно 0.

Метод наименьших квадратов

5. Исходные данные – набор *n* пар чисел $(t_k\ , x_k)$, $k = 1,2,...,n$, где t_k – независимая переменная (например, время), а x_k – зависимая (например, индекс инфляции). Предполагается, что переменные связаны зависимостью

$$x_k = a\, t_k + b + e_k,\ k = 1,2,...,n,$$

где a и b – параметры, неизвестные статистику и подлежащие оцениванию, а e_k – погрешности, искажающие зависимость.

Табл.5. Исходные данные для задачи 5.

t_k	1	3	4	7	9	10
x_k	12	20	20	32	35	42

Методом наименьших квадратов оцените параметры a и b линейной зависимости. Выпишите восстановленную зависимость.
Вычислите восстановленные значения зависимой переменной, сравните их с исходными значениями (найдите разности) и проверьте условие точности вычислений (при отсутствии ошибок в вычислениях сумма исходных значений должна равняться сумме восстановленных).
Найдите остаточную сумму квадратов и оцените дисперсию погрешностей.
Выпишите точечный прогноз, а также верхнюю и нижнюю доверительные границы для него (для доверительной вероятности 0,95).
Рассчитайте прогнозное значение и доверительные границы для него для момента t = 12.
Как изменятся результаты, если доверительная вероятность будет увеличена? А если она будет уменьшена?

Индекс инфляции

6. На основе данных табл.6 рассчитайте индекс инфляции с 14.03.1991 по 14.03.2001 на основе потребительской корзины из продуктов №№ 3, 5, 8, 13, 21, 25.

Табл.6. Номенклатура, годовые нормы потребления и цены (руб.)

№ п/п	Наименование продукта питания	Годовая норма, кг	Цена на 14.03.1991	Цена на 14.03.2001
1	Хлеб пшеничный	59,8	0-50	12
2	Хлеб ржаной	65,3	0-20	10
3	Мука пшеничная	18,5	0-46	10
4	Картофель	124,22	0-10	9
5	Капуста	30,4	0-20	8
6	Помидоры	2,8	0-85	80
7	Столовые корнеплоды	40,6	0-20	9
8	Прочие (лук)	27,9	0-50	8
9	Яблоки свежие	15,1	1-50	20
10	Сахар	19,0	0-90	21
11	Говядина	4,4	2-00	85
12	Субпродукты (печень)	0,5	1-40	45
13	Птица	16,1	2-40	52
14	Колбаса докторская	0,4	2-30	95
15	Копчености	0,3	3-70	200
16	Рыба свежая (минтай)	10,9	0-37	80
17	Сельди	0,8	1-40	40
18	Молоко, кефир	110,0	0-32	17
19	Сметана, сливки	1,6	1-70	50
20	Масло животное	2,5	3-60	70
21	Творог	9,8	1-00	45
22	Сыр и брынза	2,3	3-60	70
23	Яйца, десяток	15,2	0-90	20
24	Масло растительное	3,8	1-80	26
25	Маргарин	6,3	1-20	35

7. Гражданин Иванов в марте 1991 г. получил 150 руб., а в марте 2001 г. - 4000 руб. Во сколько раз изменился его реальный доход за 10 лет? Увеличился или уменьшился?
8. За январь индекс инфляции составил 5 % , а за февраль - 2 % . Чему равен индекс инфляции за два месяца? Каков средний уровень инфляции? Можно ли в данном случае складывать проценты инфляции?
9. Выразите текущий курс доллара США в ценах марта 1991 г.

В задачах 2 и 4 рекомендуется принять, что индекс инфляции за 10 лет (март 1991 г. - март 2001 г.) равен 40.

Упорядочения по средним рангам и по медианам

10. В таблице 7 приведены упорядочения (кластеризованные ранжировки), данные семью экспертами.

Табл.7. Исходные данные к задаче 10.

Эксперты	Упорядочения

1	2 < 3 < 6 < 7 < 1 < 4 < 5
2	3 < {2, 6} < 7< {1, 5}< 4
3	2 < 3 < 7 < 1 < 6 < {4, 5}
4	6 < {3, 7} < 2 < 5 < 4 <1
5	2 < 3 < {6,7} < 4 < 1 <5
6	{2, 3} < 1 < 6 < 4 < 5 <7
7	3 < 6 < 2 < 1 < 4 < 5 < 7

Найти:
1) упорядочение по средним рангам;
2) упорядочение по медианам;
3) согласующую их кластеризованную ранжировку.

Медиана Кемени

11. Дана матрица попарных расстояний для множества бинарных отношений из 9 элементов. Найти в этом множестве медиану для множества из 5 элементов: A_2, A_4, A_5, A_7, A_9.

Табл.8. Исходные данные для задачи 11.

0	2	13	1	7	4	10	3	11
2	0	5	6	1	3	2	5	1
13	5	0	2	2	7	6	5	7
1	6	2	0	5	4	3	8	8
7	1	2	5	0	10	1	3	7
4	3	7	4	10	0	2	1	5
10	2	6	3	1	2	0	6	3
3	5	5	8	3	1	6	0	9
11	1	7	8	7	5	3	9	0

Эконометрика качества (статистический приемочный контроль)

12. Для плана (n, 0) с $n = 27$ найти приемочный уровень дефектности.
13. Для плана (n, 0) предел среднего выходного уровня дефектности не превышает $t = 0{,}02$. Каково минимально возможное n?
14. Даны приемочный уровень дефектности $p_{пр} = 0{,}03$ и браковочный уровень дефектности $p_{бр} = 0{,}09$. Указать какой-либо допустимый план вида (n,0), т.е. план, значение оперативной характеристики которого в точке $p_{пр}$ не меньше 0,95, а в точке $p_{бр}$ не больше 0,10.