

УДК 656(075.8)  
ББК 39я73  
Б791

Рецензенты:

зав. кафедрой «Организация и безопасность движения» МАДИ (ГТУ),  
д-р техн. наук, проф. *А. И. Рябчинский*;  
зав. кафедрой «Эксплуатация транспортных средств» МГИУ,  
д-р техн. наук, проф. *В. И. Сарбаев*

**Болдин А. П.**

Б791 Основы научных исследований : учебник для студ. учреждений высш. проф. образования / А. П. Болдин, В. А. Максимов. — М. : Издательский центр «Академия», 2012. — 336 с.

ISBN 978-5-7695-7171-8

Рассмотрены основополагающие принципы и элементы научных исследований применительно к специфике технической эксплуатации автомобилей и систем наземного транспорта и транспортного оборудования. Дана характеристика и приведены примеры проведения работ в условиях пассивного и активного экспериментов. Достаточно широко представлены отдельные вопросы подготовки и обработки результатов производственных научных исследований с возможностью использования популярной программы STATISTICA (версий 5.5a и 6.0) для среды WINDOWS.

Для студентов учреждений высшего профессионального образования.

УДК 656(075.8)  
ББК 39я73

*Оригинал-макет данного издания является собственностью  
Издательского центра «Академия», и его воспроизведение любым способом  
без согласия правообладателя запрещается*

© Болдин А. П., Максимов В. А., 2012  
© Образовательно-издательский центр «Академия», 2012  
© Оформление. Издательский центр «Академия», 2012

ISBN 978-5-7695-7171-8

## ПРЕДИСЛОВИЕ

Автомобилизация, несущая много позитивного (удобство и быстрота перемещения грузов и пассажиров, облегчение и повышение производительности труда человека), имеет ряд отрицательных черт (загрязнение окружающей среды, негативное влияние на человека и животный мир, потребление материальных и энергетических ресурсов). Поэтому современное автотранспортное предприятие должно иметь не только подвижной состав, отвечающий научным требованиям к его эксплуатационным свойствам, но и построенный на научной основе производственный процесс, начиная от планирования и организации перевозок и заканчивая техническим обслуживанием и ремонтом.

Огромная роль в создании такого производства принадлежит инженеру, деятельность которого на современном этапе немыслима без умения хорошо ориентироваться в научной информации, грамотно анализировать ее, проводить теоретические и экспериментальные научные исследования и принимать на основе этого правильные инженерные решения.

В данном учебнике некоторые назревшие проблемы автомобильного транспорта решаются с позиций системного подхода. Цель данного учебника — познакомить будущих специалистов с ролью автомобилизации и ее организацией, привить им навыки по методике постановки и проведения научных исследований в области технической эксплуатации автомобилей (ТЭА) и диагностики, ознакомить их с основными методами научного исследования, их практическим применением на автотранспортных предприятиях и в сфере автотранспортного комплекса, научить анализировать и делать выводы.

## ВВЕДЕНИЕ

Темпы развития экономики страны в большей степени зависят от внедрения в производство достижений науки и техники, которые в настоящее время развиваются очень быстро. В полной мере это можно отнести к транспортному комплексу России и, в частности, к автомобильному транспорту. В условиях рыночных отношений, когда управленческие решения необходимо принимать в условиях конкуренции, необходимо всестороннее комплексное рассмотрение исследуемых процессов и явлений.

Комплексный характер решаемых задач присущ всем уровням управления сложными организациями и системами современного общества. Применение традиционных методов частных решений для них в большинстве случаев уже не дает достаточного эффекта в связи со слабым учетом возможных последствий от существенных связей с другими областями и участками хозяйственных и промышленно-производственных систем.

Одним из наиболее перспективных направлений является использование принципов так называемого системного подхода. Хотя системный подход первоначально был применен для разработки технических комплексов, последующая практика показала его большую эффективность в области хозяйственного и социального управления. Изучение принципов системного подхода позволит еще более расширить область его эффективного применения.

Сущность научно-технического прогресса (термин «научно-техническая революция» в последнее время считается не отвечающим сущности происходящих в области технического прогресса событиям) заключается в коренных качественных преобразованиях производительных сил и производственных отношений. При этом научно-технический прогресс охватывает не только науку и технику, но и производство. Наука — генератор идей. Техника — их материальное, вещественное воплощение. Производство — пространство, где развертывается функционирование техники, где научно-технические достижения используются людьми для получения им необходимых материальных благ.

Именно взаимосвязи науки, техники и производства обуславливают принципиальное отличие современного общественного развития от предыдущих периодов. Даже в XIX в. наука и техника развивались «параллельно», наука не оказывала влияния на технику, а ее развитие

в большей мере само подталкивалось техникой. В настоящее время прогресс науки вызывает прогресс техники, а достижения производства в значительной степени обуславливают прогресс науки.

Современная наука немыслима без сверхмощных ускорителей, электронных микроскопов, техники сверхнизких температур, космических ракет и другой экспериментальной техники. Таким образом, процессы в области науки, техники и производства протекают не изолированно, а связаны определенным образом. Между компонентами и связывающими процессами науки, техники и производства часто нет четких границ, они как бы «вживаются» друг в друга, образуя органическое целое.

Существенным также является то, что достижения научно-технического прогресса реализуются в социальной сфере и сфере духовного производства.

Наличие таких сложных связей в современном производстве обуславливает необходимость нового подхода к решению большинства общественно-хозяйственных проблем, стоящих в настоящее время.

Прежде всего это обусловлено тем, что одностороннее решение проблемы, дающее положительный эффект в какой-либо одной области, может давать еще большие отрицательные последствия в другой области и быть невыгодным для общества в целом.

Ярким примером этого может служить экологическая проблема, которая возникла вследствие расширения производства в целях удовлетворения возрастающих материальных потребностей общества. Чисто механические прогнозы показывают, что дальнейший рост потребления при сохранении эталонов качества сегодняшнего дня приведет к тому, что через полвека атмосфера, природа и ресурсы нашей планеты окажутся уже непригодными для жизни человеческого общества. Таким образом, задачи повышения материального уровня жизни общества должны решаться с учетом сохранения, восстановления и даже улучшения природной среды. Аналогичные ситуации характерны для решения большого числа задач более узкого профиля.

Однако возможность получения ошибок из-за недостаточного учета множественных связей в настоящее время также возросла. Это связано с тем, что, как правило, решение основных проблем невозможно в короткие периоды даже при условии неограниченных материальных и трудовых затрат. Ярким примером является проблема освоения космического пространства.

Естественно, что при длительном планировании трудно учесть влияние вновь возникших за это время факторов и связей в связи с интенсивным развитием общественного хозяйства и потребностей общества. Однако эффективность корректировки будет тем выше, чем более правильно были намечены конечная цель и программа ее достижения.



Поскольку реализация решения глобальных проблем требует огромных материальных и трудовых ресурсов, то любые, даже не очень значительные, ошибки, выявившиеся позднее из-за неполного учета комплексных связей, обходятся для общества довольно дорого. В то же время практика показывает резко возросшую опасность ошибок при решении комплексных проблем несмотря на большую вооруженность науки. Это происходит от того, что при хорошем понимании комплексности проблемы можно не получить отдачи при ее решении из-за недостатка или неправильной оценки ее возможностей.

## ГЛАВА 1

# Основные понятия и определения учебного курса «Основы научных исследований»

---

### 1.1. Понятия о науке

*Наука* — это целостная социальная система, объединяющая в себе постоянно развивающийся набор научных знаний об объективных законах природы, научную деятельность людей, направленную на создание и развитие этой системы, и учреждения, обеспечивающие научную деятельность (рис. 1.1). Система научных знаний отражена в научных понятиях, гипотезах, законах, научных фактах, теориях, идеях и т. д.

*Научная деятельность* — творческая деятельность, направленная на получение, освоение, переработку и систематизацию новых научных знаний, а следовательно, на расширение системы научных знаний.

Система знаний классифицируется по следующим направлениям:

- отрасли знаний: естественные, общественные и технические науки;
- научные дисциплины: математика, физика, химия, теория автомобиля, техническая эксплуатация автомобилей и т. д.;
- результаты научной деятельности: публикации, патенты, конструкторские разработки и т. д.

Научная деятельность классифицируется по следующим признакам:

- целевое назначение: развитие теории, разработка новой техники, совершенствование (разработка) технологии и т. д.;
- виды научных работ: фундаментальные, прикладные, разработки;
- диапазон исследовательских работ: направления в науке, научная проблема, научная тема, научный вопрос;
- методы исследования: теоретические, экспериментальные и смешанные.

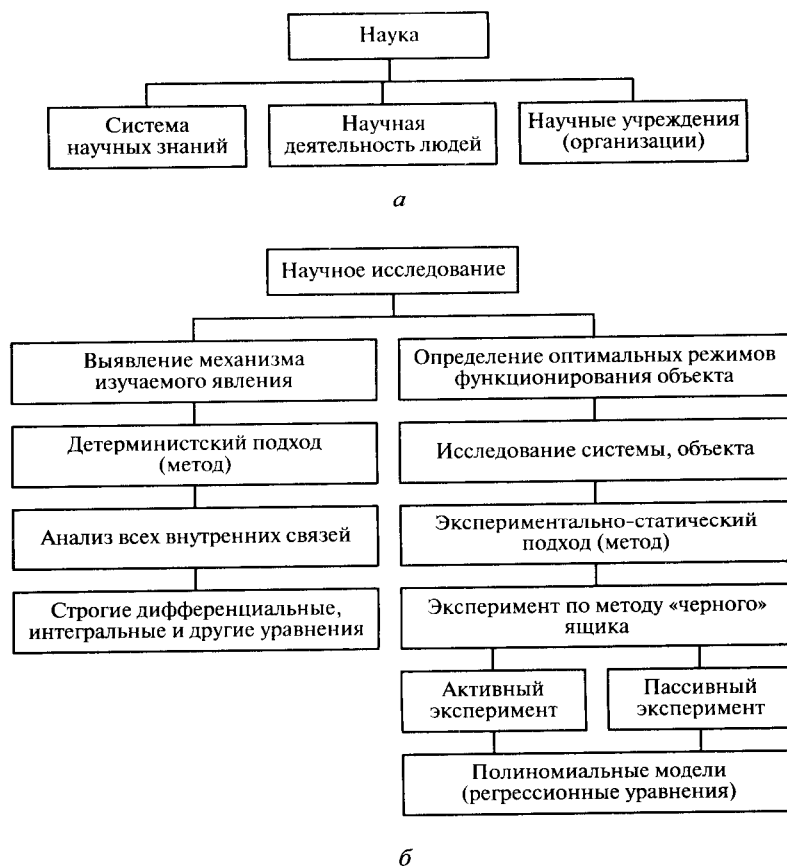


Рис. 1.1. Наука и ее составные части (а) и блок-схема основных целей и подходов научного исследования (б)

Научные учреждения обеспечивают нормальное протекание научной деятельности и включают в себя научных сотрудников, средства научной деятельности (оборудование, приборы и т.д.), объекты научной деятельности (автомобили, автотранспортные установки и т.д.), информационные фонды (библиотеки, патентные фонды и т.д.).

Научные учреждения классифицируются по тому признаку, к какой сфере человеческой деятельности в обществе они относятся:

- к непроизводственной сфере (академические институты, НИИ гуманитарного и общенаучного профилей, вузы непромышленного профиля — медицинские, юридические и т.д.);
- к производственной сфере (отраслевые институты — НИИ, КБ, НПО, технические вузы).

## 1.2. Характерные черты современной науки

Современной науке присущи следующие черты:

1. Связь с производством. Наука стала непосредственной производительной силой. Около 30 % научных достижений служат производству. В то же время наука работает и на себя (фундаментальные исследования, поисковые работы и т.д.), хотя, как показывает опыт, данное направление развивается недостаточно, особенно в области проблем автомобильного транспорта. В области технической эксплуатации следует уделять больше внимания прогностическим и поисковым работам.

2. Массовость современной науки. Наряду с увеличением численности научных учреждений и сотрудников существенно возрастают капитальные вложения в науку, особенно в передовых западных странах. Несмотря на трудности в этом отношении, связанные с переходным периодом к рыночной экономике в жизни России, в бюджетах страны, принимаемых в последнее время, наблюдается устойчивая тенденция увеличения вложений в фундаментальные исследования, имеющие государственное значение.

3. Дробление, специализация, взаимодействие и взаимопроникновение наук. На базе фундаментальных наук (философии, математики, экономики, физики, химии и т.д.) образуются специальные научные дисциплины (биофизика, биохимия, теория автомобиля, техническая эксплуатация автомобилей, транспортная логистика и т.д.); они создаются на основе как смежных наук, так и далеких друг от друга научных дисциплин.

4. Системный подход в изучении объектов исследования. Исследуемый объект рассматривается как некоторое сложное целое, состоящее из отдельных систем, подсистем и элементов. В зависимости от цели и задач исследования наблюдатель может изучать свойства объекта как единого целого, так и его составных частей. Причем в целом объект может обладать такими свойствами, которые не присущи в отдельности ни одной из его составляющих.

5. Резкое ускорение темпов научно-технического прогресса (НТП). Наука представляет собой первую фазу НТП и фактически формирует основы для развития процесса в целом. В условиях НТП основные направления научных исследований сводятся к определению момента перехода на новые качественные этапы развития, а также к определению конкретных форм и методов, позволяющих перейти на эти новые этапы развития.

6. Перевод научной деятельности на хозрасчет и самоокупаемость и в ряде случаев непосредственное участие в процессах производственно-коммерческой деятельности. Научная деятельность в сферах, не связанных непосредственно с интересами государства, должна оправдывать себя в финансовом отношении, становиться прибыльной и рентабельной. В условиях рыночной экономики это является мощ-

ным стимулом ускоренного внедрения достижений науки в общественное производство путем создания и широкого распространения новой техники и новых технологий (например, персональный компьютер, синтезирующий достижения многих наук, был создан и продолжает развиваться с нарастающей интенсивностью прежде всего для удовлетворения требований рынка — усложнения компьютерных игр).

7. Наличие различных источников финансирования:

- бюджетное финансирование (правительственные, межотраслевые, отраслевые и другие федеральные научно-технические программы, региональный бюджет, бюджет местного самоуправления и т. д.);

- внебюджетное финансирование (целевые средства специальных фондов и др.);

- частное инвестирование юридических лиц на основе льготных налогов, ценообразования, аренды;

- частное инвестирование физическими лицами (личные сбережения, льготные займы и т. п.);

- зарубежное инвестирование (целевые займы мирового банка, Европейского банка реконструкции и развития и др.);

- международная помощь и научно-техническое сотрудничество (помощь ЮНЕСКО в области транспорта, международное научно-техническое сотрудничество, помощь и безвозмездный обмен результатами научных исследований в рамках сотрудничества городов-столиц и др.).

Согласно принятым в последнее время постановлениям продукция научной организации является товаром. Для эффективной деятельности научной организации утверждаются следующие экономические нормативы:

- плата за основные производственные фонды (ОПФ), трудовые, природные ресурсы;

- отчисления от расчетной прибыли (доход) в государственный бюджет;

- отчисления от расчетной прибыли, а также от амортизации, предназначенной на полное восстановление основных фондов, в централизованный фонд развития производства, науки и техники и резервы министерства;

- образование фонда научно-технического и социального развития;

- образование фонда материального поощрения и общего фонда заработной платы для научных организаций, применяющих форму хозяйственного расчета, основанную на нормативном распределении прибыли;

- образование фонда валютных отчислений и др.

Государственные бюджетные ассигнования для фундаментальных исследований, которые не могут давать экономической отдачи в бли-

жайшее время или являются необходимой частью духовного и социального развития общества, выделяются исходя из важности конкретной научно-исследовательской темы.

### 1.3. Определение и классификация научных исследований

*Научные исследования* — это творческая деятельность человека, связанная с изучением, анализом и объяснением закономерностей развития окружающей его действительности.

Исследования включают в себя:

- научный труд или научную деятельность человека;
- предмет научного труда;
- средства научного труда.

Научная деятельность человека базируется на конкретных методах познания и связана с получением новых или уточнением старых сведений (данных) об объекте исследования или исследуемом явлении. Предметом научного исследования является объект исследования или исследуемое явление, свойство, связь, на изучение которого направлена деятельность человека.

Средством научного труда является совокупность технических средств обеспечения научного исследования (измерительное оборудование, приборы и приспособления и т. д.). По степени важности научные исследования подразделяются:

- на выполняемые по государственному плану;
- выполняемые по заданию государственных министерств, агентств и их подразделений;
- выполняемые по заданию местных органов самоуправления (например, «Экологическая программа правительства Москвы»);
- выполняемые по инициативе научно-исследовательской организации;
- выполняемые по договорным отношениям с коммерческими, государственными и негосударственными организациями, предприятиями и фирмами.

В зависимости от источников финансирования научные исследования подразделяются на госбюджетные, финансируемые из средств госбюджета, и хоздоговорные, финансируемые в соответствии с договорами, заключенными между заказчиками и исполнителями.

По длительности разработки научные исследования подразделяются на краткосрочные (срок выполнения — до одного года), среднесрочные (срок выполнения — от одного года до пяти лет) и долгосрочные (срок выполнения — более пяти лет).

По виду связи с общественным производством различают научные исследования, направленные:

- на создание новых средств производства и новых технологий;

- совершенствование производственных отношений;
- совершенствование общественных отношений, социальной сферы деятельности человека, повышение уровня духовной жизни и т.д.

Научные исследования в зависимости от целевого назначения, глубины научной проработки, степени связи с природой или промышленным производством подразделяются на фундаментальные, прикладные, научно-исследовательские и опытно-конструкторские разработки (НИР и НИОКР).

Целями фундаментальных исследований являются получение новых законов развития, вскрытие связей между явлениями (вид, форма и направление связей), создание новых теорий и открытий. Они составляют основу развития науки несмотря на то, что вероятность получения положительного результата составляет около 10 %.

Целью прикладных исследований, включая проектирование, является привязка результатов фундаментальных исследований к конкретным условиям производства и жизнедеятельности человека. Объектом их исследования являются различного рода технические системы и новые технологии. Вероятность получения положительного результата при проведении прикладных исследований составляет 20...90 %.

Целью научно-исследовательских и опытно-конструкторских разработок, включая опытное производство, является создание на базе основных результатов функциональных и прикладных исследований опытных образцов техники, новых технологических процессов или усовершенствование существующих технологий и оборудования. Вероятность получения положительного результата при проведении НИОКР составляет 50...90 %.

По способу реализации научные исследования подразделяются на теоретические и экспериментальные. В ходе теоретических исследований определяются основные методы и критерии исследования, задаются необходимые ограничения, описываются внешние и внутренние связи и т.д. В ходе экспериментальных исследований создается модель исследуемого объекта, снимаются и обрабатываются необходимые данные, проводятся проверка результатов теоретических исследований и их дальнейшее развитие и уточнение.

#### 1.4. Методы научного исследования при технической эксплуатации автомобилей

Работоспособность автомобилей и автомобильных парков обеспечивается технической эксплуатацией автомобилей.

*Техническая эксплуатация автомобилей* как область практической деятельности — это комплекс взаимосвязанных технических, экономических, организационных и социальных мероприятий, обеспечивающих:

- своевременную передачу службе перевозок или внешней клиентуре работоспособных автомобилей, требуемых технико-эксплуатационных свойств, в необходимом количестве и номенклатуре и в нужное для потребителя время;

- поддержание автомобильного парка в работоспособном состоянии при рациональных затратах трудовых и материальных ресурсов, нормативных уровнях дорожной и экологической безопасности и нормативных условиях труда персонала.

Как отрасль науки ТЭА определяет пути и методы управления техническим состоянием автомобилей и автомобильных парков для обеспечения регулярности и безопасности перевозок при наиболее полной реализации технико-эксплуатационных свойств автомобилей, заданных уровней их работоспособности и технического состояния, оптимизации материальных и трудовых затрат и минимума отрицательного влияния автомобильного транспорта на окружающую среду, население и персонал. Эффективность ТЭА обеспечивается инженерно-технической службой (ИТС), которая реализует цели и задачи ТЭА.

Под *методом научного исследования* понимают способ или совокупность способов, реализация которых позволяет достичь намеченной цели исследования.

В основе любого научного исследования прежде всего должны лежать общедиалектические и системные методологии, которые вооружают исследователя знанием общих принципов познания современного мира и являются всеобщей основой исследования. Кроме того, каждая наука, которой является и техническая эксплуатация автомобилей, имеет свои конкретные методы исследования.

В технической эксплуатации автомобилей получили развитие следующие универсальные для технических наук методы исследования.

1. *Анализ* — метод научного познания, заключающийся в том, что объект исследования мысленно расчленяется исследователем на более мелкие подобъекты или выделяются характерные свойства и качества объекта для их детального изучения. Анализ позволяет выделить главные звенья любого объекта, исследовать основные связи, т.е. понять суть происходящего. Так, при изучении надежности автомобилей в эксплуатации сначала выделяют четыре свойства надежности (безотказность, долговечность, ремонтпригодность и сохраняемость), а затем изучают их по отдельности.

2. *Синтез* — метод научного познания объекта как единого целого или присущих ему свойств. Он используется для исследования сложных систем после того, как выполнен анализ отдельных элементов системы. Анализ и синтез взаимосвязаны и дополняют друг друга.

3. *Индуктивный метод исследования*, заключающийся в том, что по результатам единичных наблюдений делают общие выводы,

на основании которых судят о связях и свойствах неизвестных объектов.

4. *Дедуктивный метод*, основанный на выводе частных положений из общих правил, законов, суждений, распространен в технике, математике, где из общих законов или аксиом выводятся частные закономерности.

5. *Научное абстрагирование* — метод, применяемый в случаях, когда необходимо сосредоточить внимание на основных элементах, связях, свойствах исследуемого объекта, не останавливаясь на частных или второстепенных его элементах или связях (например, приложение сил, действующих на автомобиль при его движении, к центру масс (центру тяжести)).

6. *Формализация*, заключающаяся в том, что исследуемый объект описывается математическими терминами и формулами. При этом конкретизируются цель и задачи исследования, более четко определяются условия их решения.

7. *Аналогия*, или *подобие* (сходство по какому-то признаку в целом различных объектов), заключающееся в том, что по сходству свойств изученных объектов делается вывод о сходстве еще не изученных свойств (например, назначение периодичности и трудоемкости технического обслуживания (ТО) и текущего ремонта (ТР) вновь созданного автомобиля по нормативам, характерным для его аналога (прототипа)).

8. *Моделирование* — метод научного исследования, при котором изучение свойств объекта проводится на упрощенной модели объекта, а не на нем самом непосредственно.

Перечисленные ранее методы научного исследования связаны между собой и в конкретном научном исследовании применяются комплексно и дополняют друг друга (физическое, имитационное моделирование и т. д.).

## 1.5. Выбор темы научного исследования

При выборе темы научного исследования необходимо оценить ее перспективность. В настоящее время приобретают особое значение численные методы оценки, среди которых можно выделить математический метод и метод экспертных оценок.

В основе *математического метода* лежат показатели (обычно экономические), определяющие перспективность исследований. Например:

$$K_3 = \frac{V_r C_{ед} P_n P_v \sqrt{T}}{3_n + 3_o + 3_n}, \quad (1.1)$$

где  $K_3$  — параметр экономической перспективности;  $V_r$  — объем продукции в год, внедряемой после освоения данной темы, ед./год;

$C_{ед}$  — стоимость единицы продукции, усл. ед.;  $P_n$  — вероятность научного успеха в разработке темы;  $P_v$  — вероятность внедрения научных разработок;  $T$  — продолжительность производственного внедрения, лет;  $3_n$  — общие затраты на научные исследования, усл. ед.;  $3_o$  — затраты на опытное и промышленное освоение, усл. ед.;  $3_n$  — затраты на производство продукции, усл. ед.

Формулу (1.1) можно представить в виде

$$K_3 = \frac{\Theta_o}{3_n} (1 - P_p), \quad (1.2)$$

где  $\Theta_o$  — общий ожидаемый экономический эффект, усл. ед.;  $P_p$  — вероятность риска.

Чем больше  $K_3$ , тем предпочтительнее тема.

В последнее время широкое применение получают и *методы экспертных оценок*. Планируемую тему оценивают специалисты-эксперты, используя при этом баллы, ранги и т. д. После соответствующей математической обработки результатов экспертизы различных направлений и тем выявляются наиболее приоритетные. При этом пользуются информационными материалами разного уровня достоверности — от высказываний отдельных специалистов, которые могут быть и заинтересованы в этих оценках, до конкретной информации по опыту эксплуатации систем и изделий, применению технологических, организационных и управленческих решений. В этих случаях возможна классификация информации (табл. 1.1), на основе которой исходный показатель  $\Pi_n$ , полученный на основе информации  $i$ -го класса, при расчетах  $\Pi_p$  корректируется с помощью коэффициента  $K_i$ , зависящего от класса информации и вида оценок:

$$\Pi_p = K_i \Pi_n. \quad (1.3)$$

Нижняя граница поправочного коэффициента  $K_i^n$  используется при расчете показателей эффективности, а верхняя граница поправочного коэффициента  $K_i^p$  — при расчете затрат.

Таблица 1.1

**Шкала количественной оценки корректирования расчетных показателей с учетом качества информации** (по данным профессора Е. С. Кузнецова)

Характеристика информации	Класс информации	Коэффициенты	
		$K_i^n$	$K_i^p$
Имеется ограниченный опыт эксплуатации изделия (системы). Проведены приемочные испытания	10	0,8	1,2

Окончание табл. 1.1

Характеристика информации	Класс информации	Коэффициенты	
		$K_i^H$	$K_i^B$
Проведены приемочные испытания в лабораторных, заводских условиях	9	0,7	1,25
Имеется опыт эксплуатации или проведены приемочные испытания прототипов или аналогов со сходными принципами работы и процессами	8	0,7	1,3
Проведены приемочные испытания прототипов или аналогов со сходными принципами работы и процессами в лабораторных или заводских условиях	7	0,6	1,4
Имеется техническое задание	6	0,5	1,4
Проведены теоретические расчеты, имеется концепция системы или изделия	5	0,4	1,6
Проведена экспертная оценка	4	0,3	1,7
Имеется зарубежная информация о создании аналогичного изделия или системы	3	0,2	1,8
Имеются систематизированные суждения специалистов	2	0,1	1,9
Публикации в отдельных литературных источниках	1	0,07	2,0
Информация отсутствует или не обнаружена	0	—	—

Примером применения экспертизы является оценка рисков в известном пакете программ Project Expert для следующих стадий проекта:

- исследования и разработки;
- приобретение и аренда земли;
- строительство, аренда или приобретение зданий и сооружений;

- приобретение и монтаж технологического и офисного оборудования;
- разработка и изготовление технологической оснастки и инструмента;
- производство;
- рынок;
- продукт, изделие;
- система распределения;
- реклама;
- сервис.

На каждой стадии эксперты оценивают уровень риска качественно (высокий, средний, низкий), рассматривая следующие показатели: реальность идеи; наличие необходимых специалистов; качество управления; финансирование; безопасность; экологичность; взаимодействие с местными властями и населением; чувствительность к законодательству; готовность среды; приспособленность к среде.

### 1.6. Этапы научного исследования

Научное исследование состоит из следующих этапов:

- 1) состояние вопроса исследования;
- 2) теоретические исследования;
- 3) экспериментальные исследования;
- 4) анализ и обобщение результатов теоретических и экспериментальных исследований;
- 5) расчет экономической эффективности и опытная апробация предлагаемых разработок.

**Состояние вопроса исследования.** В ходе реализации данного этапа проводятся патентно-лицензионный поиск, обзор и анализ НИР, НИОКР, монографий, статей по рассматриваемой проблеме. В результате формируются основные выводы и определяются цель и задачи исследования. Кроме того, в завершение данного этапа разрабатывается общая методика исследования.

Общая методика исследований представляет собой набор способов, способствующих последовательному наиболее эффективному осуществлению научного исследования.

**Теоретические исследования.** На данном этапе проводятся:

- формирование рабочей гипотезы исследования;
- обоснование, выбор и формирование целевой функции;
- анализ и выбор влияющих факторов;
- обоснование и выбор математического аппарата;
- аналитическое сравнение альтернатив развития исследуемого процесса и др.

**Экспериментальные исследования.** На данном этапе разрабатывается методика экспериментальных исследований, монтируется

экспериментальная установка, разрабатываются учетные анкеты, осуществляется сбор экспериментальных данных, обосновывается необходимость применения средств измерения, проверяется их точность, определяется количество и обратное число опытных точек, намечаются критерии и методики обработки опытных данных.

**Анализ и обобщение результатов теоретических и экспериментальных исследований.** На данном этапе проводится обработка полученного экспериментального материала, сравнение его с результатами теоретических исследований. По результатам анализа формируются новые научные положения, выводы, заключения и предложения.

**Расчет экономической эффективности и опытная апробация предлагаемых разработок.** На данном этапе проводится расчет экономической эффективности предложенных разработок или полученных результатов. Расчет экономической эффективности целесообразно проводить с государственных или общехозяйственных позиций с учетом социального эффекта (улучшение качества бытового обслуживания населения, качества перевозок пассажиров и т.д.).

### 1.7. Основные цели и подходы научного исследования, сущность пассивного и активного эксперимента

Различают две основные цели научного исследования:

- 1) выяснение механизма научного явления (поиск математических, логических и других моделей);
- 2) определение оптимальных режимов функционирования объекта, системы (используется, когда известен механизм явления).

**Детерминистский подход (ДП)** — получение функциональных зависимостей между параметрами объекта; при этом исключаются внешние связи и исследуются все внутренние связи. Под функциональной зависимостью понимается зависимость вида

$$Y = f(X_i),$$

при которой каждому значению функции  $Y$  соответствует одно вполне определенное значение фактора или аргумента  $X$ . Изучением таких зависимостей занимается математический анализ.

По результатам наблюдений составляется протокол наблюдений. Допустим, исследователь сделал 10 наблюдений и получил следующие данные:

Вход ( $X$ )	8	5	16	20	11	7	2	12	4	19
Выход ( $Y$ )	25	16	49	61	34	22	7	37	13	58

Анализ этого протокола позволяет установить, что система функционирует в соответствии с уравнением  $Y = 3X + 1$ . Достоинства ДП

закljučаются в том, что получение модели можно распространять на похожие явления, известные внутренние закономерности; недостатки — в том, что данный подход применяется лишь при несущественных допущениях, которые в практике исследования технических систем встречаются редко.

При **экспериментально-статистическом подходе (ЭСП)** одно и то же воздействие на объект исследования приводит к различным результатам, каждый из которых наступает с некоторой вероятностью. В основе ЭСП лежит эксперимент по методу «черного ящика» (рис. 1.2), идея которого заключается в следующем:

- 1) исследуемый объект рассматривается как отдельная система окружающего мира, имеющая внешнюю среду;
- 2) внешняя среда воздействует на систему через входы

$$X = (x_1, x_2, \dots, x_m);$$

- 3) система воздействует на внешнюю среду через выходы

$$Y = (y_1, y_2, \dots, y_n);$$

- 4) внутренние состояния системы характеризуются параметрами

$$S = (S_1, S_2, \dots, S_k).$$

К определенному моменту времени  $t$  будет иметь место следующая зависимость:  $Y_t = f(X_t, S_t)$ , т.е. состояния выходов определяются состояниями входов и внутренними состояниями системы. Часть входных параметров может быть управляемыми, а часть — будут составлять помехи.

Выходные параметры могут быть техническими или экономическими. В системе, считающейся «черным ящиком», структура и внутренние связи скрыты от наблюдателя. Исследователь фиксирует лишь состояния входов и выходов и анализирует наличие связи между ними; при этом используется протокол наблюдения.

Рассмотрим пример. Пусть дана система, считающаяся «черным ящиком». Она имеет один вход  $X$  и один выход  $Y$ . Наблюдения за входом и выходом показали следующие результаты:

Вход ( $X$ )	1	3	4	6	9	10	12	15	17	18
Выход ( $Y$ )	3	9	14	21	27	34	35	49	50	53



Рис. 1.2. Система и внешняя среда

Анализ результатов наблюдений показывает, что более высокой числовой характеристике входа соответствует большая числовая характеристика выхода. В данном случае  $Y$  превышает  $3I$  приблизительно в три раза, т.е. имеет место статистическая зависимость.

Математическая обработка полученных результатов (с помощью метода наименьших квадратов, см. гл. 4) дает следующее уравнение регрессии:

$$Y = 1,47 + 2,95X.$$

Данная зависимость определена для вероятностной системы и является корреляционной зависимостью.

**Корреляционная зависимость** — это зависимость, при которой случайному значению аргумента соответствует случайное значение функции.

Кроме того, в теории планирования эксперимента встречается регрессионная зависимость, при которой неслучайному значению аргумента соответствует случайное значение функции. Если фактические значения аргумента подставить в полученную формулу, то они не будут совпадать. Налицо определенные отклонения (табл. 1.2).

Таблица 1.2

Расчетные и фактические состояния выхода

Состояние входа $x_n$	Фактическое состояние выхода $y_n$	Расчетное состояние выхода $y'_n$	Отклонение фактических значений от расчетных ( $y_n - y'_n$ )
1	3	4,4	-1,4
3	9	10,3	-1,3
4	14	13,3	+0,7
6	21	19,2	+1,8
9	27	28,0	-1,0
10	34	31,0	+3,0
12	35	36,9	-1,9
15	49	45,7	+3,3
17	50	51,6	-1,6
19	53	54,6	-1,6

Таким образом, в результате ЭСП исследуется не сам объект, а результаты эксперимента. При этом возможен пассивный и активный эксперименты. При *пассивном эксперименте* (см. гл. 2, 3, 4) исследователь не влияет на вход системы.

Достоинства пассивного эксперимента:

- наблюдатель не нарушает хода процесса;
- можно использовать результаты ранее выполненных исследований.

Недостатки пассивного эксперимента:

- требуется длительный период наблюдений;
- в силу коррелированности факторов затрудняется процедура определения коэффициентов модели;
- отсутствие симметричности уровней факторов затрудняет статистическую оценку значимости коэффициентов и проверку математической модели на адекватность.

При *активном эксперименте* (см. гл. 5) исследователь вмешивается в ход эксперимента, влияя на входы системы. Достоинство активного эксперимента заключается в том, что сокращается срок проведения эксперимента; недостаток — требуется более детальная подготовка (планирования эксперимента).

Экспериментально-статистический подход используется, если:

- неизвестна зависимость;
- нужно найти коэффициенты в модели;
- нужно найти экстремум, не строя модель.

### КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Дайте определения понятия «наука» и ее составляющих.
2. По каким признакам классифицируется система научных знаний?
3. Назовите основные черты современной науки и дайте им краткую характеристику.
4. Назовите экономические нормативы, утверждаемые хозрасчетной научной организацией.
5. Дайте определение понятия «научное исследование».
6. По каким признакам классифицируются научные исследования?
7. Дайте краткую характеристику фундаментальным, прикладным исследованиям и научно-исследовательским разработкам.
8. Дайте определение понятия «научное исследование». Перечислите основные методы научного исследования.
9. В чем суть математического метода обоснования выбора темы научного исследования?
10. Дайте краткую характеристику основных этапов научного исследования.
11. Дайте краткую характеристику основных целей и подходов научного исследования.
12. Дайте краткую характеристику пассивного эксперимента.
13. Дайте краткую характеристику активного эксперимента.



**Применение закономерностей рассеяния непрерывных случайных величин при проведении исследований эксплуатационной надежности автомобилей и других показателей их работы на автотранспортных предприятиях**

**2.1. Случайные величины и возможности обработки экспериментальных данных на их основе компьютерными программами**

*Случайной* называется величина, которая при реализации определенного комплекса условий может принимать то или иное значение, но какое именно — неизвестно. Различают дискретные и непрерывные случайные величины (СВ). *Дискретные СВ* — это такие величины, которые могут принимать конечное или бесконечное счетное множество значений и которые могут быть определенным образом пронумерованы (например, целочисленные величины, количество отказов за время обкатки автомобиля и т.д.).

*Непрерывные СВ* — это такие величины, которые могут принимать бесконечное несчетное множество значений в заданных интервалах (например, время безотказной работы автомобиля, его агрегата и т.д.).

Практически подавляющее большинство экспериментальных данных, получаемых в автотранспортных предприятиях и организациях и при исследовании тех или иных характеристик работы автомобилей в эксплуатации, представляют собой непрерывные случайные величины (например, наработка до наступления отказа или неисправности для отдельного автомобиля; трудоемкость и затраты времени на устранение этого отказа или неисправности; значение диагностического параметра, отражающего техническое состояние контролируемого узла конкретного автомобиля после его определенного пробега, и т.д.).

Обработка таких данных в целях выявления их характеристик, которые затем можно распространять на другие подобные случаи использования автомобиля, для которых не проводились специальные исследования, производится методами статистики, часть из которых уже длительный период признана стандартными и не подвергается корректировке; другая часть разработана относительно недавно

(30...50 лет назад) для повышения качества решения статистических задач.

В то же время потребности современной жизни приводят к необходимости дальнейшего совершенствования статистической обработки и специфики ее использования, например для анализа результатов социологических опросов населения, что обуславливает появление новых методов. Очевидно, указанный процесс будет продолжаться и дальше, а освоение как старых, так и новых методов необходимо для проведения качественных научных исследований во многих областях, в том числе в области эксплуатации автомобильного транспорта.

Методы статистики, ранее используемые в основном в научных исследованиях и «тяжело» воспринимаемые в прежнее время из-за необходимости проведения сложных «многоэтажных» ручных расчетов, в настоящее время все чаще употребляются для решения практических задач в обычной производственной (инженерной, управленческой и коммерческой) деятельности и бытовой жизни, что привело к необходимости резкого упрощения процессов их освоения и применения. Это стало возможным благодаря универсальным и специализированным компьютерным программам статистической обработки и упрощенным методикам их изучения, которые постоянно совершенствуются наряду с созданием новых, упрощающих процесс обучения и использования.

Эти программы имеют различные уровни интеграции и степени совершенства и требуют постоянного, в основном самостоятельного, изучения. В частности, такой программой является интегрированная система STATISTICA для среды WINDOWS, многие модули которой могут быть полезны для обработки результатов в области эксплуатации автомобилей. Возможности этой программы, к сожалению попадающей к российскому пользователю в частично русифицированном варианте или только на английском языке, позволяют значительно повысить эрудицию исследователя. При дальнейшем изложении материала будут даваться ссылки на указанную программу с комментариями ее возможностей.

Российскими исследователями разработано большое количество узкоспециализированных статистических программ различного назначения. В ряде случаев они являются более простыми в использовании и дают лучшие в качественном отношении результаты, хотя и ограничены по количеству рассматриваемых вариантов (в данном учебнике основное внимание будет уделено именно этим программам). В целом уже обязательной считается обработка экспериментальных результатов путем использования компьютерных программ, дающих более широкий информационный выход при незначительных затратах времени, прежде всего из-за опасности ошибок ручного расчета. Однако освоение отмеченных ранее импортных и отечественных программ сильно зависит от теоретической подготовки

(в данном учебнике изложен минимально необходимый уровень, требующийся для выполнения исследований в области технической эксплуатации автомобилей).

## 2.2. Обработка случайных величин, связанных с рассеянием изучаемого показателя, на примере изучения долговечности автомобильных деталей, узлов и агрегатов

Согласно стандартным рекомендациям [8, 10] наиболее обоснованным подходом при оценке долговечности критических деталей, агрегатов, систем и узлов подвижного состава для конкретных условий эксплуатации является организация систематических наблюдений за работой достаточно представительной группы (выборки) автомобилей объемом порядка 20...30 единиц. За начало испытаний принимается или момент их начальной эксплуатации, или момент постановки на автомобили новых деталей, агрегатов и узлов; концом испытаний считается момент наступления отказов обследуемых элементов у всех автомобилей выборочной партии.

Такие испытания (по плану  $NUN$ ) [10] называются полными и дают в результате наиболее точные оценки показателей надежности и долговечности, в отличие от усеченных испытаний, проводимых по планам  $NRr$ ,  $NRT$  (и др.) в целях сокращения продолжительности испытаний. В результате полных испытаний обычно фиксируются значительно отличающиеся друг от друга выборочные значения  $x_i$  непрерывной случайной величины, при дальнейшей обработке которых находят ее основные характеристики, или статистики, к которым относятся:

- а) числовые характеристики СВ;
- б) закон распределения СВ.

При обработке статистических данных, в первую очередь, производят оценку числовых характеристик СВ. Различают числовые характеристики положения и рассеяния СВ. К характеристикам положения относятся математическое ожидание, медиана и мода; к характеристикам рассеяния — размах, дисперсия, стандартное (среднеквадратическое) отклонение, коэффициент вариации, асимметрия и эксцесс.

Стандартная обработка результатов статистических наблюдений представляет собой последовательные оценки математического ожидания, размаха, дисперсии, среднеквадратического отклонения, коэффициентов вариации, асимметрии и эксцесса, а понятия моды и медианы приобретают информационное значение только после графической интерпретации СВ и определения ее закона распределения.

Под *модой* понимается такое значение случайной величины, которому соответствует наибольшая вероятность для дискретной СВ, или наибольшее значение плотности вероятности для непрерывной СВ.

*Медианой* называется такое значение непрерывной СВ, относительно которого равновероятно получить как большее, так и меньшее значение. При статистических исследованиях, когда как дискретные, так и непрерывные случайные величины могут представляться в дискретном виде (например, при разбиении на интервалы и при построении гистограмм), могут возникать случаи так называемых многомодальных распределений, которые обычно обусловлены нехваткой данных или их неправильным группированием по интервалам, что обычно проверяется.

Однако в наиболее сложном случае наличие более чем одной моды может свидетельствовать о так называемой смешанной выборке, в которой находятся статистические данные, отражающие другие процессы (например: отказы автомобилей, связанные с установкой на них не только нормальных, но и дефектных деталей на заводе-изготовителе и в эксплуатации; время ремонта при устранении не только этих отказов, но и последствий дорожно-транспортных происшествий; значения диагностического параметра, отражающего как исправное, так и неисправное состояние, и др.). Обработка подобной статистики, несущей дополнительную и, как правило, более важную информацию, уже потребует усложненных и в ряде случаев специализированных подходов.

При расчетах обычно определяют не истинные значения числовых характеристик, а их статистические оценки, отличающиеся от истинных на какую-то величину ошибки. Поэтому используются различные обозначения для истинных и статистических оценок, среди которых наиболее популярными являются выделения «звездочкой» статистических оценок без изменения математического символа (или же производится замена математического символа).

*Математическое ожидание* СВ ( $\bar{X}$ ) — это число, относительно которого при неограниченном увеличении числа опытов устойчиво стабилизируется среднее арифметическое значение. Статистическое значение математического ожидания для экспериментальных данных определяется по величине выборочного среднего:

$$\bar{X}^* = \frac{\sum_{i=1}^N x_i}{N}. \quad (2.1)$$

*Размах* представляет собой разность между максимальным и минимальным значениями СВ:

$$\Delta x = x_{\max} - x_{\min}.$$

Дисперсия СВ представляет собой математическое ожидание квадрата отклонения СВ от ее математического ожидания:

$$D(x) = M[(x - \bar{X})^2]. \quad (2.2)$$

Статистическое значение дисперсии для экспериментальных данных определяется по формуле

$$\bar{D}^* = S^2(x) = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X}^*)^2. \quad (2.3)$$

В связи с тем, что дисперсия СВ имеет разную размерность по сравнению с СВ, ее использование на практике затруднено. Обычно при расчетах используют среднее квадратическое или стандартное отклонение, которое представляет собой корень квадратный из дисперсии:

$$\sigma(x) = \sqrt{D(x)}, \quad (2.4)$$

и статистическое (выборочное) значение равно:

$$S(x) = \sqrt{\frac{\sum (x_i - \bar{X}^*)^2}{N-1}}. \quad (2.5)$$

По выражениям (2.1), (2.3) и (2.5) производят статистические оценки среднего, дисперсии и стандартного отклонения не только для непрерывных, но и для дискретных СВ. Деление суммы квадратов разницы в выражениях (2.3) и (2.5) на значение  $(N-1)$  вместо логически естественного значения  $N$  в статистике обусловлено необходимостью получения так называемой несмещенной оценки для дисперсии, с тем чтобы компенсировать обычно недостаточное для получения точной оценки количество экспериментальных данных  $N$ ; с увеличением количества экспериментальных данных данная поправка становится несущественной.

*Коэффициент вариации* имеет ограниченное информационное использование для непрерывных СВ (его определение является обязательным при оценке показателей надежности работы в эксплуатации механических, электронных и других устройств) и представляет собой отношение среднего квадратического отклонения к среднему значению:

$$V_x = \frac{\sigma(x)}{\bar{X}}. \quad (2.6)$$

При этом статистическое значение оценки

$$V_x^* = \frac{S(x)}{\bar{X}^*}. \quad (2.7)$$

*Коэффициент асимметрии* характеризует асимметрию кривой распределения. Статистическое значение коэффициента асимметрии определяется по формуле

$$A_k^* = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{X}^*)^3}{S^3}. \quad (2.8)$$

*Коэффициент эксцесса* является показателем островершинности кривой закона распределения после его подбора по экспериментальным данным. Статистическая оценка коэффициента эксцесса определяется по формуле

$$E_k^* = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{X}^*)^4}{S^4} - 3. \quad (2.9)$$

Выборочные значения всех перечисленные выше показателей СВ обычно рассчитываются универсальными компьютерными программами (например, STATISTICA для WINDOWS) непосредственно по выражениям (2.1), (2.3), (2.5), (2.8) и (2.9) одновременно со значениями стандартных ошибок для средней величины и коэффициентов асимметрии и эксцесса. Сопоставление с показателями по асимметрии и эксцессу в первом приближении позволяет сделать заключение о соответствии статистического распределения нормальному закону, на котором базируются основные положения статистики. (Методика проверки на нормальность по статистическим оценкам коэффициентов асимметрии и эксцесса приведена далее; теоретически эти показатели для нормального закона равны нулю). Для специализированных программ оценки асимметрии и эксцесса и их стандартных ошибок, а также стандартной ошибки для среднего могут отсутствовать, хотя значение последней несложно получить «ручным» расчетом.

### 2.3. Графическая интерпретация случайных величин и построение гистограмм

При обработке результатов экспериментальных исследований для их наглядности и дальнейшего анализа более удобно переходить к графическому представлению. Это позволяло при «ручной» обработке (при отсутствии компьютеров и специализированных калькуляторов) значительно сокращать объем расчетов и избегать ошибок, поскольку вычисление дисперсии по выражению (2.3) становилось нереальным уже при количестве  $N > 10 \dots 15$  единиц.

Наиболее распространенным способом графического представления является построение гистограммы распределения СВ, которое «вручную» выполняется по следующей методике:

- определяются максимальное  $x_{\max}$  и минимальное  $x_{\min}$  значения показателей выборки;

- находится размах  $\Delta x$  случайной величины, представляющий собой разность между ее максимальным и минимальным значениями:

$$\Delta x = x_{\max} - x_{\min};$$

- определяется количество интервалов, которое должно стремиться к числу

$$k = 1 + 3,32 \lg N;$$

- определяется значение интервала (по формуле Стьеджера):

$$\delta = \frac{\Delta x}{k} = \frac{x_{\max} - x_{\min}}{1 + 3,32 \lg N}; \quad (2.10)$$

- подсчитывается количество наблюдений  $n_j^*$ , находящееся в каждом интервале;

- вычисляются частоты попадания наблюдений в каждый интервал:

$$p_j^* = \frac{n_j^*}{N}, \quad \varphi = 1, 2, \dots, k. \quad (2.11)$$

При этом очевидным является следующее соотношение:

$$\sum_{j=1}^k p_j^* = 1. \quad (2.12)$$

По полученным частотам более просто произвести «ручной» расчет среднего значения и дисперсии случайной величины, принимая для упрощения [2, 10], что все величины  $n_j^*$ , находящиеся в определенном интервале, становятся одинаковыми и равными среднему значению интервала:

$$\bar{X}^* = \frac{1}{N} \sum_{j=1}^k x_{jcp} n_j^* = \sum_{j=1}^k x_{jcp} p_j^*; \quad (2.13)$$

$$S^2 = \frac{1}{N} \sum_{j=1}^k (x_{jcp} - \bar{X}^*)^2 n_j^* = \sum_{j=1}^k (x_{jcp} - \bar{X}^*)^2 p_j^*. \quad (2.14)$$

По выражениям (2.13) и (2.14) легко производить оценки и для дискретных случайных величин  $x_1, x_2, \dots, x_m$ , имея для них значения вероятностей  $P^*(x_1), P^*(x_2), \dots, P^*(x_m)$ , которые для подобных случаев обычно представляются в виде таблицы.

По значениям  $n_j^*$  обычно строят гистограмму, показывая по оси абсцисс интервалы разбиения выборки, а по оси ординат — количество попавших в интервалы данных, представляя последние гра-

фически в виде столбиков от основания соответствующих интервалов.

Для примера рассмотрим построение гистограммы для выборки СВ, отражающей значения коэффициента регулярности движения автобусов, приведенные в табл. 2.1.

Таблица 2.1

Выборка случайных величин

96,5	97,7	97,8	97,6	97,5	98,3	97,5	98,4	97,8
98,0	97,8	98,3	98,5	97,4	97,0	97,0	98,5	98,4
97,5	97,7	98,5	96,8	97,8	98,0	97,0	98,4	98,3
97,7	97,0	97,9	97,6	96,8	97,8	97,9	98,3	98,1
97,4	97,6	97,5	97,7	97,5	97,3	97,7	97,9	96,9
97,2	97,5	98,2	97,8	98,1	97,3	97,3	97,5	97,4
98,3	98,8	97,4	97,4	97,0	97,4	98,7	96,9	97,5

Необходимо определить среднее арифметическое значение коэффициента регулярности, среднеквадратическое отклонение и построить гистограмму.

Для начала по данным табл. 2.1 определяем максимальное и минимальное значения коэффициента регулярности  $K_p$ :

$$K_p^{\min} = 96,5; \quad K_p^{\max} = 98,7.$$

Находим размах:

$$\Delta = K_p^{\max} - K_p^{\min} = 98,7 - 96,5 = 2,2.$$

Определяем необходимое число интервалов:

$$K \approx 1 + 3,32 \lg N = 1 + 3,32 \lg 63 = 6,75 \approx 7.$$

Определяем интервал:

$$\delta = \frac{\Delta x}{K} = \frac{2,2}{6,75} \approx 0,320;$$

Строим интервалы:

$$96,50 \dots 96,82; \quad 96,82 \dots 97,14; \quad 97,14 \dots 97,46; \quad 97,46 \dots 97,78; \\ 97,78 \dots 98,10; \quad 98,10 \dots 98,42; \quad 98,42 \dots 98,74.$$

Определяем количество наблюдений в каждом интервале  $n_j^*$  и частоты попадания наблюдений в каждый интервал  $p_j^*$  (полученные данные отражены в первых четырех столбцах табл. 2.2).

Таблица 2.2

Данные проверки нормальности распределения коэффициента регулярности движения автобусов

$K_j$ Границы интервалов коэффициента регулярности	Среднее интервальное значение коэффициента регулярности $K_{jcp}$	Опытное количество наблюдений в интервале $n_j^*$	Частоты экспериментальные $p_j^*$	$K_{jcp} p_j^*$	$(K_{cp} - K_{jcp})^2 p_j^*$	Нормированное отклонение функции Лапласа $Z_j = \frac{K_j - K_{cp}}{S_k}$	Значение функции распределения $F(Z_j) = 0,5 + \Phi(Z_j)$	$P_j = F(Z_{j+1}) - F(Z_j)$	Теоретическое число наблюдений в интервале $N_j = N p_j$	Значение критерия $\frac{n_j^* - n_j}{(n_j^* - n_j)^2}$
96,50						-1,97	0,024			
	96,66	4	0,063	6,09	0,0499			0,025	3,48	0,07
96,82						-1,37	0,085			
	96,98	7	0,111	10,76	0,0361			0,060	8,57	0,29
97,14						-0,77	0,221			
	97,30	11	0,175	17,03	0,0109			0,135	13,42	0,44
97,46						-0,17	0,434			

	97,62	15	0,238	23,23	0,0012			0,20	14,55	0,01
97,78						0,43	0,666			
	97,94	13	0,207	20,18	0,0313			0,23	11,47	0,20
98,10						1,03	0,848			
	98,26	9	0,143	14,05	0,0721			0,23	6,30	1,16
98,42						1,63	0,948			
	98,58	4	0,063	6,21	0,0668			0,06	2,47	0,96
98,74						2,23	0,987			

Примечание:  $N = \sum n_j = 63$ ;  $K_{cp} = \sum K_{jcp} p_j^* = 97,55$ ;  $D_k^* = \sum (K_{cp} - K_{jcp})^2 p_j^* = 0,2683$ ;  $S_k = \sqrt{D_k^*} = 0,532$ ;  $\chi^2 = \sum = 3,13$ .

По выражениям (2.13), (2.14) определяем среднее значение коэффициента регулярности и его среднеквадратическое отклонение (расчеты отражены в пятом и шестом столбцах табл. 2.2):

$$\bar{K}_p^* = 97,55; \sigma_k = 0,532.$$

По граничным значениям интервалов и количеству попадания в них экспериментальных данных  $n_j^*$  строим гистограмму распределения (рис. 2.1). В другом масштабе и обозначениях величин для оси ординат эта гистограмма одновременно отражает опытные (эмпирические) частоты (вероятности)  $p_j^*$  попадания СВ в заданные интервалы.

Гистограмма в первом приближении позволяет получить представление о законе распределения СВ, являющемся ее наиболее полной характеристикой. Графически указанный закон наиболее информативно представлять в виде кривой распределения в дифференциальной форме, отражающей так называемую плотность распределения. Теоретически кривая плотности получается при обработке бесконечно большого количества данных, т. е. когда  $N \rightarrow \infty$ , при этом в соответствии с выражением (2.10) ширина интервала  $\delta \rightarrow 0$ , т. е. основание столбика обращается в точку, а столбик вероятности  $p_j^*$  при графическом представлении — в прямую линию соответствующей высоты  $f_j^*$ , характеризующую плотность вероятности для данной точки диапазона рассеяния СВ.

При слиянии верхние точки прямых линий графически образуют кривую линию, отражающую функцию  $f(x)$  плотности распределения СВ, которая может быть описана соответствующими математическими выражениями. Именно математическое описание закона  $f(x)$  является универсальным представлением СВ, поскольку отражает все ее свойства, а графическая интерпретация функции  $f(x)$  дает практически наглядное представление об этом законе и его свойствах.

Приближенное представление о кривой  $f(x)$  можно получить, соединив плавной линией середины высот столбиков гистограммы. Более точное представление требует достаточно сложных расчетов,

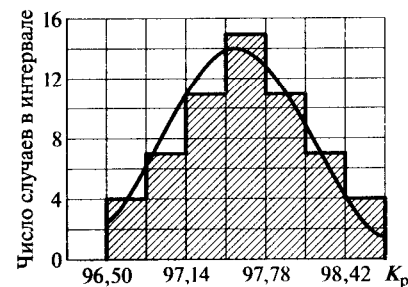


Рис. 2.1. Гистограмма и соответствующая ей теоретическая кривая распределения коэффициента регулярности движения в дифференциальной форме

рассмотренных далее, однако уже не встречает практического затруднения для современных компьютерных программ, предусматривающих обязательное и одновременное графическое представление анализируемой СВ в виде гистограммы и кривой  $f(x)$ .

На рис. 2.1 показана теоретическая кривая  $f(x)$ , отражающая нормальный закон распределения, расчет которого обычно производится автоматически при введении любых данных современными компьютерными программами.

Функция  $f(x)$  должна обладать следующими свойствами:

- плотность распределения — неотрицательная функция для всех  $x_i$ , т. е.  $f(x) \geq 0$ ;
- $f(x)$  имеет размерность, обратную размерности СВ:

$$[f(x)] = \frac{1}{[x]} = [x^{-1}];$$

- функция  $f(x)$  должна быть одномодальной.

Поскольку площадь, ограниченная кривой плотности распределения и осью абсцисс, в графической интерпретации в соответствии с выражением (2.12) должна быть равна единице, то математически это должно быть отражено соотношением

$$\int_{-\infty}^{+\infty} f(x) dx = 1. \quad (2.15)$$

Математическое ожидание СВ в соответствии с выражением (2.13) должно определяться соотношением

$$\bar{X} = \int_{-\infty}^{+\infty} x f(x) dx. \quad (2.16)$$

Дисперсия СВ в соответствии с выражением (2.14) должна определяться соотношением

$$D(x) = \int_{-\infty}^{+\infty} (x - \bar{X})^2 f(x) dx. \quad (2.17)$$

Значения  $\bar{X}$  и  $\sigma(x) = \sqrt{D(x)}$  редко отражаются в математических выражениях, описывающих наиболее распространенные законы распределения  $f(x)$  случайных величин, хотя и являются их основными характеристиками.

Вероятность того, что СВ не превысит заданного значения  $X$ , будет определяться так называемой интегральной функцией распределения по выражению

$$F(x) = P(x \leq X) = \int_{-\infty}^x f(x) dx. \quad (2.18)$$

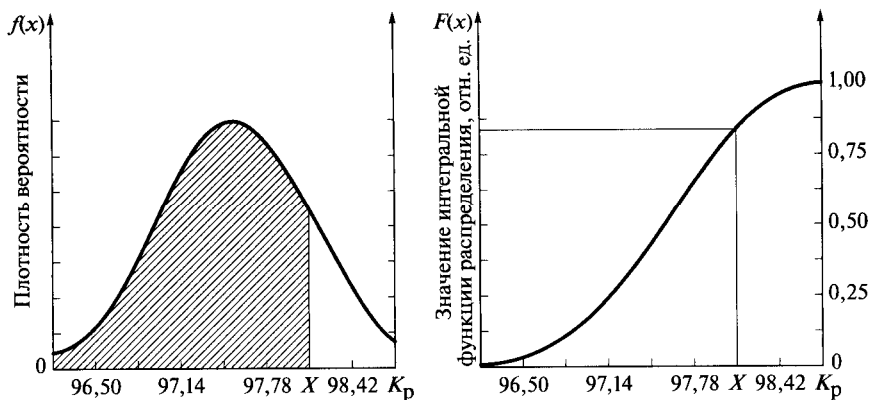


Рис. 2.2. Графическая интерпретация процесса обработки непрерывной случайной величины при использовании дифференциальной и интегральной функций распределения

Интегральная функция, или функция распределения случайной величины  $F(x)$ , является наиболее важной, поскольку в основном используется при статистических расчетах, результаты которых выводятся на экран монитора или представляются в виде таблиц. Однако графическая интерпретация функции распределения  $F(x)$ , как это можно видеть из рис. 2.2, отражающего наиболее типичное представление процессов обработки СВ современными компьютерными программами (включая заштрихованную область, ограниченную величиной  $X$  площади под дифференциальной кривой), является вспомогательным дополнением к более наглядному в информационном отношении виду функции  $f(x)$ .

Интегральная функция  $F(x)$  имеет следующие свойства:

- 1)  $F(x) \geq 0$  для всех  $x$ ;
- 2)  $0 \leq F(x) \leq 1$ ;
- 3)  $F(x)$  — неубывающая функция, т.е.  $F(x + \Delta x) \geq F(x)$ ;
- 4) при  $x = -\infty$   $F(x) = 0$ , а при  $x = +\infty$   $F(x) = 1$ ;
- 5) вероятность того, что значение СВ  $X$  заключено в интервале  $x$  и  $x + \Delta x$ , обычно определяется через функцию  $F(x)$ :

$$P(x \leq X \leq x + \Delta x) = \int_x^{x+\Delta x} f(x) dx = F(x + \Delta x) - F(x). \quad (2.19)$$

Функция  $F(x)$  используется как для непрерывных, так и для дискретных случайных величин. Для дискретной СВ  $F(x) = \sum P(x_i)$ ; при этом значения вероятностей, как правило, задаются в табличной форме.

## 2.4. Законы распределения случайных величин

В статистике разработано большое количество математических законов  $f(x)$  для описания случайных величин, имеющих описанные ранее свойства. При исследовании эксплуатационной надежности автомобилей в условиях автотранспортного предприятия (АТП) получили наибольшее распространение следующие законы: экспоненциальный, нормальный, логарифмически нормальный и Вейбулла, которые рекомендуются ГОСТами [10].

**Экспоненциальный закон распределения.** Для экспоненциального распределения справедливо следующее выражение:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x \leq 0 \end{cases} \quad (2.20)$$

где  $\lambda$  — параметр распределения.

Параметр распределения определяется по формуле

$$\lambda = \frac{1}{\bar{X}} = \frac{1}{\sigma(x)}$$

Графическая интерпретация экспоненциального распределения представлена на рис. 2.3.

При исследовании надежности автомобиля  $\lambda$  характеризует интенсивность отказа невосстанавливаемых деталей. Этим распределением описываются распределение безотказной работы деталей автомобиля, распределение времени ремонта автомобиля при восстановлении этих отказов и многое другое. При оценке показателей надежности коэффициент вариации для принятия этого закона должен находиться в диапазоне от 0,8 до 1,2 (большее значение вариации обычно свидетельствует о «смешанной» выборке, о которой упоминалось ранее).

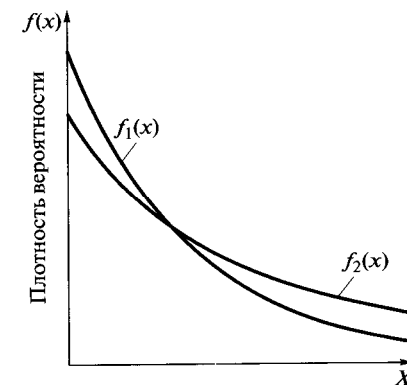


Рис. 2.3. Графическая интерпретация экспоненциального распределения ( $\lambda_1 < \lambda_2$ )

Интегральная функция экспоненциального закона имеет вид

$$F(x) = P(x \leq X) = \int_{-\infty}^x \lambda e^{-\lambda x} dx = 1 - e^{-\lambda x}. \quad (2.21)$$

**Нормальный закон распределения.** Нормальное распределение возникает в том случае, когда результат испытания является следствием влияния большого числа факторов, среди которых нет доминирующих, а объем наблюдений достаточно велик. Это двухпараметрическое распределение (параметры  $\bar{X}$  и  $\sigma$ ).

Плотность вероятностей имеет вид, широко известный как функция Гаусса, а при графической интерпретации — кривая Гаусса:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{X})^2}{2\sigma^2}}. \quad (2.22)$$

Графическая интерпретация нормального распределения представлена на рис. 2.4.

Нормальное распределение применяется для описания отказов, вызванных изнашиванием или постепенным накоплением неисправностей, когда доля внезапных отказов мала, и многих других процессов технической эксплуатации автомобилей (периодичность ТО, расход однородных эксплуатационных материалов, рассеяние значений диагностического параметра для исправного состояния и т.д.). Поскольку данный закон может обрабатывать и отрицательные СВ, его широко применяют во многих областях практической и научной деятельности. При оценке показателей надежности его рекомендуется использовать при коэффициентах вариации, находящихся в диапазоне 0,10...0,35, и значениях коэффициентов асимметрии и эксцесса, близких к нулю.

Закон является симметричным относительно среднего, и для него совпадают значения математического ожидания, моды и медианы. Хотя закон имеет самое большое распространение, его недостатком является невозможность выполнения математических расчетов по выражениям (2.16)...(2.19) из-за отсутствия способов непосредствен-

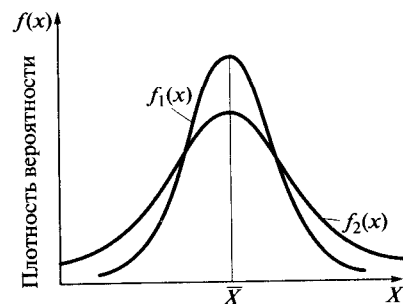


Рис. 2.4. Графическая интерпретация нормального распределения ( $\sigma_1 < \sigma_2$ )

ного интегрирования. Поэтому в статистике для этих целей используется метод, разработанный Лапласом, который ввел следующую подстановку, позволяющую перейти к так называемому центрированному нормированному распределению, имеющему  $\bar{X} = 0$  и  $\sigma = 1$ :

$$t = \frac{x - \bar{X}}{\sigma}. \quad (2.23)$$

Для нормированного распределения интегральная функция будет определяться по выражению

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{t^2}{2}} dt = 0,5 + \Phi(z), \quad (2.24)$$

где  $F(z)$  — функция распределения нормированной СВ, а нормированное значение  $z$ , соответствующее ненормированному значению  $X$ , определяется по выражению (2.23).

В выражении (2.24) показатель вероятности 0,5 в силу симметричности нормированного распределения относительно математического ожидания  $\bar{X} = 0$  будет соответствовать значению интеграла по отрицательному диапазону от  $-\infty$  до нуля, который нецелесообразно вычислять.

Для положительных аргументов соответствующее значение вероятности находится через интеграл, или функцию Лапласа:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{t^2}{2}} dt, \quad (2.25)$$

которая была вычислена ее автором через ряд Тейлора и представлена в виде таблиц. Выдержки из этих таблиц, достаточные для практических расчетов в области технической эксплуатации автомобилей, приведены в Приложении 1.

При использовании функции Лапласа справедливо соотношение

$$\Phi(-z) = -\Phi(z). \quad (2.26)$$

Рассмотрим пример. Фиксировалось время между отказами агрегата автомобиля. В результате были получены следующие статистические характеристики:  $\bar{X} = 36,3$  ч;  $\sigma = 3,3$  ч. Необходимо определить вероятность попадания СВ в интервал  $\bar{X} \pm 5$  ч.

Решение задачи сводится к следующему. Определяется нормированное значение СВ в крайних точках интервала:

$$z_1 = \frac{x_1 - \bar{X}}{\sigma} = \frac{31,335 - 36,335}{3,3} = -1,515;$$

$$z_2 = \frac{x_2 - \bar{X}}{\sigma} = \frac{41,335 - 36,335}{3,3} = 1,515.$$



По Приложению 1 определяются значения функции Лапласа в граничных точках интервала:

$$\Phi(z_1) = \Phi(-1,515) = -0,435;$$

$$\Phi(z_2) = \Phi(1,515) = 0,435.$$

Определяются значения функции распределения в граничных точках интервала:

$$F(x_1) = F(31,335) = 0,5 + \Phi(z_1) = 0,5 - 0,435 = 0,065;$$

$$F(x_2) = F(41,335) = 0,5 + \Phi(z_2) = 0,5 + 0,435 = 0,935.$$

Определяется вероятность попадания СВ в интервал  $\bar{X} \pm 5$ :

$$P\{x_1 < X < x_2\} = P\{31,335 < X < 41,335\} = F(x_2) - F(x_1) = \\ = F(41,335) - F(31,335) = 0,935 - 0,065 = 0,870.$$

Следует учитывать, что в ряде случаев могут быть даны другие табличные значения функции Лапласа, зависящие от задания пределов интегрирования, на что необходимо обращать внимание и проводить соответствующие корректировки.

На практике в первом приближении можно использовать так называемое свойство шестисигмового диапазона нормального закона, графически представленного на рис. 2.5, и показывающие вероятности попадания (в процентах) случайной величины в различные диапазоны стандартного отклонения  $\sigma$  относительно математического ожидания. Поскольку в указанном диапазоне, как это можно видеть из Приложения 1, практически находится 100 % случайной величины, за его пределами производить расчеты нецелесообразно.

В то же время диапазоны с 85 и 95 % вероятностями могут просто рассчитываться с незначительной ошибкой прибавлением к среднему значению одно- или полутора кратных величин  $\sigma$  (при нижнем пределе интегрирования  $-\infty$ ). Также при двухсторонних ограничениях диапазон  $\bar{X} \pm 2\sigma$  будет соответствовать популярной в статистике 95%-ной вероятности.

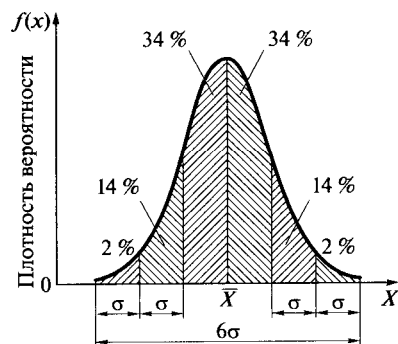


Рис. 2.5. Значения вероятностей для нормального закона распределения по диапазонам стандартных отклонений

При компьютерных расчетах значение интегральной функции Лапласа для снижения времени также можно определять по табличным данным, помещая их в программу, или же вычислять непосредственно приближенным графическим способом.

**Логарифмически нормальный закон распределения.** Непрерывная СВ считается распределенной по логарифмически нормальному закону, если ее логарифм распределен по нормальному закону. Данное распределение встречается при обработке результатов ускоренных испытаний отдельных объектов. Ему может подчиниться наработка до отказа невозстанавливаемого изделия при его усталостном разрушении, размер АТП, периодичность отказов крепежных соединений и т.д. При оценке показателей надежности его можно использовать при значениях коэффициента вариации в диапазоне 0,4...0,8.

Плотность вероятности имеет вид

$$f(x) = \frac{1}{\sigma_n \sqrt{2\pi}} e^{-\frac{(\ln x - y_0)^2}{2\sigma_n^2}}, \quad (2.27)$$

где  $\sigma_n$  — среднеквадратическое отклонение логарифма СВ;  $y_0$  — математическое ожидание логарифма СВ.

Математическое ожидание и среднеквадратическое отклонение логарифма СВ определяются по следующим формулам:

$$y_0 = \frac{1}{N} \sum_{i=1}^N \ln x_i; \quad \sigma_n = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\ln x_i - y_0)^2}. \quad (2.28)$$

Формы логарифмически нормального распределения могут быть самыми различными: редко — от почти симметричного относительно математического ожидания, а в основном — с ярко выраженной положительной асимметрией, когда значение моды существенно ниже значения медианы (графически приближается к началу координат (рис. 2.6)).

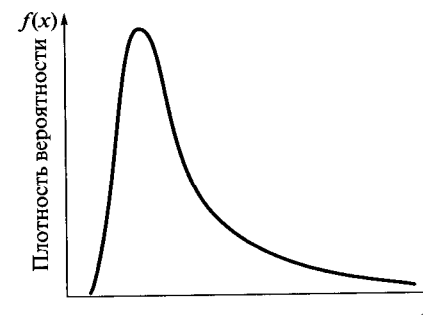


Рис. 2.6. Форма логарифмически нормального распределения

Хотя логарифмически нормальное распределение рекомендуется для применения при оценке показателей надежности нормативной документацией [10], «ручные» расчеты по нему, которые всегда целесообразно проводить для контроля качества неизвестных компьютерных программ, значительно сложнее по сравнению с нормальным законом. По этой причине в нормативных рекомендациях [10] практически «свернуты» основные статистические проработки для этого закона (за исключением определения объема выборки), поэтому целесообразно пользоваться законом Вейбулла, который по характеру распределения может заменить логарифмически нормальный закон в том же диапазоне значений коэффициентов вариации СВ.

Более важной причиной, по которой следует избегать использования логарифмически нормального закона, является теоретическая «ненормальность» изменения интенсивности отказов для него с начальным повышением и последующим снижением. Последнее противоречит физической природе процессов постепенного изнашивания, или старения, автомобильных элементов, оценка показателей надежности которых производится по экспериментальным данным.

**Закон распределения Вейбулла.** Это наиболее распространенный закон распределения при обработке экспериментальных данных по надежности автомобилей в эксплуатации; теоретически начинается с нуля и обеспечивает соответствие экспериментальным данным в диапазоне коэффициентов вариации 0,4...0,8.

Плотность распределения согласно [10] имеет вид (возможны и другие выражения)

$$f(x) = \frac{b}{a} \left(\frac{x}{a}\right)^{b-1} e^{-\left(\frac{x}{a}\right)^b} \quad (2.29)$$

при условии, что  $a$  — параметр масштаба, косвенно связанный со средним значением;  $b$  — параметр формы, косвенно связанный со значением коэффициента вариации (при  $b = 1$  закон Вейбулла вырождается в экспоненциальный).

Формы распределения закона Вейбулла представлены на рис. 2.7.

Оценка параметров закона распределения может осуществляться различными методами (максимального правдоподобия, приближенным и т. д.). Нормативной документацией [10] рекомендуется реализуемый только на ЭВМ метод последовательного приближения при статистической оценке параметров  $a$  и  $b$  по экспериментальным данным, однако при разработке программ следует учитывать редкую возможность закливания процедуры приближения при повторении определяемых значений оценок и предусматривать выход из циклового расчета с рекомендациями о корректировке исходных данных.

При «ручных» расчетах, которые зачастую необходимы, поскольку не все компьютерные программы обеспечивают 100%-ные воз-

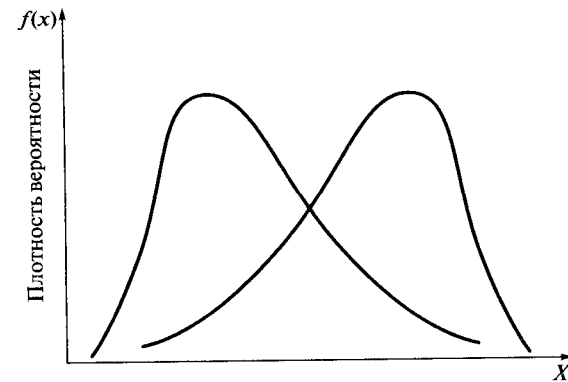


Рис. 2.7. Формы распределения закона Вейбулла

можности использования закона Вейбулла (в том числе, STATISTI-С^A), на практике получила применение следующая процедура.

1. Определяются коэффициент асимметрии и коэффициент вариации:

$$A_k = \frac{\frac{N}{(N-1)(N-2)} \sum_{i=1}^N (x_i - \bar{X})^3}{\left[ \frac{1}{N} \sum_{i=1}^N (x_i - \bar{X})^2 \right]^{1.5}} \quad (2.30)$$

2. По полученному значению из табл. 2.3 находят оценку  $\hat{b}$  параметра  $b$  и значения коэффициентов  $K_b$  и  $g_b$ . Затем определяются оценка  $\hat{a}$  параметра  $a$ :  $\hat{a} = \frac{\sigma}{g_b}$ ; и оценка  $\hat{c}$  параметра  $c$ :  $\hat{c} = \bar{X} - \hat{a}K_b$ .

Таблица 2.3

**Коэффициенты и параметры закона распределения Вейбулла**

Коэффициент вариации $V$	Коэффициент асимметрии $A_k$	$b$	$K_b$	$g_b$
0,120	-0,638	10,0	0,951	0,114
0,170	-0,455	6,9	0,935	0,159
0,20	-0,352	5,8	0,926	0,184
0,301	-0,023	3,7	0,901	0,276
0,399	0,275	2,7	0,889	0,355

Коэффициент вариации $V$	Коэффициент асимметрии $A_k$	$b$	$K_b$	$g_b$
0,500	0,567	2,1	0,886	0,443
0,605	0,865	1,7	0,892	0,540

В качестве оценки параметра  $c$  берется одно из двух значений:

- $\hat{c}$ , если оно меньше  $x_1 = x_{\min}$ , где  $x_{\min}$  — минимальное значение вариационного ряда;

- $x_1 = x_{\min}$ , если  $\hat{c}$  больше  $x_{\min}$ .

Математическое ожидание и среднее квадратическое отклонение находятся по формулам:

$$\bar{X} = aK_b + c; \quad \sigma = ag_b.$$

Удобство закона Вейбулла заключается, с одной стороны, в возможности его применения (как это можно видеть из данных табл. 2.2 и рис. 2.7) для описания случайных величин как с положительной, так и с отрицательной асимметрией, а с другой — в простоте вычисления функции распределения, поскольку выражение (2.29) легко интегрируется:

$$F(x) = 1 - e^{-\left(\frac{x}{a}\right)^b}. \quad (2.31)$$

**Гамма-распределение.** Гамма-распределение часто необходимо при описании закономерностей рассеяния диагностических параметров, поскольку теоретически оно может начинаться, в отличие от других законов, с положительного фиксированного значения  $X_0$ , называемого параметром смещения, распределяясь затем в сторону увеличения. Это соответствует физике явления многих процессов технической эксплуатации, например удельному расходу автомобилями топлива, который не может быть ниже определенной экономической величины, и др.

Плотность распределения при этом описывается выражением

$$f(x) = \frac{1}{\Gamma(r)} \lambda^2 (x - X_0)^{r-1} e^{-\lambda(x-X_0)}, \quad x \geq X_0, \quad (2.32)$$

в котором

$$\lambda = \frac{\bar{X} - X_0}{\sigma^2}; \quad r = \frac{(\bar{X} - X_0)^2}{\sigma^2}, \quad (2.33)$$

а

$$\Gamma(r) = \int_0^{\infty} u^{r-1} e^{-u} du, \quad r > 0 \quad (2.34)$$

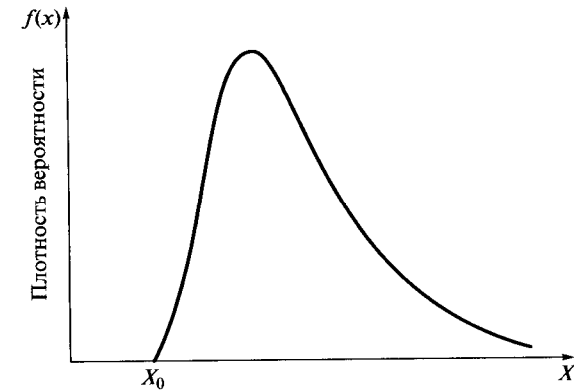


Рис. 2.8. Графическая интерпретация гамма-распределения с параметром смещения  $X_0$

носит название гамма-функции (для «ручных» расчетов она табулирована).

Обычно гамма-распределение имеет положительную асимметрию (рис. 2.8), а при  $r = 1$  (и  $X_0 = 0$ ) вырождается в экспоненциальный закон.

При компьютерных расчетах интегральная функция может определяться приближенным графическим способом (по методу прямоугольников), хотя это и приводит к увеличению длительности вычислений. Хотя программа STATISTICA предусматривает «подгонку» случайных величин под гамма-распределение (в отличие от закона Вейбулла), она не показывает значений его параметров, которые при необходимости можно вычислить вручную, используя выражения (2.33) и в дальнейшем правильно вводя результаты в вероятностный калькулятор. Вероятностный калькулятор позволяет быстро ознакомиться со многими статистическими законами и закономерностями их изменения в зависимости от задаваемых параметров, что незаметно при обучении, а также получать любые статистические значения с высокой точностью, что зачастую необходимо даже при наличии специальных таблиц, которыми удобнее пользоваться.

## 2.5. Проверка соответствия закона распределения эмпирическим данным на основе критерия Пирсона

Поскольку на основе экспериментальных данных можно использовать несколько теоретических распределений, первой из решаемых статистикой задач является выбор наилучшего закона, или «подгонка» распределения [4]. Для этого могут использоваться различные критерии, из которых наиболее употребительным и строгим [10]

является критерий согласия Пирсона. Он основан, как и другие критерии, на изучении меры расхождения между статистическим и теоретическим распределением, которая в данном случае оценивается по сумме квадратов разности этих расхождений по всем интервалам разбиений выборки СВ с учетом их «весов». К. Пирсон показал, что указанная сумма не зависит от функции распределения  $F(x)$  и числа опытов  $N$ , а зависит от числа разрядов  $k$ ; при увеличении  $N$  этот закон приближается к так называемому распределению  $\chi^2$  (хи-квадрат).

Мера расхождения при статистических оценках обычно определяется по выражению [4, 10]

$$\chi^2 = \sum_{j=1}^k \frac{(n_j^* - Np_j)^2}{Np_j}, \quad (2.35)$$

где  $p_j$  — теоретические вероятности (частоты) попадания СВ в заданные интервалы, зависящие от выбранного закона распределения и определяемые через его интегральную функцию (процедура расчета приведена далее).

Аналитическое выражение для плотности распределения *хи-квадрат*, которое чаще называют распределением Пирсона, приведено в Приложении 2 с табличными данными значений интегральной функции, взятых из нормативных документов [10], по которым в основном проводят практические расчеты. Как видно, распределение Пирсона зависит от параметра  $r$ , называемого числом степеней свободы. В данном случае оно будет равно числу независимых разрядов (интервалов разбиения выборки)  $k$  минус число независимых условий (связей), наложенных на эмпирические частоты  $p_j^*$ .

Таковыми связями для любого закона распределения будут являться условия (2.12), (2.13) и (2.14), по которым производится статистическая оценка основных параметров [10], т.е. при подборе распределения  $r = k - 3$ . Последнее значение является первым входом в таблицы Приложения 2; затем по значению величины расхождения, рассчитанной по выражению (2.35), определяется вероятность согласия теоретического закона с экспериментальными данными. Естественно, что при увеличении расхождения вероятность согласия снижается.

Для последующих практических расчетов может быть принят только теоретический закон, значение вероятности согласия по которому окажется не ниже 0,3. Соответственно при конкурирующих законах целесообразно принимать закон, имеющий наибольшую вероятность согласия.

На рис. 2.9 приведены графические формы  $\chi^2$ -распределения при различных степенях свободы, которые необходимо учитывать при его использовании для других целей (например, при оценке дисперсии нормальной выборки). При значениях  $r = 3$  и выше дифферен-

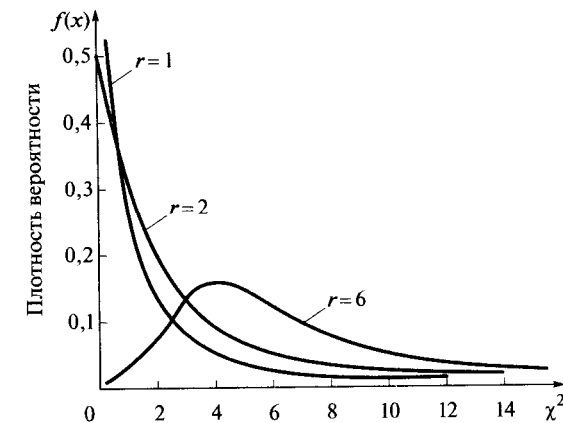


Рис. 2.9. Графическая интерпретация  $\chi^2$ -распределения при различных значениях степеней свободы

циальная функция уже представляет собой одномодальную асимметричную кривую. Также необходимо обращать внимание на пределы интегрирования, по которым составляются табличные данные, поскольку при этом меняются значения вероятностей. В нашем случае взяты пределы интегрирования, рекомендуемые в учебном пособии [2, 3], которые наилучшим образом отражают физическую сущность подбора распределения по критерию согласия.

Для примера рассмотрим проверку на нормальность закона распределения коэффициента регулярности движения автобусов, заданного случайными величинами табл. 2.1. При этом целесообразно продолжить расчеты (см. табл. 2.2), уже отражающие эмпирические данные  $n_j^*$  и соответствующие им частоты  $p_j^*$  попадания СВ в независимые интервалы, а также статистические оценки для среднего значения и дисперсии.

Дальнейшие расчеты в соответствии с выражением (2.35) при условии принятия нормального закона приведены в столбцах 7... 11 табл. 2.2 с расчетными формулами.

Так, в столбце 7 показаны нормированные отклонения функции Лапласа для границ интервалов; в столбце 8 — значения интегральной функции для границ интервалов, найденных при помощи интеграла Лапласа; в столбце 9 — теоретические вероятности попадания СВ в данные интервалы; в столбце 10 — произведение величин  $Np_j$ , отражающее теоретическое количество  $n_j$  попадания СВ в данный интервал (последние удобно использовать для графических построений функции, как это показано для нашего случая на рис. 2.1). В столбце 11 приведены значения несоответствия между теоретическим законом и экспериментальными данными для каждого интервала и их сумма, которая в данном случае равна 3,13.

По Приложению 2 находим, что при степени свободы  $r = 7 - 3 = 4$  такое значение будет соответствовать вероятности согласия не менее чем 0,5, что выше принятого для практики критического значения 0,3. Поэтому гипотеза о нормальности закона распределения коэффициента регулярности движения не отвергается.

Проверка нормальности распределения может проводиться с помощью коэффициентов асимметрии и эксцесса. При этом реализуется следующая процедура.

Вычисляются значения коэффициента асимметрии

$$A_k = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{X})^3}{\left[ \frac{1}{N} \sum_{i=1}^N (x_i - \bar{X})^2 \right]^{3/2}},$$

эксцесса

$$E_k = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{X})^4}{\left[ \frac{1}{N} \sum_{i=1}^N (x_i - \bar{X})^2 \right]^2}$$

и их стандартных отклонений:

$$S(A_k) = \sqrt{\frac{6(N-1)}{(N+1)(N+3)}}; \quad S(E_k) = \sqrt{\frac{24N(N-2)(N-3)}{(N+1)^2(N+3)(N+5)}}$$

(указанные расчеты, как отмечалось ранее, автоматически выполняются программой STATISTICA). На основе полученных данных проверяется гипотеза о нормальности распределения. Если выполняется условие  $|A_k| < 1,5S(A_k)$  и  $|E_k| < 1,5S(E_k)$ , то гипотеза принимается; если  $|A_k| \geq 2S(A_k)$  и  $\left| E_k + \frac{6}{N+1} \right| \geq 2S(E_k)$ , то гипотеза отвергается.

Если  $1,5S(A_k) < |A_k| < 2S(A_k)$ ;

$$1,5S(E_k) < \left| E_k + \frac{6}{1+N} \right| < 2S(E_k),$$

то применяются другие гипотезы.

Использование критерия  $\chi^2$  (хи-квадрат, или Chi-Square) с соответствующей ему вероятностью согласия (соответствия) является

стандартной процедурой программы STATISTICA при «подгонке» законов распределения, предусматривающей графическое построение гистограмм и теоретического распределения. Однако существенным недостатком данной программы при решении задач в области технической эксплуатации автомобилей является, с одной стороны, отсутствие для анализа закона Вейбулла, а с другой — несоблюдение или даже невыполнение строгих научных рекомендаций по использованию критерия Пирсона [10]. Поэтому необходимо использовать специализированные программы для обработки распределений, не имеющих указанных недостатков.

В частности, программа, разработанная на кафедре «Эксплуатация автомобильного транспорта и автосервис» МАДИ, предусматривает обработку для полных статистических выборок по экспоненциальному, нормальному, гамма-распределению и закону Вейбулла. При этом на экран монитора выводятся данные о параметрах распределения, среднему значению и стандартному отклонению, коэффициенту вариации (последний не рассчитывается программой STATISTICA), показывается гистограмма и выбираемый для анализа теоретический закон распределения, эмпирические  $n_j^*$  и теоретические  $n_j$  значения попадания данных в рассчитываемые по формуле Стьеджера (2.10) интервалы разбиения, значение оценки  $\chi^2$  и соответствующей этой оценки вероятности согласия.

Кроме того, расширены возможности обработки указанными распределениями для любых случайных величин, встречающихся в практике технической эксплуатации автомобилей, в частности для диагностических параметров, которые могут иметь отрицательные значения, не обрабатываемые программой STATISTICA. В этом случае программой предусматривается преобразование исходных данных путем их суммирования со специально определяемой величиной, при



Рис. 2.10. Графическая интерпретация преобразованного гамма-распределения с отрицательной асимметрией

которой все они становятся положительными и могут быть обработаны по экспоненциальному закону, закону Вейбулла и гамма-распределению, а на экран выводятся графические изображения, соответствующие исходному (отрицательному) диапазону чисел.

Для подобных СВ коэффициент вариации, который введен для удобства оценки показателей надежности, может принимать любое значение. Также за счет преобразования путем вычитания выборки из наибольшего значения СВ можно искусственно получить гамма-распределение с отрицательной асимметрией (рис. 2.10), удобное при обработке диагностических параметров, например значений силы тяги на колесах или тормозных сил, которые теоретически не могут превышать определенных значений. Таким образом, отмеченная ранее специфика программы практически обеспечивает решение любых задач технической эксплуатации, связанных с обработкой распределений, хотя при этом целесообразно использовать дополнительные возможности программы STATISTICA.

## 2.6. Понятие доверительного интервала и доверительной вероятности при статистической оценке характеристик рассеяния случайных величин

Как уже отмечалось ранее, при обработке выборки эмпирических данных случайных величин о надежности автомобилей или других показателей их работы основные числовые характеристики, прежде всего значения средней и дисперсии, определяются с некоторыми ошибками  $\epsilon$ , величины которых будут зависеть, в первую очередь, от объема (представительности) выборки  $N$ , а также от вида закона распределения. Конкретное представление о величинах этих ошибок и, таким образом, о точности и надежности статистических оценок получают на основе дополнительных расчетов так называемого доверительного интервала для несмещенной оценки показателя  $\tilde{a}$ , задаваемого с достаточным уровнем доверительной вероятности  $\beta^*$ , как правило, от 0,8 и выше [10].

При этом условии для оценки истинного значения  $a$  запишется в виде

$$P(\tilde{a} - \epsilon < a < \tilde{a} + \epsilon) = \beta^*. \quad (2.36)$$

Это условие означает, что с вероятностью  $\beta^*$  неизвестное значение показателя  $a$  попадает в интервал

$$I_{\beta^*} = (\tilde{a} - \epsilon; \tilde{a} + \epsilon). \quad (2.37)$$

(Знаком «\*» здесь обозначена так называемая двухсторонняя доверительная вероятность в отличие от односторонней (также исполь-

зуемой в статистике), которая в нормативной документации [10] обозначается без знака «\*»).

Поскольку здесь величина  $a$  не является случайной (случайными являются интервал  $I_{\beta^*}$ , его положение на оси абсцисс определяется его центром  $\tilde{a}$ , случайна и сама длина интервала  $2\epsilon$ , поскольку величина  $\epsilon$  вычисляется по опытным данным), то величину  $\beta^*$  принято толковать как вероятность того, что случайный интервал  $I_{\beta^*}$  накроет точку  $a$ .

Границы интервала называются нижними и верхними доверительными границами. Сложность задачи определения указанных границ заключается в том, что закон распределения оценки  $\tilde{a}$  зависит от закона распределения СВ и его неизвестных параметров, в частности от самого параметра  $a$ . Наиболее достоверно данный процесс теоретически разработан в статистике для нормального закона распределения. Прежде всего это относится к оценкам математического ожидания и дисперсии, а также к другим показателям.

Сущность теоретических представлений заключается в следующем. Согласно центральной предельной теореме [4] случайная величина — средняя арифметическая  $\bar{X}^*$  любой большой выборки  $N$  случайной величины  $X$  — будет распределена асимптотически нормально с параметрами

$$\bar{X} = M\bar{X}^* \text{ и } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}}, \quad (2.38)$$

т.е. ее центр совпадает с теоретическим математическим ожиданием СВ, а дисперсия в  $N$  раз меньше теоретического значения дисперсии случайной величины  $X$  (корень квадратный из этой дисперсии, представленный в выражении (2.38), носит название стандартной ошибки [4]). В первом приближении указанное свойство нормальности можно использовать для определения верхних и нижних границ доверительного интервала для центра распределения (по схеме, представленной рис. 2.5), непосредственно используя значение функции Лапласа, как это автоматически производится программой STATISTICA для двухсторонней доверительной вероятности 0,95.

Доказано [4, 10], что при нормальном распределении величины  $\tilde{a}$  случайная величина

$$T = \sqrt{N} \frac{\bar{X}^* - \bar{X}}{\sqrt{D^*}} \quad (2.39)$$

(где величины  $\bar{X}^*$  и  $D^*$  представляют собой несмещенные статистические оценки средней и дисперсии, рассчитанные по выражениям (2.1) и (2.3)), подчиняется так называемому закону распределения Стьюдента с  $(N - 1)$  степенями свободы.

Плотность распределения этого закона имеет вид

$$S_k(t) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{\pi k}} \left(1 + \frac{t^2}{k}\right)^{-\frac{k+1}{2}}, \quad (2.40)$$

где  $k$  — число степеней свободы;  $\Gamma(x)$  — известная гамма-функция.

Из выражения (2.40) видно, что  $S_k(t)$  — четная функция, графически симметричная относительно оси ординат (напоминает в этом отношении нормированный нормальный закон), однако ее размах зависит от числа степеней свободы и при малых выборках заметно шире нормального закона, как это можно видеть из рис. 2.11.

Хотя при больших  $k$  распределение Стьюдента практически совпадает с нормированным нормальным законом, в настоящее время считается целесообразным проводить именно по нему расчеты верхней и нижней границ доверительного интервала для среднего [10] (по аналогии, как при использовании нормального закона), с тем чтобы избежать завышенных оценок. При этом значения математического ожидания и дисперсии задаются согласно выражению (2.38) при определенном уровне доверительной вероятности. Однако здесь возможно получить разные ответы в зависимости от одностороннего или двухстороннего способа задания доверительной вероятности, как это можно видеть из рис. 2.12, 2.13.

Наиболее жесткие условия (т.е. дающие увеличенные значения диапазона доверительного интервала) получаются при задании двухстороннего способа доверительной вероятности  $\beta^*$  из условия, что параметр  $T$  распределения Стьюдента по абсолютному значению не превысит определенной величины  $t_{\beta^*}$ :

$$P(|T| < t_{\beta^*}) = \int_{-t_{\beta^*}}^{t_{\beta^*}} S_k(t) dt = \beta^*, \quad (2.41)$$

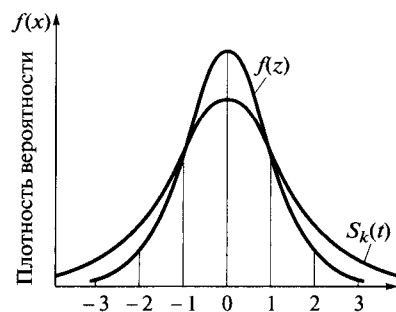
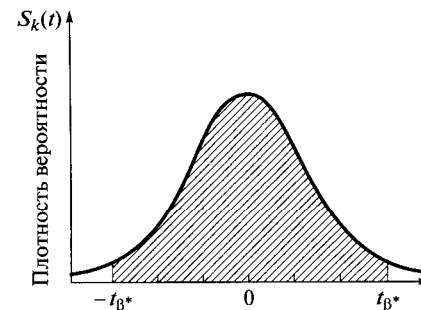


Рис. 2.11. Графическая интерпретация распределения Стьюдента в сравнении с нормированным нормальным законом

Рис. 2.12. Схема определения доверительных границ для средней величины при двухстороннем способе задания доверительной вероятности по распределению Стьюдента



или с учетом симметричности (более распространенное выражение):

$$P(|T| < t_{\beta^*}) = 2 \int_0^{t_{\beta^*}} S_k(t) dt = \beta^*. \quad (2.42)$$

Для величин  $t_{\beta^*}$ , которые обычно называют значением  $t$ -критерия или критерием Стьюдента, составлены специальные таблицы для практического использования при решении нескольких важных для статистики задач. Так, для условия (2.42) нижние и верхние значения границ доверительного интервала для центра распределения  $\bar{X}$  согласно схеме, представленной на рис. 2.12, а также согласно выражениям (2.36) и (2.39) будут находиться по выражению

$$P(\bar{X}^* - t_{\beta^*, N-1} \frac{S(x)}{\sqrt{N}} < \bar{X} < \bar{X}^* + t_{\beta^*, N-1} \frac{S(x)}{\sqrt{N}}) = \beta^*. \quad (2.43)$$

При одностороннем способе задания доверительной вероятности, как это предписывается нормативной документацией [10], значение  $t_{\beta}$  в выражении (2.43) согласно схеме, представленной на рис. 2.13, определяется из условия

$$P(T < t_{\beta}) = \int_{-\infty}^{t_{\beta}} S_k(t) dt = \beta, \quad (2.44)$$

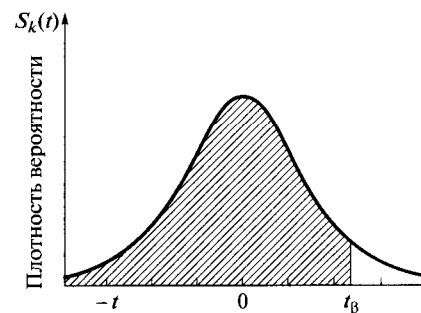


Рис. 2.13. Схема определения доверительных границ для средней величины при одностороннем способе задания доверительной вероятности по распределению Стьюдента

которое будет давать, как это видно из сопоставления со схемой, представленной на рис. 2.12, меньшие значения доверительного интервала и завышенную точность оценки для среднего.

Значение  $t_{\beta}$ , определяемое из условия (2.44), называется *квантилем*, под которым понимается односторонний правый доверительный интервал любого теоретического закона распределения и который широко используется при оценке точности статистических показателей. В частности, таблицы квантилей для распределения Стьюдента приводятся в нормативной документации [10] и в незначительно сокращенном виде приведены в Приложении 3. Могут быть и другие варианты представления таблиц для распределения Стьюдента; для правильного их применения необходимо обращать внимание на пределы интегрирования, отражающие схемы его использования. Так, важное значение имеет так называемый уровень значимости  $q$ , обычно определяемый в процентах (в программе STATISTICA он обозначается буквой  $p$  и выражается в десятичных долях единицы), который определяется только для случая двухстороннего способа задания доверительной вероятности по выражениям (2.41) или (2.42) соотношением [4]

$$q = 1 - \beta^* = 2 \int_{t_{q^*}}^{\infty} S_k(t) dt. \quad (2.45)$$

Для исключения ошибок целесообразно объединить три основных способа использования таблиц Стьюдента (по выражениям (2.42), (2.44), (2.45)) в один через дополнительные ключи входа, как это сделано в Приложении 3. Хотя программой STATISTICA расчет 95%-ных границ доверительного интервала производится по нормальному закону, оценка уровня значимости осуществляется на основе распределения Стьюдента при двухстороннем способе задания доверительной вероятности.

Удобной для практики специфической особенностью программы STATISTICA является графическое представление расчетов по границам доверительного интервала в виде «ящика с усами» различного вида, зависящего от того, какая информация отображается. Для примера на рис. 2.14 показан наиболее популярный вариант представления для двух выборок случайных величин, когда среднее значение отображено выделенной точкой, длина ящика равна двойной величине стандартной ошибки, а усы показывают 95%-ные двухсторонние доверительные границы, вычисленные по нормальному закону. В первом приближении здесь можно сделать часто необходимое для практики заключение о равенстве или различии средних, однако в случае перекрытия границ доверительных интервалов более надежную оценку можно получить на основе так называемого  $t$ -теста, определяющего уровень значимости различия на основе распределения Стьюдента.

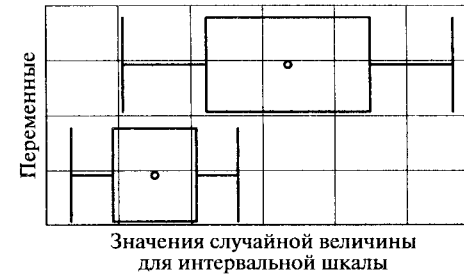


Рис. 2.14. Представление результатов обработки случайных величин в виде «ящиков с усами» в программе STATISTICA

Реже на практике выполняется построение доверительного интервала для дисперсии  $D(x)$  и среднеквадратического отклонения  $\sigma(x)$  по величинам их статистических оценок  $S^2(x)$  и  $S(x)$  (программой STATISTICA автоматического выполнения подобной операции не предусматривается).

Доказано [4, 10], что для нормальной совокупности случайная величина

$$V = \frac{NS^2}{\sigma^2} \quad (2.46)$$

распределена по закону  $\chi^2$  с  $(N - 1)$  степенями свободы.

При этом для заданного уровня доверительной вероятности  $\beta$ , который можно выразить через уровень значимости  $q$  соотношением  $\beta = 1 - q$ , можно построить бесконечное количество интервалов, которым отвечала бы вероятность, в точности равная  $\beta$ . Однако в статистике в силу несимметричности распределения  $\chi^2$  рекомендуется выбирать доверительный интервал для дисперсии таким образом, чтобы вероятности выхода величины  $V$  за его пределы вправо и влево (заштрихованные области на рис. 2.15) были одинаковы и равны, т. е.

$$\frac{q}{2} = \frac{1 - \beta^*}{2}. \quad (2.47)$$

Условие для нахождения доверительного интервала при этом будет задано соотношением

$$P\left(\chi_1^2 < \frac{NS^2}{\sigma^2} < \chi_2^2\right) = 1 - P(\chi^2 < \chi_1^2) - P(\chi^2 > \chi_2^2) = \beta^*, \quad (2.48)$$

из которого следует:

$$P\left(\frac{NS^2}{\chi_2^2} < \sigma^2 < \frac{NS^2}{\chi_1^2}\right) = \beta^* = 1 - q. \quad (2.49)$$



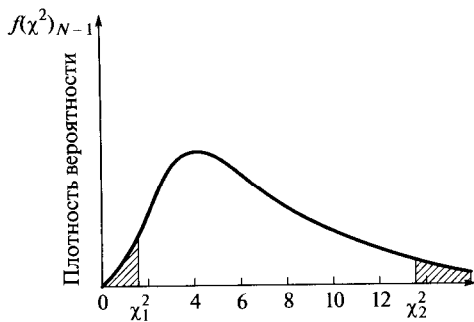


Рис. 2.15. Построение доверительного интервала для оценки  $\sigma^2$  по эмпирической дисперсии  $S^2$

Отсюда для практических расчетов границ доверительного интервала для  $\sigma$  необходимо сначала по таблицам Приложения 3 найти при числе степеней свободы  $(N - 1)$  значения  $\chi_2^2$  и  $\chi_1^2$  для  $P_2 = \frac{q}{2}$  и  $P_1 = 1 - \frac{q}{2}$ , а после их подстановки в выражение (2.49) и соответствующих вычислений извлечь из полученных результатов квадратный корень. Однако при этом необходимые данные для больших значений  $N$  могут отсутствовать. В последнем случае не следует интерполировать данные Приложения 2 для избежания ошибок, а следует взять из них ближайшие значения (по возможности для меньшего числа степеней свободы), изменив соответствующим образом величину  $N$  в выражении (2.49), или по возможности вычислить необходимые значения при помощи вероятностного калькулятора программы STATISTICA.

**Пример.** На выборке из  $N = 136$  значений СВ произведенные оценки для дисперсии и стандартного отклонения составили:  $S^2(x) = 278,9$  и  $S(x) = 16,7$ . Требуется определить при доверительной вероятности  $\beta^* = 0,95$  доверительные интервалы для дисперсии  $D(x)$  и стандартного отклонения  $\sigma(x)$ .

**Решение.** Имеем:  $q = 1 - \beta = 0,05$ ;  $q/2 = 0,025$ . По Приложению 2 для ближайшей степени свободы  $r = 130$  при  $P_2 = 0,025$  значение  $\chi_2^2 = 167,5$  и при  $P_1 = 1 - q/2 = 0,975$  значение  $\chi_1^2 = 100,3$ . Неравенство из условия (2.49) для данного случая запишется в виде

$$\frac{131 \cdot 278,9}{167,5} < \sigma^2 < \frac{131 \cdot 278,9}{100,3}.$$

После преобразований получаем:

$$218,6 < \sigma^2 < 364,3, \text{ или } 14,8 < \sigma < 19,1.$$

Получаемые на основе доверительных интервалов оценки точности определения математического ожидания и стандартного отклонения позволяют при необходимости задаваться наиболее жесткими условиями для ряда практических расчетов, например принимая значения математического ожидания по нижней границе и стандартного отклонения — по верхней, и наоборот.

**Для экспоненциального закона распределения.** Согласно нормативным документам [10] при исследованиях надежности для математического ожидания для данного распределения рекомендуется следующий метод оценки доверительных границ при задании доверительной вероятности  $\beta^*$  или уровня значимости  $q = 1 - \beta^*$ :

$$\frac{2(N-1)\bar{X}^*}{\chi_{\frac{q}{2}}^2(2N)} < \bar{X} < \frac{2(N-1)\bar{X}^*}{\chi_{1-\frac{q}{2}}^2(2N)}, \quad (2.50)$$

где соответствующие значения  $\chi^2(2N)$  при степенях свободы  $2N$  и вероятностях  $P_2 = \frac{q}{2}$  и  $P_1 = 1 - \frac{q}{2}$  определяются по Приложению 2 (для распределения Пирсона).

**Для логарифмически нормального закона распределения.** В простых случаях производится перевод исходных данных СВ в новый вариационный ряд значений ее натуральных логарифмов  $\ln(X_i)$ , по которому уже затем выполняются все необходимые расчеты на основе методов, принятых для нормального закона. В ответственных случаях, когда производятся испытания на надежность, необходимо пользоваться сложными методами, рекомендуемыми нормативными документами [10], в соответствии с которыми определяются нижние и верхние границы доверительных оценок основных показателей надежности: средней наработки до отказа, гамма-процентной наработки до отказа, гамма-процентного ресурса. Эти показатели должны быть связаны с исходным вариационным рядом, хотя расчет по ним, по всей видимости, должен проводиться через преобразованный ряд по нормальному закону с последующим возвратом к исходным показателям.

**Для распределения Вейбулла.** Поскольку распределение Вейбулла является основным законом при исследовании надежности, нормативными документами [10] для него даются все необходимые статистические разработки, связанные с оценкой доверительных интервалов, включающие в себя достаточно сложные расчеты по формулам при одновременном использовании специально разработанных таблиц и графиков. Однако это касается только основных показателей надежности, упомянутых ранее, а непосредственно для самих параметров закона Вейбулла ( $a$  и  $b$ ) рекомендаций по оценке доверительных интервалов не дается.

Очевидно, что даже при возможности получения таких оценок за основу при исследованиях надежности должны быть взяты разработ-

ки нормативных документов [10]. Однако в других случаях, например при разработке имитационных моделей, где закон Вейбулла широко используется, могут потребоваться скорректированные в более жесткую сторону значения его параметров  $a$  и  $b$ . Получить эти значения можно исходя из изложенных ранее возможностей компьютерных программ (STATISTICA) на основе того обстоятельства [4], что отсутствие нормальности не приводит к фатальным последствиям, а при увеличении размера выборки, используемой для получения распределения выборочного среднего, это распределение согласно центральной предельной теореме приближается к нормальному, а при  $N = 30$  оно уже почти нормально. Поэтому для нахождения скорректированных параметров  $a$  и  $b$  на первом этапе необходимо на основе нормального закона получить скорректированные оценки для средней и дисперсии, а затем по изложенной в подразд. 2.4 методике, обобщенной табл. 2.2, найти новые значения параметров  $a$  и  $b$ .

### 2.7. Определение объема выборки и организация наблюдений за автомобилями при изучении показателей их работы в эксплуатации

Одной из основных задач, решаемых исследователем при проведении пассивного эксперимента, является определение объема выборки  $N$ , который устанавливается с учетом следующих условий:

- экономичность (объем выборки должен быть минимальным);
- получение достоверных данных об изучаемом явлении.

При этом выборка представляет собой совокупность элементов  $x_1, x_2, \dots, x_N$ , взятых случайным образом из генеральной совокупности. Под *генеральной совокупностью* следует понимать конечное или бесконечное множество элементов, которые объединены между собой качественным или количественным аргументом.

Объем выборки всегда меньше объема генеральной совокупности, поэтому оценки, получаемые в результате анализа выборочных совокупностей, как было отмечено ранее, всегда имеют какую-то погрешность. С учетом этой погрешности и определяется объем выборки. Кроме того, может быть задан и закон распределения.

Нормативные документы, связанные с оценкой показателей надежности [10], дают подробные рекомендации по выбору числа объектов, в первую очередь, для наиболее сложных случаев так называемых усеченных испытаний, сокращающих продолжительность и стоимость исследований, в зависимости от того, какие показатели и с какой достоверностью должны быть получены. Эти рекомендации не должны нарушаться. Для наиболее простых случаев полных испытаний, которые не обязательно могут быть связаны с надежностью (например, продолжительность простоя автомобилей в ремонте, обусловленных организационными факторами; распределение зна-

чений диагностических параметров), необходимый объем выборки может быть получен на основе изложенных в подразд. 2.6 понятий о доверительном интервале и доверительной вероятности.

В общем случае методы расчета необходимого объема выборки могут быть:

- непараметрическими (когда вид закона распределения неизвестен);
- параметрическими (когда вид закона распределения известен).

При *непараметрическом методе* объем выборки для оценки неизвестного закона надежности определяется по формуле

$$N = \frac{\ln(1-\beta)}{\ln P(x)}, \quad (2.51)$$

где  $P(x)$  — требуемая вероятность безотказной работы в течение некоторой наработки  $x$ .

В формуле (2.51) доверительная вероятность  $\beta$  задается с учетом того, чтобы за наработку  $x$  не произошло отказа.

Например, для изучения характеристик надежности узлов и агрегатов, обеспечивающих безопасность движения автомобиля, объем выборки должен составлять не менее 28:

$$N = \frac{\ln(1-0,95)}{\ln 0,9} = 28.$$

Применение *параметрических методов* определения объема выборки связано с использованием параметров закона распределения. Для нормального распределения объем выработки рассчитывается из условия

$$N \geq \frac{t^2 \sigma^2}{\varepsilon^2}, \quad (2.52)$$

где  $t$  — обращенное значение функции Лапласа;  $\sigma^2$  — дисперсия оценки;  $\varepsilon$  — точность вычисления выборочной средней ( $\bar{X}$ ).

Учитывая, что точность  $\varepsilon$  может быть выражена через относительную точность  $\varepsilon_0 = \varepsilon / \bar{X}$ , формулу (2.52) можно записать следующим образом:

$$N \geq \frac{t^2 \sigma^2}{\varepsilon_0^2 \bar{X}^2} \geq \frac{t^2 V^2}{\varepsilon_0^2}, \quad (2.53)$$

где  $V$  — коэффициент вариации.

Обычно используют пяти- и однопроцентный уровень значимости, которому соответствуют значения  $t$ , равные 1,96 и 2,58. Отсюда по выражению (2.53) можно легко рассчитать объемы выборки для рекомендуемых нормативными документами [10] значений относительной точности 0,05, 0,10, 0,15 и 0,20. Результаты расчета для предельно расширенного диапазона коэффициентов вариации 0,1 ... 0,4 приведены в табл. 2.4.

Таблица 2.4

Объем выборки для нормального распределения

Уровень значимости $q$	Коэффициент вариации $V$	Относительная точность $\epsilon_0$			
		0,05	0,10	0,15	0,20
0,05	0,1	16	4	2	1
	0,2	62	16	7	4
	0,3	138	36	16	9
	0,4	245	62	28	16
0,01	0,1	27	7	3	2
	0,2	107	27	12	7
	0,3	240	60	27	15
	0,4	697	174	78	27

Можно видеть резкое возрастание объемов эксперимента при повышении точности и уровня значимости, а также коэффициента вариации. При невозможности реализации такого эксперимента для целей технической эксплуатации можно ограничиться уровнем значимости 0,05, считающимся достаточным при использовании программы STATISTICA, и относительной точностью 0,10... 0,15 при коэффициенте вариации 0,3. Обычно этого оказывается достаточным; при необходимости можно произвести дополнительные корректировки результатов на основе распределения Стьюдента, не повторяя сложных опытов.

В нормативных документах [10] приведены более строгие расчеты объемов выборки для нормального распределения при диапазоне уровней значимости в 20, 10, 5 и 2 %, которые необходимо использовать для более ответственных (полных) испытаний на надежность.

Что касается экспоненциального закона распределения, то в случае полных испытаний на надежность объем выборки в соответствии с нормативной документацией [10] рекомендуется определять по методике, разработанной для распределения Вейбулла, при величине коэффициента вариации  $V = 1$ . При этом значение объема выборки  $N$  получается из решения уравнения

$$\frac{2N}{\chi_{\beta}^2(2N)} = (\epsilon_2 + 1)^b, \quad (2.54)$$

где  $\chi_{\beta}^2(2N)$  — значение функции хи-квадрат при степени свободы  $2N$  и условии задания односторонней доверительной вероятности  $\beta$ , которое определяется по Приложению 2;  $\epsilon_0$  — предельная верхняя относительная ошибка в определении средней наработки до отказа;  $b$  — параметр распределения Вейбулла.

В табл. 2.5 представлены немного скорректированные по сравнению с [10] результаты расчета числа объектов полных испытаний для

Таблица 2.5

Число объектов испытаний при плане [NUN] в распределении Вейбулла по предельной относительной ошибке и различных значениях коэффициента вариации

Предельная относительная ошибка $\epsilon_0$	Уровень значимости $q$	Доверительная вероятность $\beta$	Число объектов $N$ при коэффициенте вариации $V$								
			0,4	0,5	0,6	0,7	0,8	0,9	1,0		
0,05	0,20	0,90	105	200	250	400	500	500	500	500	650
	0,10	0,95	170	250	400	500	650	800	1000	1000	1000
	0,05	0,975	235	375	500	1000	1000	1000	1000	1000	1000
	0,02	0,99	315	500	800	1000	1000	1000	1000	1000	1000
0,10	0,20	0,90	32	50	65	100	100	125	150	150	200
	0,10	0,95	50	80	100	150	200	200	250	250	400
	0,05	0,975	65	100	160	215	295	295	375	375	450
	0,02	0,99	100	150	200	315	400	400	500	500	650
0,15	0,20	0,90	15	25	32	40	40	65	80	80	80
	0,10	0,95	25	40	50	80	100	100	125	125	150
	0,05	0,975	32	50	80	110	140	140	175	175	210
	0,02	0,99	40	65	100	150	200	200	250	250	315
0,20	0,20	0,90	10	15	20	32	40	40	40	40	50
	0,10	0,95	15	25	32	40	40	50	50	80	100
	0,05	0,975	20	32	47	64	80	80	110	110	125
	0,02	0,99	25	40	65	80	125	125	150	150	150

Число объектов испытаний при плане [NUN] и логарифмически нормальном распределении по предельной относительной ошибке и различных значениях вариации

Предельная относительная ошибка $\varepsilon$	Уровень значимости $q$	Доверительная вероятность $\beta$	Число объектов $N$ при коэффициенте вариации $V$						
			0,4	0,5	0,6	0,7	0,8	0,9	1,0
0,05	0,20	0,90	100	150	250	315	400	500	650
	0,10	0,95	150	250	400	500	650	800	1000
	0,05	0,975	245	381	546	735	949	1000	Более 1000
0,10	0,20	0,99	315	500	800	Более 1000	Более 1000	Более 1000	Более 1000
	0,10	0,90	25	40	65	80	100	125	150
	0,05	0,95	40	65	100	125	150	200	250
0,15	0,20	0,975	62	96	137	184	238	296	359
	0,10	0,99	80	125	200	250	315	400	500
	0,02	0,90	13	20	25	40	50	55	65
0,20	0,20	0,95	20	32	40	50	80	100	110
	0,10	0,975	28	43	61	82	106	132	160
	0,02	0,99	40	50	80	125	150	200	200
0,20	0,20	0,90	6	10	15	20	25	32	40
	0,10	0,95	10	15	25	32	40	50	65
	0,05	0,975	16	24	35	46	60	74	90
0,20	0,20	0,99	20	32	50	65	80	100	125

вместо рекомендуемого диапазона предельной относительной ошибки  $\varepsilon$  и уровней значимости  $q$  при различных значениях коэффициентов вариации  $V$ .

Аналогичная по структуре табл. 2.6 отражает результаты расчета для логарифмически нормального распределения по выражению [10]

$$N = \frac{U_{\beta}}{\varepsilon} \ln(V^2 + 1) \left[ 1 + \frac{\ln(V^2 + 1)}{2} \right], \quad (2.55)$$

где  $U_{\beta}$  — квантиль нормированного нормального закона при односторонней доверительной вероятности  $\beta$ ;  $\varepsilon$  — предельная относительная ошибка при оценке среднего значения;  $V$  — коэффициент вариации.

При организации наблюдений за работой автомобилей на АТП согласно изложенной ранее методике определяется их количество (при исследовании показателей надежности обычно ограничиваются 15...25 автомобилями) и разрабатывается бланк наблюдений, в котором указываются гаражные номера автомобилей подконтрольной группы и значения измеряемых показателей (факторов) или входов и значения функции отклика или выходов. Кроме того, устанавливается интервал наблюдения.

Под *интервалом наблюдения* следует понимать промежуток времени между двумя соседними наблюдениями. Как правило, применяется равномерный интервал наблюдений. Величина интервала в часах, сутках, месяцах, годах или единицах пробега подвижного состава (км, тыс. км и т.д.) назначается исследователем с учетом обеспечения необходимого объема наблюдений.

Информация о надежности автомобилей может быть получена путем периодического опроса водителей, механиков или обработки соответствующей документации, имеющейся на АТП (личная карточка автомобиля, эксплуатационный отчет, журнал учета и движения материальных средств и т.д.), или комбинированно (использованием обоих способов).

### КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Перечислите основные характеристики случайных величин.
2. Дайте характеристику показателей рассеяния случайных величин.
3. Дайте определение гистограммы распределения, опишите процедуру ее построения.
4. Дайте краткую характеристику функции распределения и ее плотности.
5. Сформулируйте экспоненциальный закон распределения, дайте его краткую характеристику.
6. Сформулируйте нормальный закон распределения, дайте его краткую характеристику.

7. Сформулируйте логарифмически нормальный закон распределения, дайте его краткую характеристику.

8. Сформулируйте закон распределений Вейбулла, дайте его краткую характеристику.

9. Опишите процедуру определения параметров закона распределения Вейбулла.

10. Дайте определения критерия Пирсона и опишите процедуру проверки закона распределения на нормальность с помощью критерия Пирсона.

11. Опишите процедуру проверки нормальности закона распределения с помощью коэффициентов асимметрий и эксцесса.

12. Дайте определения доверительного интервала и доверительной вероятности.

13. Дайте определение и графическую интерпретацию односторонней и двухсторонней доверительной вероятности, квантиля распределения и уровня значимости.

14. Сформулируйте закон распределения Стьюдента. Каковы особенности его использования в статистических расчетах?

15. Приведите основные расчетные формулы для определения границ доверительных интервалов для параметров нормального закона распределения.

16. Приведите основные формулы для расчета границ доверительного интервала для экспоненциального закона распределения.

17. Каким образом определяется объем выборки для непараметрического подхода?

18. Каким образом определяется объем выборки для нормального закона распределения?

19. Каким образом определяется объем выборки для экспоненциального закона распределения?

20. Каким образом определяется объем выборки для закона распределения Вейбулла?

21. Каким образом определяется объем выборки для логарифмически нормального закона распределения?

22. Дайте характеристику процедуры организации наблюдений за работой автомобилей на АТП.

## ГЛАВА 3

# Использование критериев Стьюдента, Фишера и дисперсионного анализа при выявлении расхождения сравниваемых выборок случайных величин и обосновании возможности их объединения. Разделение смешанных выборок

### 3.1. Простейший случай проверки «нулевой» гипотезы о принадлежности двух выборок одной генеральной совокупности

В гл. 2 описан самый первый и необходимый этап обработки экспериментальных выборок случайных величин, который необходимо проводить практически во всех задачах научных исследований в области технической эксплуатации автомобилей, включая корреляционный и регрессионный анализы и активный многофакторный эксперимент. Однако необходимо учитывать, что очень редко удастся непосредственно получить достоверный конечный результат при простой статистической обработке данных по причине или их недостаточного количества, или наличия в них мешающих или неудобных данных, не связанных с грубыми ошибками, а отражающих качественно другие процессы, которые на начальном этапе сложно исключить из рассмотрения. Поэтому, как правило, необходимы дополнительные более сложные статистические расчеты, позволяющие в конечном счете с определенной степенью приближения решать поставленные задачи.

В этом отношении достаточно плодотворным является путь сведения из нескольких небольших экспериментальных выборок, полученных в различных условиях, одну более мощную выборку, повышающую точность конечных статистических оценок. Однако это можно сделать только при условии, что все они принадлежат одной генеральной совокупности, т. е. для этого необходимо обоснованное статистическое подтверждение. Также в случае, когда по статистическим показателям различных выборок пытаются произвести сравнительную оценку качества каких-либо изучаемых процессов, простое сопоставление средних величин не может служить надежным результатом изменения (повышения или снижения) качества, несмотря на существенные различия между этими оценками.

В случае же наличия в этих данных смешанных результатов, отражающих различные процессы, такое сравнение вообще неправомерно проводить без предварительной «фильтрации» результатов. В целом могут быть и другие, более сложные, ситуации, которые необходимо решать в процессе исследования. Следовательно, обоснованный вывод может быть получен лишь путем иногда достаточно сложного научно поставленного анализа статистических данных. Поэтому необходимо обратить особое внимание на данную сторону научных исследований в сфере технической эксплуатации автомобилей и показать ее значимость на примере уже выполненных работ.

К сожалению, ранее указанным методам не уделялось должного внимания, несмотря на их достаточную теоретическую проработку, что можно объяснить сложностью «ручных» вычислений. Однако появление компьютерных программ (STATISTICA и др. [4]), как уже отмечалось ранее, значительно упростило процедуру расчетов даже по таким сложным задачам, как многофакторный дисперсионный анализ. Современному исследователю необходимо максимально использовать эти возможности, чего можно добиться только при правильном понимании процессов в целом.

Описанные в гл. 2 методы определения доверительных интервалов для параметров распределений случайных величин редко используются непосредственно, но являются основой косвенных способов различных проверок для многих статистических исследований. При этом наиболее часто ставится задача оценки степени расхождения между полученными в эксперименте основными параметрами независимых распределений случайных величин — средними значениями и дисперсиями, отражающими разные, но технологически или физически подобные процессы, как правило, отличающиеся друг от друга. В этом случае должна проверяться так называемая нулевая гипотеза  $H_0$  о принадлежности этих выборок одной генеральной совокупности, поскольку может оказаться, что эти различия являются несущественными и объясняются случайностью характера распределений.

В простейшем случае проверка нулевой гипотезы может быть выполнена для двух выборок  $N_1$  и  $N_2$  независимых случайных величин  $X$  и  $Y$ , распределенных по нормальному закону, на основе равенства их центров распределений, с одной стороны, и дисперсий — с другой. Наиболее популярной является проверка о равенстве центров по критерию Стьюдента ( $t$ -критерию). В основе ее лежит положение теории [2] о том, что случайная величина, отражающая разность двух выборочных средних  $(\bar{Y}^* - \bar{X}^*)$  для однородных случайных величин с равными дисперсиями  $\sigma_X^2 = \sigma_Y^2$ , будет распределена по закону Стьюдента с центром  $(\bar{Y}^* - \bar{X}^*)$ , числом степеней свободы  $k = N_1 + N_2 - 2$  и дисперсией

$$\sigma_{\bar{Y}-\bar{X}}^2 = \frac{\sigma_X^2}{N_1} + \frac{\sigma_Y^2}{N_2}. \quad (3.1)$$

Задаваясь определенным уровнем значимости  $q$ , можно построить доверительный интервал для разности центров распределения  $(\bar{Y} - \bar{X})$  с доверительной вероятностью  $P = 1 - q$  вида

$$(\bar{Y}^* - \bar{X}^*) \pm t_{q,k} \sqrt{\frac{(N_1 + N_2)(N_1 S_X^2 + N_2 S_Y^2)}{N_1 N_2 (N_1 + N_2 - 2)}}, \quad (3.2)$$

где  $S_X^2, S_Y^2$  — статистические оценки дисперсий.

При этом если значение точки 0 попадает внутрь этого интервала, то нулевая гипотеза о равенстве центров распределений  $\bar{X}$  и  $\bar{Y}$  и о принадлежности выборок одной генеральной совокупности принимается. Если же значение точки 0 выходит за пределы интервалов формулы (3.2), то нулевая гипотеза бракуется, поскольку при этом вероятность того, что точка 0 попадет во внутрь интервала формулы (3.2), будет равна  $q$  (0,05 и ниже) и будет представлять собой событие практически невозможное.

Практически проверка сводится к вычислению критерия

$$t_{q, N_1 + N_2 - 2} = \frac{\bar{Y}^* - \bar{X}^*}{\sqrt{N_1 S_X^2 + N_2 S_Y^2}} \sqrt{\frac{N_1 N_2 (N_1 + N_2 - 2)}{N_1 + N_2}} \quad (3.3)$$

и установлению того, выполняется ли неравенство

$$|t| \leq t_{q,k} \quad (3.4)$$

для принятия нулевой гипотезы  $H_0$  или неравенство

$$|t| \geq t_{q,k} \quad (3.5)$$

для ее выбраковки. Наиболее распространена проверка условия (3.5) о существенной мере расхождения между средними, по которой делается заключение об изменении качества контролируемых процессов. Такая проверка является стандартной в системе STATISTICA и называется  $t$ -тестом.

Особенностью использования данной проверки является предварительное задание условия зависимости или независимости для проверяемых выборок. В первом случае необходимо обязательное равенство количества случайных чисел для сравниваемых выборок, т.е.  $N_1 = N_2$ , а при вводе неодинакового количества программа автоматически ведет расчет по меньшему  $N$ , отбрасывая остальные числа большей выборки. Зависимыми выборками можно считать результаты повторно полученных статистических показателей случайных величин какого-либо процесса после изменения его качества, например расхода топлива у опытной группы автомобилей после проведения определенных мероприятий по его снижению. Считается [4], что для зависимых выборок  $t$ -тест обладает большей чувствительностью. Для проведения расчетов непосредственно по выражению (3.3) необходимо задать условие независимости выборок.

В результате расчета программы STATISTICA не только выдает значение  $t$  с указанием числа степеней свободы, но и определяет уровень значимости (в вероятностном виде) полученного критерия Стьюдента исходя из двухсторонней доверительной вероятности. При этом критическим порогом считается уровень значимости 5%, по достижении или превышении которого все результаты выводятся красными цифрами, свидетельствующими о выполнении  $t$ -теста о существенности различия средних. Однако для практики технической эксплуатации автомобилей приемлемым уровнем значимости можно считать показатель 10...20% (последний соответствует уровню односторонней доверительной вероятности  $p = 0,9$ ).

Более редко используется при «ручных» расчетах проверка нулевой гипотезы  $H_0$  на основе показателя о равенстве или существенном отличии дисперсий двух выборок, оценки которого производится на основе так называемого распределения Фишера ( $F$ -распределения). Дополнительно к этому распределению можно прибегать в случаях, если критерий Стьюдента не выявляет желательной значимости различия между выборками на основе средних и хотелось бы получить дополнительную аргументацию.

Гипотезы о дисперсиях играют очень большую роль, поскольку измеряемая дисперсией величина рассеяния как раз и характеризует такие конструкторские и технологические показатели, как точность машин и приборов, погрешность показателей измерительных приборов, точность технологических процессов [2]. Очевидно, что при равенстве средних значений выходных технологических показателей более совершенным будет считаться более стабильный процесс, имеющий меньший диапазон рассеяния, и т.д. Для проверки гипотезы о равенстве дисперсий для двух генеральных совокупностей, оценка которых производится по независимым выборкам, необходимо знать такую функцию для их статистических показателей, распределение которой не зависело бы от каких-либо неизвестных параметров.

Этому последнему условию удовлетворяет распределение отношения двух несмещенных оценок дисперсий, вычисленных по выражению (2.3) и полученных из независимых выборок, распределенных по нормальному закону, т.е. распределение величины

$$\frac{S_1^2}{S_2^2} = F, \quad (3.6)$$

причем обычно в качестве числителя  $S_1^2$  берется наибольшая из двух несмещенных оценок дисперсии. В целом  $F$ -распределение зависит только от числа степеней свободы  $k_1 = N_1 - 1$  и  $k_2 = N_2 - 1$ , если выборка имеет объемы: первая (с большей дисперсией)  $N_1$  и вторая  $N_2$ .

Вид распределения Фишера в дифференциальной форме для определенных значений  $k_1$  и  $k_2$  приведен на рис. 3.1. При помощи вероятностного калькулятора системы STATISTICA можно найти

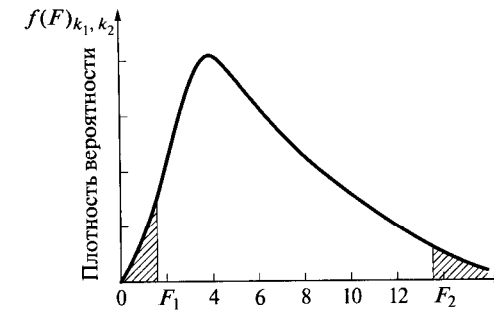


Рис. 3.1.  $F$ -распределения для определенных значений  $k_1$  и  $k_2$  и критические границы  $F_1$  и  $F_2$  [2]

интегральную и дифференциальную функции этого распределения для любых значений  $k_1$  и  $k_2$ , однако для практики более удобным, как и в случае с критерием Стьюдента, является табличное представление этого показателя, введенное самим Р.Фишером (Приложение 4).

Основой таблиц Фишера является статистическое понятие, характерное для любого доверительного интервала, критической области для оценки  $F$ , за которую принимают показанные на рис. 3.1 два заштрихованных интервала: интервал больших значений, удовлетворяющий неравенству  $F > F_2$ , и интервал малых значений:  $0 < F < F_1$ . Граничные точки подбираются таким образом, чтобы при уровне значимости  $q$  вероятности попадания в верхние и нижние критические области были одинаковы, поскольку при этом обеспечивается наибольшая чувствительность критерия Фишера:

$$P(F > F_2) = q/2 \text{ и } P(0 < F < F_1) = q/2. \quad (3.7)$$

Гипотеза о равенстве дисперсий  $\sigma_1^2 = \sigma_2^2$  принимается, если значение отношения (3.6) не попадает в критическую область, т.е. находится внутри диапазона  $F_1 \dots F_2$  (рис. 3.1) для любых вариантов вычисления: большей дисперсии к меньшей, и наоборот. Если же выборочное значение  $F$  оказывается в критической области, т.е. превышает значение  $F_2$  для отношения большей дисперсии к меньшей, или же оказывается ниже значения  $F_1$  для отношения меньшей дисперсии к большей, то гипотеза о равенстве дисперсий  $\sigma_1^2 = \sigma_2^2$  должна быть отвергнута, также как и нулевая гипотеза  $H_0$  о принадлежности двух выборок одной генеральной совокупности.

Учитывая условие (3.7), более удобно производить проверку по верхней границе  $F_2$ , вычисляя отношение наибольшей дисперсии к меньшей. Исходя из этого Р.Фишером были рассчитаны односторонние критические верхние одно- и пятипроцентные границы для любых комбинаций числа степеней свободы  $k_1$  и  $k_2$  (см. Приложение 4), которые соответствуют двух- и десятипроцентным уровням

значимости  $q$ . Последнее считается недостатком этих таблиц, который для ответственных случаев в настоящее время может быть устранен компьютерными вычислениями при помощи вероятностного калькулятора.

Необходимо учитывать, что многие последующие теоретические и практические работы в области статистики были ориентированы именно на эти таблицы и отработанные критерии будут сохраняться в будущем, в частности в области регрессионного анализа и активного многофакторного эксперимента. Поэтому в Приложении 4 представлен необходимый для практического использования наиболее полный вариант таблиц Фишера.

В программе STATISTICA при выполнении  $t$ -теста автоматически производится проверка и по  $F$ -критерию с выдачей значения двухсторонней вероятности попадания показателя в критический диапазон (уровня значимости); при вероятности  $p < 0,10$  в соответствии с таблицами Фишера уже можно ставить вопрос о непринятии нулевой гипотезы о равенстве дисперсий, хотя в зависимости от целей исследования может быть использован и более жесткий уровень ( $p < 0,01$  и ниже).

Очень часто в задачах статистического контроля качества производственных процессов необходимо осуществлять проверку однородности ряда дисперсий из  $k$  последовательных выборок одинакового объема, например характеризующих степень разладки отдельной ячейки или производственного процесса в целом в различные периоды времени. Поскольку попарное сравнение всех комбинаций выборок в этом случае является достаточно трудоемкой и даже сложно интерпретируемой операцией, были предложены (М. С. Бартлеттом [2]) упрощенные приемы проверки, основанные на рассмотрении последовательности величин

$$G_i = \frac{S_i^2}{S_1^2 + S_2^2 + \dots + S_k^2}, \quad (3.8)$$

где  $k$  — номер выборки.

Для максимального из членов этой последовательности  $G_{\max}$ , которой отвечает наибольшее значение  $S_i^2$ , т.е.  $\max[S_i^2]$ , найден закон распределения, с помощью которого Кохраном [4] (в другом варианте перевода Кохраном [1, 2]) рассчитаны специальные таблицы. В них при числе  $k$  выборок от 2 до 120 и объемах выборок  $r$  от 2 до 145 даны значения показателя  $G_{\max}$ , отвечающие пяти- и однопроцентному уровню значимости (Приложение 5).

Если найденное из наблюдений значение  $G_{\max}$  окажется больше указанного в таблице для соответствующих  $k$  и  $r - 1$ , то гипотеза об однородности (равенстве) ряда дисперсий должна быть отвергнута и свидетельствовать о разладке контролируемого производственного процесса или других обстоятельствах. Проверка по критерию (3.8) широко используется при обработке данных активного многофак-

торного эксперимента (см. гл. 5), где количество опытов  $k$  может быть достаточно большим, а количество повторных измерений в каждом опыте  $r$  обычно бывает небольшим — два или три. Выход за пределы, установленные таблицами Кохрана, в данном случае будет указывать на неравноточность измерений в опытах вследствие ошибок в установлении уровней факторов и ошибок регистрации и будет ставить под сомнение результаты исследования в целом.

### 3.2. Однофакторный и многофакторный дисперсионные анализы как общие методы проверки расхождения между средними при большом количестве статистических выборок

Дисперсионный анализ, основы которого были разработаны Р. Фишером в 1920-х гг., помимо того, что позволяет избежать многочисленных повторных сравнений различных групп выборок с помощью серий  $t$ -критериев, является более эффективным и, что особенно ценно, более информативным для малых выборок. Если даже овладеть самой простой техникой средств дисперсионного анализа, реализованной в программе STATISTICA, то можно ощутить его заметные преимущества в конкретных исследованиях.

Популярное изложение особенностей дисперсионного анализа дается одним из авторов работы [4] в виде следующего примера (табл. 3.1), показывающего результаты статистических наблюдений за двумя группами, состоящими из трех объектов. Здесь средние значения двух групп существенно различны (2 и 6), хотя имеют одинаковую сумму квадратов отклонений — 2, в сумме по двум группам — 4. Если

Таблица 3.1

Пример дисперсионного анализа

Параметры	Группа 1	Группа 2
Наблюдение 1	2	6
Наблюдение 2	3	7
Наблюдение 3	1	5
Среднее	2	6
Сумма квадратов отклонений от среднего (СК)	2	2
Общее среднее (без учета принадлежности к группе)	4	
Общая сумма квадратов отклонения от общего среднего	28	



вычислить сумму квадратов отклонений без учета групповой принадлежности, то получим 28. Таким образом, очевидно, что дисперсия (сумма квадратов), основанная на внутригрупповой изменчивости, показывает гораздо меньшие значения, чем при вычислении на основе общей изменчивости (относительно общего среднего), и это более ощутимое различие в сумме квадратов объясняется различием между средними, имеющим гораздо меньшую чувствительность.

В общем случае теория [2] дает следующее обоснование схемы простейшего случая однофакторного дисперсионного анализа. Пусть имеется  $m$  обследованных по определенному признаку групп (например, предприятий) из  $n$  объектов в каждой группе, по которым получены определенные статистические оценки (например, измерены показатели эффективности). При этом предполагается, что внутри групп статистические оценки нормально распределены с одним и тем же стандартным отклонением  $\sigma_Y$ , но с различными центрами распределения  $v_1, v_2, \dots, v_m$ . Результаты оценок и их обработка приведены в табл. 3.2.

Обозначим через  $\bar{Y}_i^*$  — среднее по  $i$ -й строке:

$$\bar{Y}_i^* = \frac{1}{n} \sum_{j=1}^n Y_{ij} \quad (i=1, \dots, m; j=1, \dots, n); \quad (3.9)$$

а также общее среднее:

$$\bar{Y}^* = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n Y_{ij} = \frac{1}{m} \sum_{i=1}^m \bar{Y}_i^*. \quad (3.10)$$

Таблица 3.2

**Расположение исходных данных при однофакторном дисперсионном анализе**

Номер группы	Номер измерения (элемента группы)						Среднее по группам
	1	2	.	$j$	.	$n$	
1	$Y_{11}$	$Y_{12}$	.	$Y_{1j}$	.	$Y_{1n}$	$\bar{Y}_1^*$
2	$Y_{21}$	$Y_{22}$	.	$Y_{2j}$	.	$Y_{2n}$	$\bar{Y}_2^*$
.	.	.	.	.	.	.	.
$i$	$Y_{i1}$	$Y_{i2}$	.	$Y_{ij}$	.	$Y_{in}$	$\bar{Y}_i^*$
.	.	.	.	.	.	.	.
$m$	$Y_{m1}$	$Y_{m2}$	.	$Y_{mj}$	.	$Y_{mn}$	$\bar{Y}_m^*$
Общее среднее							$\bar{Y}^*$

Даже если в генеральной совокупности математические ожидания совпадают ( $v_1 = v_2 = \dots = v_m$ ) и, следовательно, выборка произведена из единой нормально распределенной совокупности, то в общем случае выборочные средние по группам не равны между собой:  $\bar{Y}_1^* \neq \bar{Y}_2^* \neq \dots \neq \bar{Y}_m^*$ . В данном случае проверка нулевой гипотезы  $H_0$  о том, что  $v_1 = v_2 = \dots = v_m$ , будет осуществляться на основе полученных статистических данных, отраженных в табл. 3.2. Если данная гипотеза верна, то выборочные средние значения по группам  $\bar{Y}_1^*, \bar{Y}_2^*, \dots, \bar{Y}_m^*$  не должны существенно различаться. Но если различия будут признаны существенными, то исходная гипотеза  $H_0$  неверна и выборка неоднородна.

Существенность или не существенность различия между средними может оцениваться, как уже отмечалось ранее, отношением частных составляющих, на которые раскладывается общая дисперсия выборки. Для этого используется способность общей суммы квадратов отклонений в исходной выборке распределяться на среднюю сумму квадратов отклонений внутри групп и суммой квадратов отклонений между группами. Последнее можно видеть при помощи следующего преобразования общей суммы квадратов, при вводе в нее с противоположными знаками групповых средних  $\bar{Y}_i^*$ :

$$Q = \sum_{i=1}^m \sum_{j=1}^n (Y_{ij} - \bar{Y}^*)^2 = \sum_{i=1}^m \sum_{j=1}^n (Y_{ij} - \bar{Y}_i^* + \bar{Y}_i^* - \bar{Y}^*)^2 = \sum_{i=1}^m \sum_{j=1}^n (Y_{ij} - \bar{Y}_i^*)^2 + 2 \sum_{i=1}^m \sum_{j=1}^n (Y_{ij} - \bar{Y}_i^*)(\bar{Y}_i^* - \bar{Y}^*) + \sum_{i=1}^m \sum_{j=1}^n (\bar{Y}_i^* - \bar{Y}^*)^2, \quad (3.11)$$

причем

$$S = \sum_{i=1}^m \sum_{j=1}^n (Y_{ij} - \bar{Y}_i^*)(\bar{Y}_i^* - \bar{Y}^*) = \sum_{i=1}^m (\bar{Y}_i^* - \bar{Y}^*) \sum_{j=1}^n (Y_{ij} - \bar{Y}_i^*). \quad (3.12)$$

Но  $\sum_{j=1}^n (Y_{ij} - \bar{Y}_i^*) = 0$ , поскольку представляет собой сумму отклонений наблюдений в  $i$ -й группе от средней по этой группе, поэтому  $S = 0$ . Дополнительно учитывая, что

$$\sum_{i=1}^m \sum_{j=1}^n (\bar{Y}_i^* - \bar{Y}^*)^2 = n \sum_{i=1}^m (\bar{Y}_i^* - \bar{Y}^*)^2, \quad (3.13)$$

получим окончательные результаты для преобразования выражения (3.11):

$$Q = n \sum_{i=1}^m (\bar{Y}_i^* - \bar{Y}^*)^2 + \sum_{i=1}^m \sum_{j=1}^n (Y_{ij} - \bar{Y}_i^*)^2 = Q_1 + Q_2. \quad (3.14)$$

В выражении (3.14) сумма  $Q_1 = n \sum_{i=1}^m (\bar{Y}_i^* - \bar{Y}^*)^2$  называется суммой квадратов отклонений между группами и характеризует степень рас-

хождения групповых средних; ее также называют рассеиванием по факторам [2]. Сумма  $Q_2 = \sum_{i=1}^m \sum_{j=1}^n (Y_{ij} - \bar{Y}_i^*)^2$  называется суммой квадратов отклонений внутри групп, или остаточной суммой квадратов, и характеризует остаточное рассеяние.

В математической статистике доказывается [2], что величины  $Q_1$  и  $Q_2$  при любых предположениях о соотношениях между  $v_i$  независимы и распределены по закону  $\chi^2$  Пирсона с  $(m - 1)$  и  $m(n - 1)$  степенями свободы. В таких случаях существование или не существование расхождений между средними на фоне остаточных внутригрупповых отклонений определяется с помощью дисперсионного соотношения

$$F = \frac{Q_1 / (m - 1)}{Q_2 / (n - m)} = \frac{S_{\text{МГ}}^2}{S_{\text{ср.вг}}^2}, \quad (3.15)$$

подчиняющегося  $F$ -распределению с  $(m - 1)$  и  $(n - m)$  степенями свободы. Таким образом, в выражении (3.15)  $S_{\text{МГ}}^2$  представляет собой оценку межгрупповой дисперсии, а  $S_{\text{ср.вг}}^2$  — оценку средней внутригрупповой дисперсии.

Гипотезу об однородности исходной выборки  $H_0$  отвергают, если полученное по выражению (3.15) значение  $F$ -критерия оказывается выше критического уровня ( $F > F_{\text{кр}}$ ), найденного из таблиц Фишера (см. Приложение 4), или по другим соображениям. Очевидно, что при равенстве групповых средних величина межгрупповой дисперсии будет практически совпадать с величиной общей дисперсии, подсчитанной без учета групповой принадлежности.

Схему однофакторного дисперсионного анализа удобно представить в виде табл. 3.3 [2]. Число степеней свободы, на основе которых рассчитываются составляющие дисперсий, легко проверяется следующим тождеством:

$$mn - 1 = (m - 1) + (n - m). \quad (3.16)$$

В общем случае количество объектов внутри групп может быть различным, однако это не изменяет схему расчета, при которой соответствующим образом корректируются числа степеней свободы. Покажем последнее на примере, заимствованном из работ кафедры прикладной математики МАДИ (ГТУ) [2], который будет использован и далее.

Здесь решалась задача построения модели зависимости производительности труда  $Y$  (тыс. руб./чел. в год) от фондовооруженности  $X$  предприятия (тыс. руб./чел.). Для этого были собраны данные по 20 различным предприятиям, разделенным по качественному признаку на три группы: новые; старые, но модернизированные; просто старые предприятия. Исходные данные приведены в табл. 3.4 (данные условные). Если указанные группы предприятий однородны, то их

Таблица 3.3

Схема однофакторного дисперсионного анализа

Компоненты дисперсии	Сумма квадратов отклонений	Число степеней свободы	Выражения для оценок дисперсий
1	2	3	4
Полная (общая)	$\sum_{i=1}^m \sum_{j=1}^n (Y_{ij} - \bar{Y}^*)^2$	$mn - 1$	$\frac{1}{mn - 1} \sum_{i=1}^m \sum_{j=1}^n (Y_{ij} - \bar{Y}^*)^2$
Межгрупповая составляющая	$\sum_{i=1}^m n_i (\bar{X}_i - \bar{X})^2$	$m - 1$	$\frac{1}{m - 1} \sum_{i=1}^m \sum_{j=1}^n (\bar{Y}_i^* - \bar{Y}^*)^2$
Внутригрупповая составляющая	$\sum_{i=1}^m \sum_{j=1}^n (Y_{ij} - \bar{Y}_i^*)^2$	$(n - m)$	$\frac{1}{(n - m)} \sum_{i=1}^m \sum_{j=1}^n (Y_{ij} - \bar{Y}_i^*)^2$

Таблица 3.4

Данные по 20 различным предприятиям

I. Новые предприятия			II. Старые модернизированные предприятия			III. Старые предприятия		
Номер предприятия	$Y_i$	$X_i$	Номер предприятия	$Y_i$	$X_i$	Номер предприятия	$Y_i$	$X_i$
1	112,0	100,0	8	34,0	29,0	14	26,4	25,9
2	83,3	76,7	9	36,2	28,5	15	29,5	28,7
3	100,0	90,0	10	32,9	26,5	16	30,5	30,0
4	75,0	66,7	11	32,9	28,0	17	34,5	32,8
5	85,7	71,4	12	33,0	27,0	18	27,7	26,9
6	66,6	50,0	13	33,6	27,8	19	27,0	26,8
7	87,5	75,0				20	29,2	28,7

можно объединить в одну общую группу и строить единую модель, в противном случае такая модель должна быть построена для каждой группы предприятий.

Предварительные расчеты показывают, что коэффициент вариации (по выходному показателю  $Y$ ) для объединенной группы состав-

ляет 0,56, что выше критической величины для нормального закона (0,33); в то же время для выделенных групп он составил 0,17, 0,037 и 0,063 соответственно. Таким образом, выборка в целом неоднородна и предварительное разбиение на три группы с нормальными законами распределения является целесообразным.

Результаты расчетов при выполнении дисперсионного анализа представлены в табл. 3.5. Здесь сумма квадратов межгрупповых отклонений  $Q_1$  при различном числе данных в группах ( $n_1 = 7$ ;  $n_2 = 6$ ;  $n_3 = 7$ ) определялась в соответствии с формулой (3.14) по выражению

$$Q_1 = n_1(\bar{Y}_1^* - \bar{Y}^*) + n_2(\bar{Y}_2^* - \bar{Y}^*) + n_3(\bar{Y}_3^* - \bar{Y}^*) = 7(87,16 - 50,86)^2 + 6(33,77 - 50,86)^2 + 7(29,21 - 50,86)^2 = 14\,258.$$

Следовательно, межгрупповая дисперсия с учетом того, что число степеней свободы равно  $m - 1 = 2$ , составит:

$$S_{\text{мг}}^2 = 14\,258 : 2 = 7\,129.$$

Таблица 3.5

**Результаты расчетов при выполнении дисперсионного анализа**

Составляющие оценок	Количество расчетных данных	Средние значения	Сумма квадратов отклонений	Число степеней свободы	Оценка дисперсий
1	2	3	4	5	6
<i>А. Основные показатели при статистическом анализе</i>					
1-я группа	7	87,16	1369,6	6	228,3
2-я группа	6	33,77	8,1	5	1,62
3-я группа	7	29,21	41,7	6	6,95
Общая	20	50,86	15 667	19	824,1
<i>Б. Составляющие (суммы квадратов и др.) при дисперсионном анализе</i>					
Между группами	3	—	14 265	2	7 129
Внутри групп	7 + 6 + 7	—	1 419	17	83,47
Общая	20	—	15 677	19	824,1

**Примечание.** Сумма первых трех чисел 4-го столбца таблицы А должна быть равна (с точностью до округления) второму числу этого же столбца таблицы Б; сумма двух первых чисел 4-го столбца таблицы Б должна быть равна третьему числу этого столбца.

Межгрупповая дисперсия при неодинаковом числе объектов в группах должна определяться не просто как  $\sum_{i=1}^m (\bar{Y}_i^* - \bar{Y}^*)^2 / (m - 1)$ , а с учетом числа измерений в каждой группе.

Сумму квадратов межгрупповых отклонений можно получить, сложив данные по группам из табл. 3.5(А):

$$Q_2 = 1\,369,6 + 8,1 + 41,7 = 1\,419.$$

Внутригрупповая дисперсия, для которой число степеней свободы в соответствии с табл. 3.3 и условием (3.16) будет равно

$$(n - m) = 20 - 3 = 17,$$

составит:  $S_{\text{ср.вр}}^2 = 1\,419 : 17 = 83,47$ . Поскольку межгрупповая дисперсия больше средней внутригрупповой, то проверка по  $F$ -критерию обоснована:

$$F = 7\,129 : 83,47 = 85,38.$$

Из Приложения 4 можно видеть высокий уровень значимости (менее 2%) полученной оценки (точную величину можно получить при помощи вероятностного калькулятора). Таким образом, разбиение на группы целесообразно. В то же время из табл. 3.5(А) можно видеть близость данных 2-й и 3-й групп, которую можно проверить по критерию Стьюдента (формула (3.3)). Здесь попарный  $t$ -тест необходимо применять потому, что дисперсионный анализ показал неоднородность выборки. Проверка по тесту (3.3) показала, что, несмотря на кажущуюся близость, различие между группами значительное как по средним (уровень значимости меньше 5%), так и по дисперсиям; объединение указанных групп предприятий в одну общую (такая возможность рассматривалась в старых вариантах примера) приведет к искажению конечных результатов.

В случае, когда дисперсионный анализ показывает однородность групп, для оценки расхождения между группами необходимо рассмотреть другой тест, который должен учитывать внутригрупповую дисперсию и число степеней свободы при ее оценке. При одинаковом числе  $N$  компонентов в каждой группе данный тест имеет вид [2]:

$$t_{q, m(n-1)} = \frac{\bar{Y}_{i(1)}^* - \bar{Y}_{i(2)}^*}{S_{\text{ср.вр}}} \sqrt{\frac{n}{2}}. \quad (3.17)$$

В случае неодинаковости компонентов  $n_{i(1)}$  и  $n_{i(2)}$  значение  $n$  вычисляется как среднее между  $n_{i(1)}$  и  $n_{i(2)}$ :  $n = (n_{i(1)} + n_{i(2)}) / 2$ .

Процедура для двух-, трех- и более факторного дисперсионного анализа по схеме не отличается от описанной ранее, только имеет более сложные расчеты. В частности, основному тождеству (3.14) для однофакторного анализа в случае двухфакторного будет отвечать тождество

$$Q = Q_1 + Q_2 + Q_3, \quad (3.18)$$

где  $Q_1$  и  $Q_2$  условно называется суммой квадратов разностей между строками и между колонками исходной матрицы с числом степеней свободы  $(r - 1)$  и  $(v - 1)$ ;  $r$  и  $v$  — соответственно число строк и число колонок матрицы;  $Q_3$  называется остаточной суммой квадратов и имеет число степеней свободы  $(r - 1)(v - 1)$ .

Для проверки степени значимости расхождений, обнаруженных в средних по строкам и по колонкам, вычисляются критерии:

$$F_A = \frac{Q_1/(r-1)}{Q_3/((r-1)(v-1))} = \frac{S_1^2}{S_3^2}; \quad (3.19)$$

$$F_B = \frac{Q_2/(r-1)}{Q_3/((r-1)(v-1))} = \frac{S_2^2}{S_3^2}. \quad (3.20)$$

Все приведенные расчеты легко выполняются программой STATISTICA при условии правильного задания исходных данных. Кроме того, указанной программой проводятся значительно более сложные виды дисперсионного анализа — с так называемыми вложенными планами, планами с повторными измерениями и др.

### 3.3. Применение кластерного анализа и метода подбора закона распределения в ограниченном диапазоне данных для разделения смешанных выборок

Широкие возможности программы STATISTICA позволяют провести разделение исходной выборки табл. 3.4 и в случае отсутствия точной информации о качественной принадлежности данных, только на основе предположения об ее неоднородности в результате смешивания статистических данных из различных генеральных совокупностей. (Зачастую неоднородность выборки может быть следствием попадания в нее грубых ошибочных данных, отсеивание которых также может быть осуществлено изложенными далее методами.)

Подобные обстоятельства обычно качественно понимаются до начала исследований, хотя количество уровней разделения по смешанности может быть заранее неизвестно или задано приблизительно на основе общих соображений. На примере данных табл. 3.4 была показана целесообразность разделения предприятий на три группы, хотя в первоначальном варианте работы [2] обосновывалась целесообразность разделения на две группы за счет объединения всех старых предприятий в одну группу. Покажем возможность решения задач разделения на основе кластерного анализа и методом подбора

закона распределения для части данных, находящихся в ограниченном диапазоне рассеяния.

Процедура кластерного анализа позволяет осуществлять разделение с тем или иным качеством с любым контролируемым уровнем (число кластеров и методика разделения задаются самим исследователем); при этом по полученным результатам может проверяться и само качество разделения. Необходимо отметить, что строгого определения кластера не существует вообще. Под *кластером* обычно понимают группу объектов, обладающих плотностью (плотность объектов внутри кластера выше, чем вне его), дисперсией, отличимостью от других кластеров, формой (например, кластер может иметь очертания гиперболы или эллипсоида), размером [2]. Сложность задач кластерного анализа заключается в том, что реальные объекты являются многомерными, т.е. описываются не одним, а несколькими параметрами, и объединение объектов в группы производится в пространстве многих измерений, что достаточно трудно поддается представлению. Кроме того, данные могут носить нечисловой характер.

В целом методы кластерного анализа подразделяются на агломеративные (от *лат.* *agglomerato* — скопление) и итеративные дивизивные (от *фр.* *division* — деление, разделение). В агломеративных, или объединительных, методах происходит последовательное объединение наиболее близких объектов в один кластер; процесс такого объединения можно показать на графике в виде дендрограммы или дерева объединения. В основу объединения в той или иной степени положены характеристики расстояния между объектами, которые будут являться основой сходства для одномерных объектов или одной из мер сходства для многомерных. При этом для оценки расстояний удобно использовать обычную евклидову метрику, при которой для двухпараметрического объекта он может быть изображен точкой на плоскости, а расстояние между объектами вычисляется по теореме Пифагора.

Другим способом оценки является так называемое манхэттенское расстояние, или расстояние городских кварталов, связанное с перемещением человека по улицам города, а не с движением по ровной местности. Также могут быть использованы и другие методы оценки сходства. Для рассматриваемого нами случая одномерной выборки разделение легко производится агломеративным методом *tri clustering* (дерево объединения) с использованием евклидовой метрики, результаты которого затем показываются графически.

Из итеративных наиболее удобным является метод группировки на основе  $k$ -средних (*k-means clustering*), который работает непосредственно с объектами, а не с матрицей сходства. При нем объект относится к тому классу, расстояние до которого в евклидовом пространстве минимально. Итоговый результат представляется в виде таблиц о принадлежности переменных к определенному кластеру и расстояний до центра кластеров. Проверка работы программы по-

казала, что метод *tri clustering* давал разделение только на два кластера с совпадением с данными табл. 3.4: новые и старые предприятия. В то же время метод *k-means clustering* при задании количества кластеров, равного трем, приводил к разделению с полным совпадением с разделением данных табл. 3.4: новые, старые модернизированные и просто старые предприятия.

Другим методом разделения, обладающим графической наглядностью по плотности, является подбор закона распределения с достаточной вероятностью согласия для части выборки в определенном диапазоне ее изменения, границы которого задаются исследователем. Данную процедуру в программе STATISTICA позволяет осуществлять модуль «Непараметрическая статистика / Распределения» (*Nonparametrics/Distribution*), а еще лучше — разработанные специализированные программы.

Покажем процедуру обработки на примере данных по  $Y$  табл. 3.4. После их ввода и уточнения границ и количества интервалов по правилу (2.10) на экране появится гистограмма (рис. 3.2), имеющая две ярко выраженные моды, с аппроксимирующим законом нормального распределения. По полученной гистограмме уже очевидно предварительное разделение предприятий на две группы, причем все старые предприятия попали в крайний левый столбец, а новые — в крайние правые три столбца.

Вероятность согласия аппроксимирующего нормального закона по критерию  $\chi^2$  Пирсона, подсчитанная в соответствии с правилом (2.35), здесь практически равна нулю. В то же время обработка в диапазоне 55...115, куда попадают все новые предприятия (рис. 3.3), уже дает достаточно высокую вероятность согласия ( $P_{\text{cor}} = 0,7$ ), подтверждающую правильность данной группировки. С другой стороны, обработка в диапазоне 22,0...36,5, в который попадают все старые предприятия (рис. 3.4), дополнительно выявляет еще двухмодальное смешанное распределение, не аппроксимирующееся нормальным

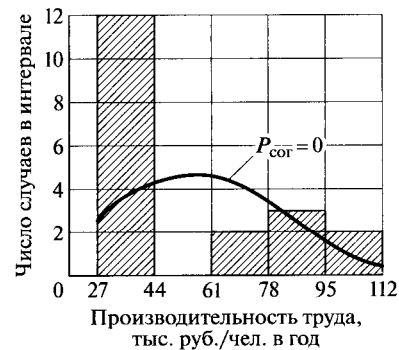


Рис. 3.2. Распределение производительности труда при  $P_{\text{cor}} = 0$

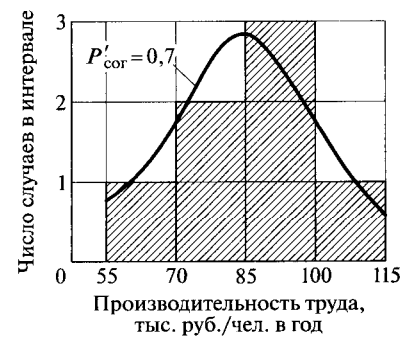


Рис. 3.3. Распределение производительности труда при  $P'_{\text{cor}} = 0,7$

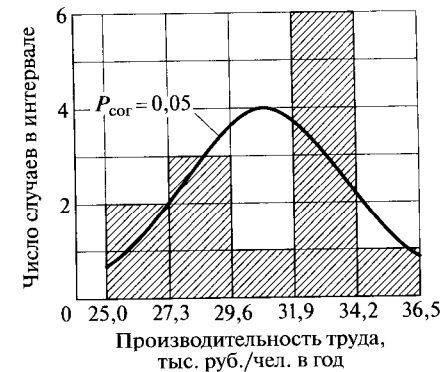


Рис. 3.4. Распределение производительности труда при  $P_{\text{cor}} = 0,05$

зконом ( $P_{\text{cor}} < 0,05$ ), и указывает на целесообразность дальнейшего разделения выборки. Окончательное разделение можно видеть из данных рис. 3.5 при последовательной обработке статистики в диапазонах 25...31 для старых предприятий, и 30,5...36,5 — для модернизированных; при этом получены высокие вероятности согласия ( $P_{\text{cor}} = 0,9$ ) с аппроксимирующими нормальными законами.

Из последних данных видно, что следует считать ошибочным включение предприятия № 17 в 3-ю группу (по распределению его следует отнести ко 2-й группе). С другой стороны, предприятие № 16 по своим характеристикам может соответствовать как 2-й, так и 3-й группам. На эти предприятия при дальнейшей обработке следует обратить особое внимание, поскольку они могут исказить окончательную статистику. Таким образом, для данного примера процедура разделения на основе подбора закона распределения с достаточной вероятностью согласия для части смешанной выборки показала бо-

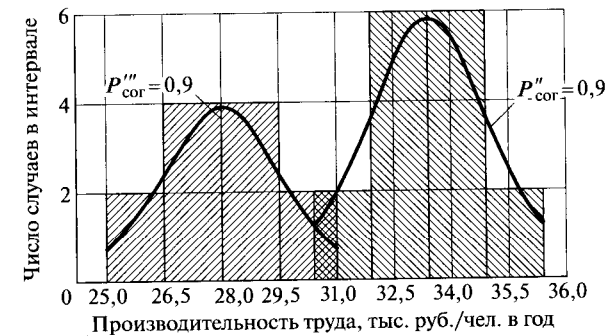


Рис. 3.5. Распределение производительности труда при  $P'''_{\text{cor}} = 0,9$  и  $P''_{\text{cor}} = 0,9$

лее высокую чувствительность по сравнению с методами кластерного анализа, поэтому ее целесообразно использовать для одномерных смешанных выборок.

### 3.4. Пример использования принципов разделения и объединения выборок для определения нормативов метода диагностирования экологической безопасности карбюраторных автомобилей при их испытаниях на ненагруженных беговых барабанах

Как известно [6], нормативные значения диагностических параметров необходимы для оценки технического состояния механизмов, систем и узлов автомобиля, осуществляемой путем сравнения с нормативами измеренных величин диагностических параметров. Поскольку для автомобильного транспорта основным распространением получила бинарная оценка состояний («Исправно» — «Неисправно»), то в качестве норматива может быть использована так называемая предельно допустимая величина параметра  $\Pi_{пр}$ . При этом нахождение измеренного параметра за пределами установленного допуска будет классифицироваться как неисправное состояние, а в пределах допуска — как исправное.

К нормативным значениям неправильно относить величины ограничений, задаваемых ГОСТами, в частности по безопасности движения и по экологии, поскольку выход за эти ограничения недопустим и классифицируется как нарушение законов государства. Очевидно, что в этом случае предельно допустимые значения должны быть более «жесткими» по сравнению с величинами ГОСТов. Также в качестве нормативов не всегда можно принимать и рекомендуемые заводами-изготовителями номинальные допуски на зазор в контактах прерывателя, тепловые зазоры клапанов, давления впрыска дизельного топлива и другие им подобные величины, на которые ориентируются при выполнении ремонтно-регулирующих операций в целях устранения неисправностей. Таким образом, значение параметров в эксплуатации должно находиться в диапазоне от номинального  $\Pi_n$  до предельно допустимого  $\Pi_{пр}$ ; своевременный возврат в указанный диапазон осуществляется на основе принятой системы организации и технологии технического обслуживания (по номенклатуре диагностических операций и периодичности их проведения).

В соответствии с теорией распознавание образов в бинарном классе возможно, если соответствующие диапазоны распределений значений диагностического параметра  $\Pi$ , аппроксимируемые теоретическими законами  $f(\Pi/D_1)$  для исправного состояния  $D_1$  и  $f(\Pi/D_2)$ , а для неисправного состояния  $D_2$ , достаточно разделены

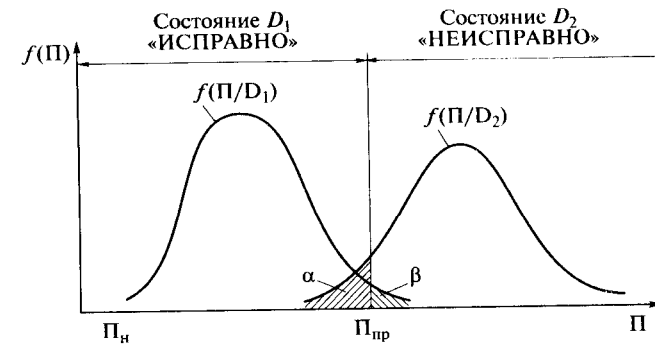


Рис. 3.6. Схема распределения значений диагностического параметра  $\Pi$  для исправного  $D_1$  и неисправного  $D_2$  состояний и появления вероятностей ошибок постановки диагноза 1-го ( $\alpha$ ) и 2-го ( $\beta$ ) родов при введении относительного норматива  $\Pi_{пр}$

и незначительно пересекаются друг с другом (рис. 3.6). (Полное разделение и отсутствие пересечений может иметь место только при наличии промежуточных (одного или нескольких) классов состояний.) Из рис. 3.6 можно видеть, что в этом случае на практике возможно относительное распознавание, поскольку при введении любого нормативного значения  $\Pi_{пр}$ , разграничивающего диапазоны состояний, неизбежно появление ошибок 1-го рода («пропуск неисправности») и 2-го рода («ложная неисправность»).

Вероятности этих ошибок  $\alpha$  и  $\beta$  (заштрихованные области на рис. 3.6) могут меняться в зависимости от установления величины  $\Pi_{пр}$ . Так, для наиболее ответственных систем и механизмов, отказы которых могут привести к значительным нежелательным в социальном или экономическом отношении последствиям (например, дорожные отказы тормозов), целесообразно минимизировать или практически исключить вероятность ошибки 1-го рода  $\alpha$ . Это может быть достигнуто за счет более «жесткого» ограничения допустимого диапазона рассеивания значений параметра для исправного состояния, аппроксимированного теоретическим законом  $f(\Pi/D_1)$ . При этом неизбежно существенное увеличение вероятности ошибок 2-го рода  $\beta$  и связанных с ними преждевременных ремонтно-регулирующих воздействий, хотя в целом это приемлемо с социальных и экономических позиций.

Для менее ответственных систем и механизмов, особенно для тех, обслуживание и ремонт которых связаны с большими материальными и трудовыми затратами (например, ремонт цилиндро-поршневой группы двигателя, замена топливного насоса высокого давления (ТНВД) дизеля), целесообразна минимизация или практическое исключение вероятности ошибок 2-го рода  $\beta$ . Это достигается путем менее «жесткого» ограничения теоретического закона рассеивания

$f(\Pi/D_1)$  (на рис. 3.6 норматив  $\Pi_{пр}$  смещается вправо, в сторону увеличения), а значительная вероятность ошибок 1-го рода  $\alpha$  и связанных с ними нежелательных последствий в данном случае будет также приемлемой с экономических позиций (например, двигатель некоторый период будет эксплуатироваться с повышенным расходом масла и топлива из-за износа цилиндра-поршневой группы). Указанный период будет коротким, поскольку при дальнейшем неизбежном ухудшении состояния значение диагностического параметра выйдет за пределы менее «жесткого» ограничения.

Сформулированные ранее положения закономерны только при достаточной информативности диагностического параметра, который прежде всего будет определяться по определенному несовпадению диапазонов его рассеивания для исправного и неисправного состояний. В случае, когда эти диапазоны в значительной мере совпадают, как это показано на рис. 3.7, параметр и соответствующий ему метод диагностирования являются неинформативными и их бессмысленно использовать для эксплуатационной оценки состояний в связи с отсутствием чувствительности. Другим признаком неинформативности параметра является его нестабильность, при которой полученное на практике (эмпирическое) распределение значений параметра для исправного состояния не может быть аппроксимировано теоретическим законом  $f(\Pi)$  в связи с отсутствием естественного физически устойчивого диапазона значений параметра, характерного для исправной работы контролируемых механизмов.

Применительно к специфике эксплуатации автомобильного транспорта нормативное значение может быть найдено путем ограничения теоретического закона рассеивания для исправного состояния  $f(\Pi/D_1)$  практически целесообразными границами: с вероятностью 0,85 для ответственных узлов и механизмов («жесткое» ограничение); с вероятностью 0,95 — в остальных случаях (менее «жесткое» ограничение). Нахождение закономерности  $f(\Pi/D_1)$  на практике не вы-

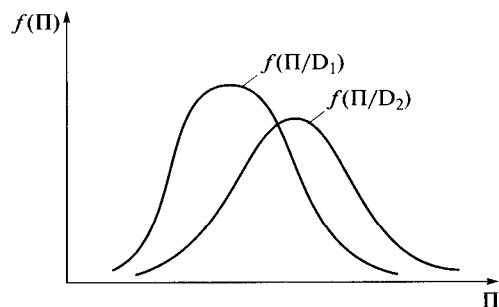


Рис. 3.7. Схема распределения значений диагностического параметра  $\Pi$  для исправного  $D_1$  и неисправного  $D_2$  состояний при отсутствии информативности метода диагностирования

являет принципиальных затруднений, поскольку автомобили, подлежащие диагностике перед очередным обслуживанием (ТО-1 или ТО-2), являются не только работоспособными, но и преимущественно исправными (по законам надежности рациональная вероятность исправного состояния при заходе на ТО находится в диапазоне 0,20...0,30).

Измерив значения диагностического параметра у статистически достаточной группы автомобилей (40...50 единиц), можно получить путем разделения этой исходной смешанной выборки (на основе подбора теоретического закона  $f(\Pi/D_1)$  для определенных границ диапазона, соответствующего исправному состоянию) необходимую выборку данных (желательно в 15...25 единиц) для оценки нормативного значения параметра. Зона исправного состояния при этом легко определяется на основе физических соображений, которые могут быть пояснены схемой, представленной на рис. 3.8.

Здесь в зависимости от физической основы параметра при нахождении норматива может быть верхнее, нижнее или двухстороннее ограничение с вероятностью 0,85 или 0,95, отражающей «жесткость» задания норматива. При этом если в первых двух случаях в качестве номинального значения (наилучшего) значения можно принимать соответственно наименьшее или наибольшее из измеренных у статистически представительной совокупности автомобилей значений параметра, то в третьем случае номинальное значение  $\Pi_n$  должно быть предварительно известно и находиться внутри диапазона рассеивания  $f(\Pi/D_1)$ , хотя не обязательно посередине. Соответственно полученный таким образом диагностический норматив должен быть представлен оператору участка диагностики следующим образом:

- допустимое значение параметра — не более  $\Pi_{пр}$ ;
- допустимое значение параметра — не менее  $\Pi_{пр}$ ;
- допустимое значение параметра — в диапазоне от  $\Pi_{пр}^{\min}$  до  $\Pi_{пр}^{\max}$ .

Диапазон рассеивания значений диагностического параметра для исправного состояния зависит от реальных условий эксплуатации, в первую очередь, от уровня технической культуры автотранспортного производства, а также от интенсивности эксплуатации конкретных агрегатов, узлов и систем автомобиля. При значительной технической культуре реально складывающийся диапазон отклонения будет меньше и на предприятии может поддерживаться более высокий уровень технического состояния автомобиля за счет обоснованного введения более «жестких» диагностических нормативов. Однако очевидно, что использование подобных нормативов для предприятия с низким уровнем технической культуры будет экономически нецелесообразно, поскольку в достаточном количестве случаев последующие ремонтно-регулирующие воздействия не будут давать положительного эффекта, а вероятность выбраковки будет повышенной. Следовательно, необходима оперативная оценка диагностических

нормативов для конкретных условий эксплуатации и их последующей корректировки по мере повышения уровня технического состояния парка автомобилей.

Таким образом, процедура нахождения диагностических нормативов для автомобилей состоит из двух этапов:

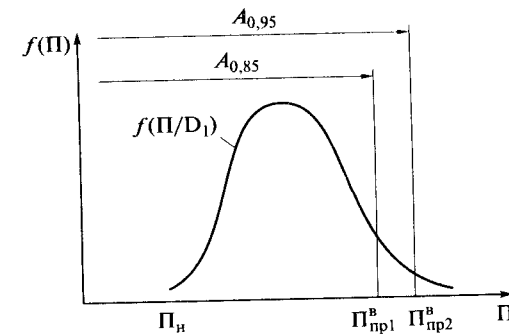
- определение теоретического закона рассеяния  $f(\Pi/D_1)$  значений диагностического параметра для исправного состояния в реальных условиях эксплуатации;

- численная оценка норматива на основе соответствующего ограничения допустимого диапазона рассеяния по найденному закону  $f(\Pi/D_1)$  (по схеме, представленной на рис. 3.8).

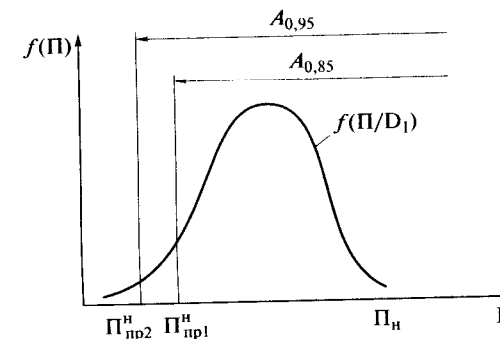
Использование компьютерных программ, обеспечивающих помимо сокращения времени трудоемких расчетов дополнительную графическую интерпретацию их результатов при поэтапной процедуре подбора теоретического закона  $f(\Pi/D_1)$  путем изменения границ диапазона смешанной выборки, резко упрощает весь этот сложный процесс определения норматива и делает его доступным для широкой производственной практики. Основная трудность при этом обычно связана со сложностью получения достаточно представительной статистики, и объединение нескольких независимых выборок (при условии подтверждения об их принадлежности к одной генеральной совокупности) в ряде случаев обеспечивает решение поставленной задачи.

Покажем это на примере разработки нового метода диагностирования карбюраторного двигателя по содержанию в отработавших газах токсичных веществ при скоростных испытаниях автомобиля на ненагруженных беговых барабанах. Данный метод приводит к резкому снижению затрат на оборудование (отпадает необходимость использования дорогого динамометрического стенда) и упрощению технологического процесса диагностирования (не требуется проводить измерения расхода топлива [6]) при одновременном повышении информативности за счет более качественной локальной оценки состояния главной дозирующей системы карбюратора, эффективности работы зажигания для конкретного двигателя, состояния его цилиндро-поршневой группы, а также за счет общей оценки уровня экологической безопасности автомобиля на нагруженных режимах. (Согласно некоторым оценкам на указанные режимы приходится около 90 % общего баланса загрязнения атмосферы крупных городов.)

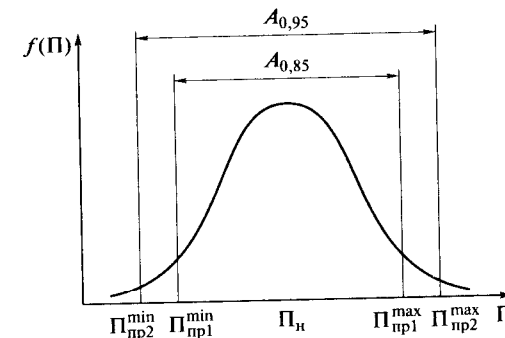
Применение именно ненагруженных беговых барабанов обеспечивает стабильность тестовых режимов испытания для отдельных автомобилей, а величина загрузки двигателя на прямой передаче при этом оказывается достаточной как для проверки качества смесеобразовательных процессов 1-й (основной) камеры карбюратора, так и для оценки состояния двигателя по выбросам  $\text{NO}_x$  (на данный компонент приходится не менее 80 % от общего баланса загрязнения



а



б



в

Рис. 3.8. Схема определения нормативного значения диагностического параметра при одностороннем верхнем (а), одностороннем нижнем (б) и двухстороннем (в) ограничениях



Результаты статистической оценки уровня экологической безопасности автомобилей ГАЗ-3110 по содержанию в выхлопных газах CO (%), CH и NO<sub>x</sub> (млн<sup>-1</sup>) при скоростных испытаниях на ненагруженных беговых барабанах в эксплуатации

Режим испытания	Номер автомобиля												
	1	2	3	4	5	6	7	8	9	10	11	12	13
40 км/ч:													
CO	0,22	0,29	2,90	0,62	4,26	0,30	1,60	3,2	0,26	1,30	0,79	0,20	0,55
CH	198	134	1500	260	302	2500	250	217	220	204	207	246	360
NO	540	381	800	490	660	1100	400	480	480	700	510	580	500
60 км/ч:													
CO	0,32	0,20	1,60	0,50	2,73	0,40	1,80	1,7	0,20	1010	0,58	0,27	0,46
CH	158	130	900	180	233	2500	200	180	160	190	160	125	220
NO	677	510	1210	770	990	1600	550	700	745	1100	760	690	840
80 км/ч:													
CO	0,39	0,23	1,20	0,40	3,51	0,45	1,50	1,5	0,66	0,90	0,58	0,36	0,64
CH	119	160	600	164	235	2500	170	160	180	180	160	117	156
NO	760	790	1600	1000	1400	2000	725	850	940	1510	1000	830	1200

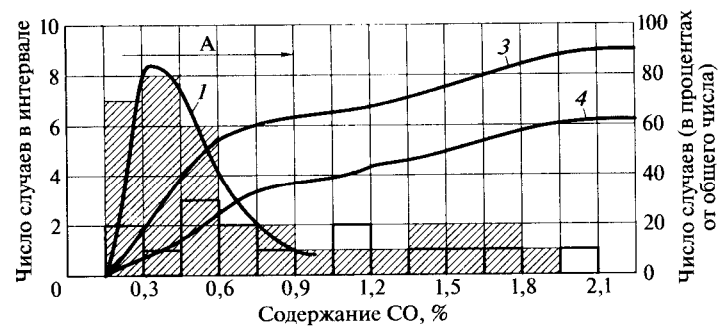
атмосферы по контролируемым для бензинового двигателя показателям). Анализ этого компонента стал возможным благодаря достижениям научно-технического прогресса.

В табл. 3.6, 3.7 приведены соответственно результаты эксплуатационных испытаний на содержание трех токсичных веществ в выхлопных газах на скоростных режимах 40, 60 и 80 км/ч по 13 автомобилям ГАЗ-3110 и восьми автомобилям М-2141, которые по отдельности не являются достаточными для определения норматива. Естественным путем повышения представительности статистики здесь является объединение выборок по трем скоростным режимам для каждой модели автомобиля с последующим объединением двух новых выборок в одну общую. Одновременно это позволит упростить процесс диагностирования за счет снижения количества нормативных показателей.

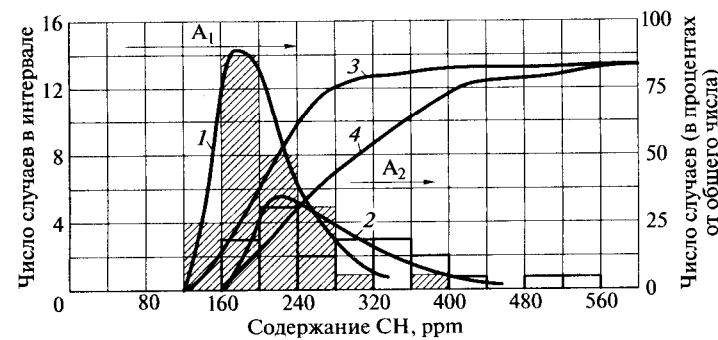
Таблица 3.7

Результаты статистической оценки уровня экологической безопасности автомобилей М-2141 по содержанию в выхлопных газах CO (%), CH и NO<sub>x</sub> (млн<sup>-1</sup>) при скоростных испытаниях на ненагруженных беговых барабанах в эксплуатации

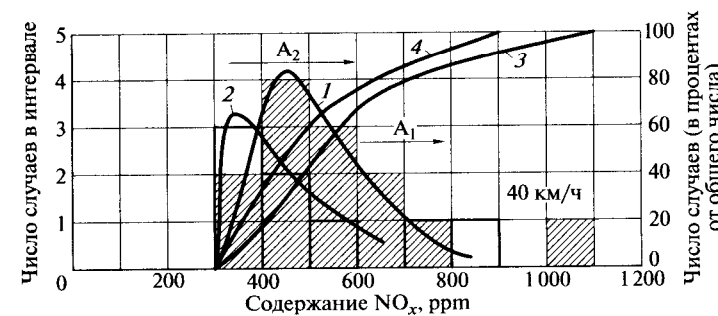
Режим испытания	Номер автомобиля							
	1	2	3	4	5	6	7	8
40 км/ч:								
CO	1,15	0,40	0,18	5,09	2,04	0,80	3,50	3,58
CH	300	200	331	347	2500	370	309	331
NO	900	760	310	440	586	310	373	480
60 км/ч:								
CO	1,20	0,60	0,74	4,50	1,71	0,58	3,10	3,06
CH	219	200	371	412	1500	330	308	371
NO	1200	1000	520	520	810	382	365	640
80 км/ч:								
CO	1,45	0,73	0,50	3,70	1,53	0,30	2,50	2,78
CH	200	180	394	520	800	260	260	474
NO	1600	1400	730	810	1100	470	540	765



а



б



в

Рис. 3.9. Распределение значений содержания CO (а), CH (б) и NO<sub>x</sub> (в) в выхлопных газах автомобилей ГАЗ-3110 (заштрихованные гистограммы, кривые 1, 3) и М-2141 (незаштрихованные гистограммы, кривые 2, 4) при их испытаниях на прямой передаче на ненагруженных беговых барабанах динамометрического стенда (40, 60 и 80 км/ч):

1, 2 — теоретические распределения для исправного состояния; 3, 4 — интегральные кривые для всей выборки; А — диапазоны допустимого значения изменения диагностического параметра для целей технологического диагностирования

Однако непосредственное использование дисперсионного анализа для трех скоростных режимов в данном случае потребовало предварительного разделения смешанных выборок путем подбора закона распределения (в данном случае гамма-распределения вида (2.32) с параметром смещения) для диапазона исправного состояния; проверка на однородность осуществлялась только по части данных, попавших в диапазон статистической обработки при подборе закона. Затем по объединенным смешанным выборкам проводилось повторное более качественное разделение и использовался критерий Стьюдента вида (3.3) для проверки однородности оставшейся части выборок по двум моделям автомобиля.

Нормативное значение параметра определялось путем разделения уже общей смешанной выборки на основе схем, представленных на рис. 3.8. Окончательные результаты представлены на рис. 3.9.

Здесь полное объединение удалось осуществить только по параметру CO, показавшему высокую информативность и чувствительность (за пределами норматива оказалось более половины автомобилей из-за неправильного подбора воздушных и топливных жиклеров 1-й камеры, а отклонение параметра от норматива у отдельных автомобилей достигало пятикратного значения).

В то же время по показателю CH объединение выборок удалось осуществить только по отдельным моделям из-за различия в работе систем зажигания (контактные и бесконтактные), а по NO<sub>x</sub> какое-либо объединение вообще оказалось невозможным ввиду увеличения выбросов при повышении скорости. На рис. 3.9 по последнему показателю приведены данные только для скорости 40 км/ч, на которой еще маловероятно зашкаливание измерительного прибора. Данная статистика оказалась малопредставительной, особенно по автомобилям «Москвич», и необходимы дополнительные исследования, тем не менее она позволила устойчиво определить потребность в замене цилиндропоршневой группы для двигателя с большим пробегом. Таким образом, статистическая обработка с использованием обоснования для объединения независимых и при этом даже смешанных выборок позволила решить поставленную задачу при небольшом количестве объектов исследования. Данный опыт может оказаться полезным при отсутствии возможности обследования больших групп автомобилей, особенно импортных.

### КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Дайте определение понятия *t*-теста для определения различия между средними величинами двух статистических выборок.
2. В чем сущность критерия Фишера для выявления различия между статистическими выборками с одинаковыми средними?
3. В каких случаях целесообразно использовать критерий Кохрана для сравнения статистических выборок?

4. Каким образом используются таблицы Фишера для оценки расхождения между статистическими выборками?

5. Назовите основные принципы кластерного анализа для проверки однородности статистической выборки и целесообразности ее разделения на отдельные статистически независимые группы.

6. Какова целесообразность объединения статистических выборок при решении задач практического характера?

7. Каким образом используется критерий Пирсона для разделения выборки на статистически независимые группы?

8. Каким образом принцип разделения смешанных выборок на статистически независимые группы может быть использован при определении нормативных значений диагностических параметров?

9. Как статистически понимается предельно допустимое значение диагностического параметра, определенного на основе разделения смешанной выборки на основе критерия Пирсона, при нормировании технологических процессов диагностирования?

## ГЛАВА 4

### Сглаживание стохастических зависимостей. Корреляционный и регрессионный анализы

#### 4.1. Сглаживание стохастических экспериментальных зависимостей по методу наименьших квадратов для случая однофакторной линейной регрессии

Как отмечалось в гл. 1, в естествознании и технике часто приходится иметь дело с функциональными или детерминистическими зависимостями, при которых какая-либо физическая величина определяется как однозначная функция одной или нескольких величин, т. е. имеет место функция

$$y = f(x, z, \dots, u), \quad (4.1)$$

в которой величина  $y$  определена значениями  $x, z, \dots, u$ .

Функциональная связь может существовать и между случайными величинами. Особенность этой связи заключается в том, что одна из них реагирует на изменение другой изменениями своего закона распределения. Такая связь называется стохастической (вероятностной). В этом случае математическое выражение зависимости одной случайной величины  $Y$  (называемой функцией отклика, параметром оптимизации, результирующим признаком или следствием) является результатом воздействия на нее нескольких других независимых переменных  $X_1, X_2, \dots, X_n$  (называемых аргументами, факториальными признаками, объясняющими факторами и т. п.) и называется уравнением регрессии, или математической моделью явления.

В общем виде такую зависимость записывают в виде

$$Y = \varphi(X_1, X_2, \dots, X_n), \quad (4.2)$$

где  $\varphi$  — форма зависимости функции отклика от объясняющих факторов.

В зависимости от формы связи различают линейную регрессию, выражаемую линейной функцией, и нелинейную регрессию, выражаемую нелинейной функцией. Смысл регрессии (от *англ.* to regress —

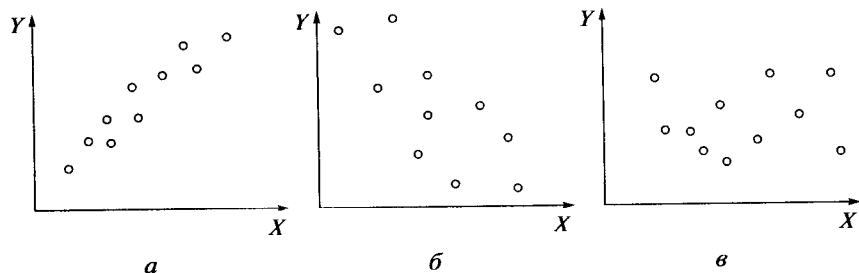


Рис. 4.1. Диаграммы рассеивания наблюдений при различной тесноте связи между случайными величинами  $X$  и  $Y$  [2]:  
 $a$  — тесная связь;  $б$  — слабая связь;  $в$  — отсутствие связи

двигаться в обратном направлении) заключается в том, что результаты наблюдений при большом числе испытаний стремятся к среднему значению.

С задачей определения уравнения регрессии связана задача определения коэффициентов корреляции (от *англ.* *correlation* — связь, соотношение, зависимость)  $r_{XY}$ , устанавливающих зависимость результативного признака от факториальных. При этом указанная зависимость может быть не только между функцией отклика и факториальными признаками, но и между самими факториальными признаками.

Разница между регрессией и корреляцией заключается в том, что основной задачей регрессионного анализа является определение формы зависимости  $Y$  от  $X_1, X_2, \dots, X_n$ , а основной задачей корреляционного анализа является определение силы этой зависимости. Получаемое в результате регрессионного анализа уравнение регрессии позволяет предсказать значение  $\hat{Y}$  в зависимости от связанных с ним независимых переменных  $X_1, X_2, \dots, X_n$  [2].

Указанные ранее положения наглядно представляются в графическом виде для случаев однофакторных зависимостей (рис. 4.1). Здесь попарные измерения значений случайных величин  $X$  и  $Y$ , обычно оказывающихся при пассивном эксперименте распределенных по нормальному закону, по масштабируемому расположению экспериментальных точек может указывать на наличие как тесной (рис. 4.1,  $a$ ), так и слабой (рис. 4.1,  $б$ ) функциональной стохастической связи или ее отсутствие (рис. 4.1,  $в$ ), хотя все подобные выводы должны быть обоснованы соответствующими математическими расчетами.

Такие расчеты обычно начинаются с задания формы связи между функцией отклика и независимой переменной, которая в общем случае для однофакторной модели обычно задается в виде полинома  $n$ -й степени или другой нелинейной зависимости вида

$$\hat{Y} = \varphi(X; a, b, c, \dots), \quad (4.3)$$

где  $a, b, c, \dots$  — неизвестные коэффициенты модели, значения которых определяются в процессе регрессионного анализа.

В статистике доказывается [2], что в соответствии с принципом максимального правдоподобия сам вид функции (4.3) и ее коэффициенты наиболее целесообразно устанавливать (подбирать из возможных вариантов) путем минимизации суммы квадратов отклонений наблюдаемых значений  $Y_i$  от сглаженных, задаваемых функцией  $\varphi(X_i; a, b, c, \dots)$ , как это схематически показано на рис. 4.2, в соответствии с условием, которое называется методом наименьших квадратов:

$$\sum_{i=1}^n [Y_i - \varphi(X_i; a, b, c, \dots)]^2 \Rightarrow \min. \quad (4.4)$$

Для определения значений  $a, b, c, \dots$ , обращающих левую часть выражения (4.4) в минимум, необходимо продифференцировать ее по  $a, b, c, \dots$  и приравнять производные нулю:

$$\begin{aligned} \sum_{i=1}^n [Y_i - \varphi(X_i; a, b, c, \dots)] \left( \frac{\partial \varphi}{\partial a} \right)_i &= 0; \\ \sum_{i=1}^n [Y_i - \varphi(X_i; a, b, c, \dots)] \left( \frac{\partial \varphi}{\partial b} \right)_i &= 0; \\ \sum_{i=1}^n [Y_i - \varphi(X_i; a, b, c, \dots)] \left( \frac{\partial \varphi}{\partial c} \right)_i &= 0, \end{aligned} \quad (4.5)$$

где  $\left( \frac{\partial \varphi}{\partial a} \right)_i$  — значение частной производной функции  $\varphi$  по параметру  $a$  в точке  $X_i$ ,  $\left( \frac{\partial \varphi}{\partial a} \right)_i = \varphi'_a(X_i; a, b, c, \dots)$ .

Система уравнений содержит столько же уравнений, сколько неизвестных  $a, b, c, \dots$ , и для ее решения необходимо задаться конк-

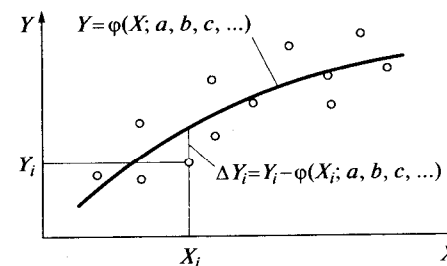


Рис. 4.2. Схема сглаживания стохастической зависимости между случайными величинами на основе метода наименьших квадратов

ретным значением функции  $\varphi$ . Наиболее простым и популярным для практического применения способом является задание линейной зависимости вида

$$Y = \varphi(X; a, b) = aX + b. \quad (4.6)$$

Для такой зависимости имеем:

$$\frac{\partial \varphi}{\partial a} = X; \left( \frac{\partial \varphi}{\partial a} \right)_i = X_i; \frac{\partial \varphi}{\partial b} = 1; \left( \frac{\partial \varphi}{\partial b} \right)_i = 1. \quad (4.7)$$

Подставляя последние значения в формулы (4.5), получаем два уравнения для определения коэффициентов  $a$  и  $b$ :

$$\sum_{i=1}^n [Y_i - (aX_i + b)]X_i = 0; \sum_{i=1}^n [Y_i - (aX_i + b)] = 0,$$

или, раскрывая скобки и производя суммирование, получаем:

$$\begin{aligned} \sum_{i=1}^n Y_i - a \sum_{i=1}^n X_i - bn &= 0; \\ \sum_{i=1}^n X_i Y_i - a \sum_{i=1}^n X_i^2 - b \sum_{i=1}^n X_i &= 0. \end{aligned} \quad (4.8)$$

Система уравнений (4.8) уже дает схему непосредственного вычисления коэффициентов  $a$  и  $b$  линейной модели на основе выполнения простейших статистических расчетов над случайными величинами  $X$  и  $Y$  (их суммирования, а также суммирования попарных произведений и квадратов значений  $X_i$ ). В то же время при делении на  $n$  и после проведения соответствующих преобразований к привычным статистическим характеристикам — выборочным средним  $\bar{X}^*$  и  $\bar{Y}^*$  и дисперсиям  $S_x^2$  и  $S_y^2$  — получим следующее решение:

$$a = \frac{K_{XY}^*}{S_x^2}; b = \bar{Y}^* - a\bar{X}^* = \bar{Y}^* - \frac{K_{XY}^*}{S_x^2} \bar{X}^*, \quad (4.9)$$

где  $K_{XY}^* = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}^*)(Y_i - \bar{Y}^*)$  называется *коэффициентом ковариации*, или *моментом связи* между случайными величинами (обозначением величины его математического ожидания может являться  $\text{cov}(X, Y)$  [2]).

Для независимых случайных величин значение коэффициента ковариации равно нулю; для зависимых теоретически оно будет определяться соотношением

$$\text{cov}(X, Y) = M(XY) - M(X)M(Y). \quad (4.10)$$

Величина  $\text{cov}(X, Y)$  зависит от единиц измерения  $X$  и  $Y$  и сама по себе еще не может служить показателем связи. Для получения безразмерного показателя рассматривают ковариации нормирован-

ных отклонений, получаемые при делении на стандартные отклонения  $\sigma_X$  и  $\sigma_Y$  случайных величин:

$$r_{XY} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}. \quad (4.11)$$

Показатель  $r_{XY}$  называется *коэффициентом корреляции* случайных величин  $X$  и  $Y$ , значение которого может находиться в диапазоне от  $-1$  до  $+1$ . Для независимых величин он равен нулю, поскольку для них  $\text{cov}(X, Y) = 0$ . Однако обратного заключения сделать нельзя: величины могут быть связаны даже функционально, но при нахождении центра распределения независимой переменной около начала координат коэффициент корреляции окажется равным нулю (в данном случае рекомендуется произвести повторные расчеты при переносе начала координат). В общем случае коэффициент корреляции между линейными функциями величин  $X$  и  $Y$  будет по абсолютному значению равен коэффициенту корреляции между самими этими величинами.

Выборочное значение коэффициента корреляции определяется по выражению

$$r_{XY}^* = \frac{K_{XY}^*}{S_X S_Y}, \quad (4.12)$$

и его значимость проверяется по критерию Стьюдента [2]:

$$t_{q,n-2} = \frac{r_{XY}^* \sqrt{n-2}}{\sqrt{1-(r_{XY}^*)^2}}. \quad (4.13)$$

Соответственно может быть найден и доверительный интервал для истинного значения коэффициента корреляции [2]:

$$r_{XY}^* - t_{q,n-2} \frac{1-(r_{XY}^*)^2}{\sqrt{n-2}} < r_{XY} < r_{XY}^* + t_{q,n-2} \frac{1-(r_{XY}^*)^2}{\sqrt{n-2}}. \quad (4.14)$$

Полученное по выражению (4.12) выборочное значение коэффициента корреляции  $r_{XY}^*$  используется при оценке коэффициента  $a$  линейного уравнения регрессии (4.6) через преобразования выражения (4.9):

$$a^* = \frac{K_{XY}^*}{S_x^2} = r_{XY}^* \frac{S_Y}{S_x} = a_{Y/X}^*. \quad (4.15)$$

Последнее обозначение называется *коэффициентом регрессии*, для которого также могут быть найдены доверительные интервалы [2]:

$$a_{Y/X}^* - t_{q,n-2} \frac{S_Y}{S_x} \frac{1-(r_{XY}^*)^2}{\sqrt{n-2}} < \beta_{Y/X} < a_{Y/X}^* + t_{q,n-2} \frac{S_Y}{S_x} \frac{1-(r_{XY}^*)^2}{\sqrt{n-2}}. \quad (4.16)$$

Следует иметь в виду, что при перемене мест данных  $X$  и  $Y$ , т.е. при изменении условия зависимой и независимой переменной, будет получена совсем другая зависимость с коэффициентом регрессии, определяемым по выражению

$$\beta_{X/Y} = r_{XY}^* \frac{S_X}{S_Y}, \quad (4.17)$$

а также с другим значением свободного члена линейного уравнения. На графике указанные линии регрессии будут пересекаться в точке с координатами, равными значениям средних, причем угол пересечения будет увеличиваться при снижении значения коэффициента корреляции. В предельных случаях, при  $r_{XY} = \pm 1$ , т.е. при наличии жесткой (детерминированной) функциональной связи, обе прямые сливаются в одну [2].

#### 4.2. Коэффициент детерминации и его использование для оценки точности и адекватности однофакторной модели линейной регрессии

Обобщая зависимости (4.6), (4.9), (4.11), (4.12) и (4.15) и переходя к математическому ожиданию  $m_X$  и  $m_Y$  случайных величин и их стандартным отклонениям  $\sigma_X$  и  $\sigma_Y$ , уравнение минимизирующей (для суммы квадратов отклонений) прямой приближенной регрессии можно записать в виде [2]

$$\hat{Y} = m_Y + r_{XY} \frac{\sigma_Y}{\sigma_X} (X - m_X). \quad (4.18)$$

Показателем качества найденной модели регрессии будет являться дисперсия ошибок  $\delta_e^2$ , под которой понимают усредненное значение (математическое ожидание) суммы квадратов отклонений функции отклика  $Y$  от предсказанных значений, определяемых уравнением (4.18). В статистике доказывается [2], что указанная величина будет равна:

$$\sigma_e^2 = M(Y - \hat{Y})^2 = \sigma_Y^2 (1 - r_{XY}^2). \quad (4.19)$$

Она называется остаточной, не коррелированной с объясняющей переменной  $X$ , дисперсией и задает точность прогноза значений  $\hat{Y}$  с обычно принятой в статистике 95%-ной доверительной вероятностью.

Соотношение (4.19) имеет важное значение при разложении общей дисперсии  $\sigma_Y^2$  на ее составляющие, как в дисперсионном анализе. В статистике также доказывается [2], что в последнем случае имеет

место следующая зависимость, легко получаемая при соответствующей интерпретации выражения (4.19):

$$\sigma_Y^2 = r_{XY}^2 \sigma_Y^2 + \sigma_e^2 (1 - r_{XY}^2). \quad (4.20)$$

Отсюда видно, что  $r_{XY}^2$  (статистическая оценка  $(r_{XY}^*)^2$ ) представляет собой ту долю дисперсии величины  $Y$ , которая обусловлена линейно прогнозируемой при каждом значении  $X$  компонентой  $Y$ , и на остаточную и не коррелированную с  $X$  компонентой падает доля дисперсии, определяемая через так называемый коэффициент неопределенности  $(1 - r_{XY}^2)$ .

Величина  $r_{XY}^2$  также имеет важное значение в статистике и называется *коэффициентом детерминации*. В силу свойства (4.19) статистическая оценка коэффициента детерминации (и коэффициента множественной детерминации) легко вычисляется на основе опытных данных для регрессионных моделей любых видов (однофакторных нелинейных и многофакторных линейных и нелинейных) по выражению

$$(r_{XY}^*)^2 = \frac{S_Y^2 - S_e^2}{S_Y^2} = 1 - \frac{S_e^2}{S_Y^2} \quad (4.21)$$

или же путем непосредственного использования суммы квадратов отклонений, как это принято в дисперсионном анализе (см. табл. 4.3, столбец 2), для всей выборки  $Q_{\text{общ}}$  и для регрессионной модели (объясненные отклонения)  $Q_{\text{объясн}}$ , которые вычисляются в процессе промежуточных расчетов:

$$(r_{XY}^*)^2 = \frac{Q_{\text{общ}} - Q_{\text{ост}}}{Q_{\text{общ}}} = \frac{Q_{\text{общ}} - (Q_{\text{общ}} - Q_{\text{объясн}})}{Q_{\text{общ}}} = \frac{Q_{\text{объясн}}}{Q_{\text{общ}}}. \quad (4.22)$$

Графическая интерпретация взаимосвязей между всеми описанными ранее показателями для случая линейной регрессии между случайными величинами  $Y$  и  $X$  и при условии нормальных законов распределения для этих величин приведена на рис. 4.3. Здесь одномерное нормальное распределение  $\phi(Y)$  с параметрами  $m_Y$  и  $\sigma_Y$  выражает априорное знание о показателе  $Y$ . (В общем случае величины  $X$  и  $Y$  совместно дают двумерное нормальное распределение; показатели рассеяния этого распределения графически могут быть представлены в виде эллипса, главные оси которого имеют все больший поворот относительно координатных осей с увеличением силы корреляционной связи [2].) Условное нормальное распределение  $\phi(Y/X)$  характеризует уже неопределенность показателя  $Y$  не вообще, а при условии, что фактор  $X$  принял конкретное значение  $\hat{X}$ .

Линия условных средних  $\hat{Y}$  (4.18) при нормальном законе двумерного распределения всегда является прямой линией (обратное утверждение неверно). Кривая плотности распределения  $\phi(Y/X)$  предполагается симметрично по отношению к  $\hat{Y}$ , соответствующих конкретному  $\hat{X}$ . При наличии статистической связи ( $r_{XY} \neq 0$ ) условное

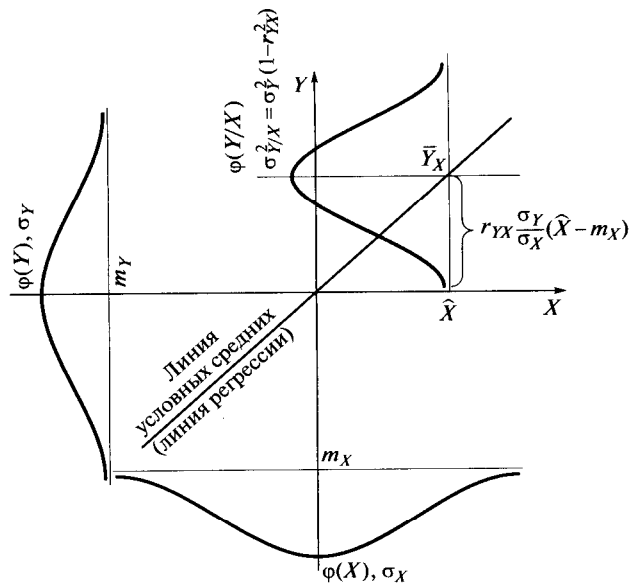


Рис. 4.3. Схема двумерной регрессии [2]

распределение  $\varphi(Y/X)$  всегда имеет меньший «размах» по сравнению с априорным распределением  $\varphi(Y)$ ; этот размах тем меньше, чем ближе значение  $|r_{XY}|$  к единице.

В целом для каждого фиксированного значения  $\hat{X}$  положение условной плотности распределения и значения его среднеквадратического отклонения свое. Этим изменением закона распределения случайная величина  $Y$ , как уже отмечалось ранее, реагирует на изменение статистически связанной с ней величины  $X$  как аргумента. Условная функция распределения описывает остаточную неопределенность показателя  $Y$ , не раскрытую с помощью фактора  $X$ . При этом сама величина  $\hat{X}$  считается известной точно [2]. Указанная неопределенность в конечном счете отражается на точности прогноза (предсказанного значения) показателя  $Y$  в зависимости от величины  $\hat{X}$  с принятой в статистике 95%-ной доверительной вероятностью, которая графически (рис. 4.4) представляет собой двойной раструб с самой узкой его частью вблизи зоны «центра масс» двумерного нормального распределения.

Связь  $X$  и  $Y$  через  $r_{XY}$  является основой всех косвенных измерений и оценок в технике и экономике. При этом относительная точность косвенных оценок, как это можно видеть из выражений (4.18) и (4.21), зависит только от величины  $r_{XY}$ , что отражено в табл. 4.1. Здесь оценка степени тесноты статистической связи базируется на

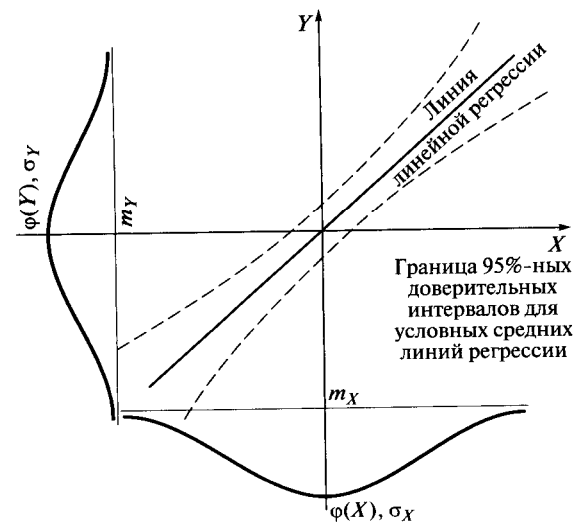


Рис. 4.4. Раструб прогноза предсказанных значений для показателя отклика

инженерных требованиях к точности косвенных практически приемлемых оценок. В инженерной практике считается допустимой априорная погрешность 10 %, при которой практически при косвенной оценке можно знать величину  $Y$  по измеренной величине  $X$ . Граница  $r_{XY} = 0,7$  выбрана из условия не хуже, чем 50 %-ная относительная погрешность. Если  $r_{XY} = 0,3$ , то с инженерной точки зрения такими статистическими связями можно пренебречь [2]. Более распространенным является граничное значение  $r_{XY} = 0,5$ , ниже которого установленную статистическую связь не рекомендуется использовать для инженерных расчетов.

Таблица 4.1

Относительная точность косвенных оценок в зависимости от значений коэффициента корреляции [2]

Диапазон значений коэффициента корреляции	Относительная остаточная ошибка, %	Теснота статистической связи
$0,95 \leq  r_{XY}  < 1,00$	До 10	Очень тесная (сильная)
$0,70 \leq  r_{XY}  < 0,95$	До 50	Тесная
$0,30 \leq  r_{XY}  < 0,70$	До 90	Слабая
$0 \leq  r_{XY}  < 0,30$	~100	Практически отсутствует

Приведенные ранее теоретические и расчетно-практические соотношения для коэффициентов детерминации и корреляции дают основу для оценки точности и достоверности регрессионной модели, без которых невозможно ее принятие и использование вообще.

Здесь прежде всего имеет значение разложение на составляющие общей дисперсии вида (4.20), которым можно дать толкование, сходное с дисперсионным анализом (остаточная как внутригрупповая и объясненная как межгрупповая дисперсии), и оценить неслучайность такого разделения с помощью критерия Фишера (3.15), который в данном случае может быть выражен через коэффициент детерминации, с учетом степеней свободы, подсчитываемым по правилам дисперсионного анализа (табл. 3.3, столбец 3) с соответствующими корректировками, обусловленными количеством  $m$  коэффициентов модели вместо количества групп (для случая линейной регрессии  $m = 2$ ).

В результате получим соотношение

$$F_{q, m-1, n-m} = \frac{Q_{\text{объясн}}/(m-1)}{Q_{\text{ост}}/(n-m)} = \frac{Q_{\text{общ}}(r_{XY}^*)^2(n-m)}{(m-1)Q_{\text{общ}}(1-(r_{XY}^*)^2)} = \\ = F_{q, 1, n-2} = \frac{(r_{XY}^*)^2(n-2)}{1-(r_{XY}^*)^2}. \quad (4.23)$$

Соотношение (4.23) может быть использовано для всех видов статистической связи независимо от их линейности или нелинейности, так же как и толкование коэффициента детерминации  $r_{XY}$  как отношения объясненной статистической связью части дисперсии к полной дисперсии. Проверка достоверности регрессионной модели по выражению (4.23) называется *проверкой на адекватность* и иногда интерпретируется как оценка ее качества по соотношениям между уровнями полезного сигнала и шума [2].

Точность модели принято оценивать через значимость величины коэффициента регрессии на основе  $t$ -критерия [2], который в первом приближении (более подробно об этом будет сказано далее) с учетом выражения (4.16) [2] для линейной модели может быть представлен в виде

$$t_{q, n-2} = \frac{r_{XY}^* \sqrt{n-2}}{1-(r_{XY}^*)^2}, \quad (4.24)$$

т.е. имеет место то же выражение, что и для оценки значимости коэффициента корреляции (4.13) [2], что естественно ввиду зависимостей (4.15) и (4.17).

В общем случае при получении недостаточного уровня значимости коэффициента регрессии считается равным нулю, что для однофакторной линейной модели будет соответствовать отсутствию связи, однако в последнем случае такое заключение может быть сделано

ранее на основе критерия Фишера (4.23). В программе STATISTICA производится оценка уровня значимости и для свободного члена уравнения регрессии, однако последнее, вероятно, обусловлено общими методикой оценки значимости линейных коэффициентов модели любой сложности, которая будет изложена далее, и не применяется в отечественных исследованиях [2] в связи с тем, что при нахождении оценки свободного члена около начала координат расчетный уровень значимости, как уже отмечалось ранее, всегда будет значительно заниженным.

#### 4.3. Матричные способы определения коэффициентов уравнений многофакторных регрессий, представляемых полиномами $n$ -й степени

В общем случае в соответствии с выражением (4.2) может иметь место многофакторная регрессия, причем каждый фактор может быть выражен полиномом  $n$ -й степени:

$$\hat{Y} = B_0 + B_1 X_1 + \dots + B_n X_1^n + \dots + B_{1K} X_K + \dots + B_{nK} X_K^n + \\ + B_{1,2} X_1 X_2 + \dots + B_{1,K} X_1 X_K + \dots + B_{K-1,K} X_{K-1} X_K + \dots \quad (4.25)$$

при учете связей для комбинаций из трех параметров первой степени или комбинаций из двух и более параметров  $n$ -й степени. (Обычно принято анализировать только парные комбинации параметров первой степени, как это отражено записью (4.25) для уравнения регрессии.) В последней записи также использованы обозначения  $B$  для коэффициентов модели, принятые в программе STATISTICA, хотя для их нумерации в указанной программе используется натуральный ряд чисел.

Для определения искомых коэффициентов модели вида (4.25) применяется, как и для простой регрессии, метод наименьших квадратов. Этот метод в конечном счете на основе зависимостей (4.5) приводит к так называемой системе нормальных алгебраических линейных уравнений (первой степени), которые принято решать по методу Крамера [2]. (По методу Крамера для определения неизвестного вычисляется определитель матрицы, составленной из величин коэффициентов при неизвестных членах, на который затем делится определитель скорректированной исходной матрицы, получаемой при замене столбца коэффициентов при искомом неизвестном на столбец из свободных членов.) Решение такой системы уже для трех неизвестных приводит к достаточно трудоемким вычислениям, объем которых в дальнейшем еще более возрастает из-за необходимости разложения полученного сложного определителя на сумму более простых определителей 2-го или 3-го порядков, имеющих простые методы вычисления.



В этих условиях для вычисления коэффициентов регрессии более удобными оказались матричные подходы, легко реализуемые в компьютерных алгоритмах [2] и в максимальной степени используемые программой STATISTICA. Данная программа позволяет получить методом итерации (последовательным подбором) для однофакторной зависимости необходимые коэффициенты задаваемого уравнения любой сложности, а затем на основе зависимости (4.22) и, возможно, (4.23) определить коэффициент детерминации и оценить адекватность модели; однако при этом будут отсутствовать оценки значимости коэффициентов (даже если их определение возможно для линейных моделей), без которых ее дальнейшее применение будет считаться недостаточно обоснованным. Поэтому в любом случае необходимо сначала получить наиболее подходящую модель с оценкой значимости для ее коэффициентов, а затем уже подбирать другие чисто нелинейные модели, физически больше соответствующие изучаемому процессу, ориентируясь по получаемым при этом коэффициентам детерминации и другим статистическим характеристикам на показатели предыдущей модели, с тем чтобы хотя бы косвенно дать оценку точности (и значимости) чисто нелинейной модели.

Методически для освоения матричного подхода полезно вспомнить некоторые операции, которые можно производить с матрицами. В общем случае под матрицей понимается прямоугольная таблица, составленная из элементов произвольной формы. Элементы матрицы располагаются в строки и столбцы (иногда их называют колонками), которые вместе называют рядами матрицы. Элементы матрицы часто обозначают двойным индексом  $a_{ij}$ , где первый индекс  $i$  обозначает номер строки, а второй  $j$  — номер столбца, на пересечении которых находится элемент  $a_{ij}$ .

В символическом обозначении матрица обычно заключается в круглые скобки или (реже) в двойные вертикальные черточки. Используются также краткими обозначениями матриц:  $(a_{ij})$  или  $\|a_{ij}\|$ , а также обозначают ее строчным символом увеличенного размера [2], например  $X$  для матрицы случайных наблюдений  $X_i$  независимого фактора, представленных в табл. 3.4, и  $Y$  для матрицы наблюдений  $Y_i$  фактора отклика. Основной операцией при матричном подходе определения коэффициентов модели будет являться перемножение матриц. Оно возможно, если размер строки первой матрицы будет равен размеру столбца второй матрицы. В результате получается новая матрица, у которой количество строк будет определяться размерами 1-й матрицы, а количество столбцов — размерами 2-й матрицы.

Для примера умножим матрицу  $(A)$  размером  $(M \cdot N)$  на матрицу  $(B)$  размером  $(N \cdot L)$ . В результате получим матрицу  $(C)$  размером  $(M \cdot L)$ , элементы  $C_{ML}$  которой будут определяться как сумма попарных произведений элементов строки 1-й матрицы на соответствующие элементы столбца 2-й матрицы. Данные процедуры можно записать следующим образом:

$$(A)(B) = (C); (MN)(NL) = (ML); \quad (4.26)$$

$$C_{ML} = \sum_{N=1}^N a_{MN} b_{NL}. \quad (4.27)$$

Из соотношений (4.26), (4.27) следует, что при перемене мест множителей (такое возможно для квадратных матриц одинакового размера) в результате будет получена совсем другая матрица, т. е.

$$(A)(B) \neq (B)(A). \quad (4.28)$$

Особенности указанных ранее процедур можно также видеть на примере перемножения матриц  $(1 \cdot 10)(10 \cdot 1)$ , элементы которых представлены натуральным рядом чисел от 1 до 10, и наоборот. Если в первом случае будет получена матрица размером  $(1 \cdot 1)$ , единственным элементом которой равен сумме квадратов натуральных чисел, то во втором — школьная таблица умножения.

Для квадратных матриц, широко используемых при решении системы линейных алгебраических уравнений, обычно вычисляется определитель. Важными свойствами определителя являются: неизменность его величины при прибавлении (или вычитании) к элементу какой-либо строки матрицы элементов другой строки или линейной комбинации других строк; равенство нулю, если элементы двух его строк равны или пропорциональны или одна из строк является линейной комбинацией каких-либо остальных строк.

При планировании активного многофакторного эксперимента, особенности которого рассмотрены в гл. 5, широко используются так называемые матрицы Адамара, элементами которых являются значения  $\pm 1$  и 0. Среди них особое значение имеет так называемая *единичная* квадратная матрица, у которой диагональные элементы равны единице, а все остальные — нулю. Такая матрица имеет обозначение  $E$ , а ее определитель равен единице. При помощи этой матрицы обеспечивается решение системы нормальных алгебраических уравнений 1-й степени на основе метода, который может быть пояснен на примере простейшей системы из двух неизвестных:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1; \\ a_{21}x_1 + a_{22}x_2 &= b_2. \end{aligned} \quad (4.29)$$

В матричной форме данное уравнение может быть записано в виде

$$AX = B, \quad (4.30)$$

или

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}. \quad (4.31)$$

Для получения решения матричного уравнения (4.30) необходимо осуществить такое его преобразование, чтобы в левой части оказалось произведение единичной матрицы на матрицу  $X$ . Для этого обе части этого уравнения необходимо умножить на так называемую обратную матрицу  $A^{-1}$  (по отношению к матрице коэффициентов; при этом произведение исходной и обратной матриц независимо от их последовательности будет приводить к единичной матрице  $E$ ), т.е. выполнить преобразование:

$$A^{-1}AX = EX = A^{-1}B, \quad (4.32)$$

в результате которого левая часть выражения (4.32) примет вид

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \quad (4.33)$$

Из равенства (4.33) видно, что после раскрытия символьных обозначений и перемножения матриц правой части уравнения (4.32) в столбце результирующей матрицы будут записаны непосредственные решения системы линейных уравнений (4.29).

Основная трудность данного подхода при всей его наглядности будет связана с вычислением обратной матрицы  $A^{-1}$ , при котором в компьютерных программах, в том числе и в программе STATISTICA, реализуется достаточно сложный алгоритм, в процессе которого сначала необходимо составить так называемую транспонированную матрицу  $A^T$ , у которой строки исходной матрицы будут являться столбцами, и наоборот; вычисляется определитель транспонированной (или исходной) матрицы (они равны); составляется присоединенная матрица, элементы которой равны алгебраическим дополнениям транспонированной матрицы; осуществляется деление присоединенной матрицы на определитель [2].

В то же время на основе единичной матрицы возможен более простой алгоритм решения системы уравнений (4.29)... (4.31), удобный для «ручных» расчетов, который можно видеть на следующем простом примере линейной системы из двух неизвестных:

$$\begin{aligned} x_1 + 2x_2 &= 1; \\ x_1 + 3x_2 &= 0. \end{aligned} \quad (4.34)$$

К. Гауссом предложена более простая по сравнению с (4.34) запись системы в виде расширенной матрицы, в которой неизвестные подразумеваются, как и знак равенства, а показываются только коэффициенты при неизвестных и свободные члены. Для решения необходимо преобразовать расширенную матрицу таким образом, чтобы в ее левой части получить единичную матрицу, тогда в правой части будет записан готовый ответ:

$$\left| \begin{array}{cc|c} 1 & 2 & 1 \\ 1 & 3 & 0 \end{array} \right| \sim \left| \begin{array}{cc|c} 1 & 2 & 1 \\ 0 & 1 & -1 \end{array} \right| \sim \left| \begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & -1 \end{array} \right|. \quad (4.35)$$

Здесь на первом этапе преобразования из второй строки расширенной матрицы вычли первую, а на втором этапе из первой строки вычли вторую, умноженную на два.

Аналогичный алгоритм может быть использован и при нахождении обратной матрицы (без вычисления определителей и присоединенных матриц). Для этого исходная матрица дополняется единичной, записываемой в правую часть расширенной матрицы, которая затем преобразуется на основе деления строк и их вычитания с соответствующими множителями до вида, где в левой части уже будет записана единичная матрица, а в правой — искомая обратная.

Покажем правильность данного алгоритма на параллельном расчете примера, заимствованного из [2]:

$$\begin{aligned} & \left| \begin{array}{ccc|ccc} 2 & 2 & 3 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ -1 & 2 & 1 & 0 & 0 & 1 \end{array} \right| \sim \left| \begin{array}{ccc|ccc} 1 & 1 & 1,5 & 0,5 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 1 & -2 & -1 & 0 & 0 & -1 \end{array} \right| \sim \\ & \sim \left| \begin{array}{ccc|ccc} 1 & 1 & 1,5 & 0,5 & 0 & 0 \\ 0 & -2 & -1,5 & -0,5 & 1 & 0 \\ 0 & -3 & -2,5 & -0,5 & 0 & -1 \end{array} \right| \sim \left| \begin{array}{ccc|ccc} 1 & 1 & 1,5 & 0,5 & 0 & 0 \\ 0 & 1 & 0,75 & 0,25 & -0,5 & 0 \\ 0 & 1,2 & 1 & 0,2 & 0 & 0,4 \end{array} \right| \sim \\ & \sim \left| \begin{array}{ccc|ccc} 1 & 0 & 0,75 & 0,25 & 0,5 & 0 \\ 0 & 1 & 0,75 & 0,25 & -0,5 & 0 \\ 0 & 0 & 0,1 & -0,1 & 0,6 & 0,4 \end{array} \right| \sim \left| \begin{array}{ccc|ccc} 1 & 0 & 0,75 & 0,25 & 0,5 & 0 \\ 0 & 1 & 0,75 & 0,25 & -0,5 & 0 \\ 0 & 0 & 1 & -1 & 6 & 4 \end{array} \right| \sim \\ & \sim \left| \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -4 & -3 \\ 0 & 1 & 0 & 1 & -5 & -3 \\ 0 & 0 & 1 & -1 & 6 & 4 \end{array} \right|. \quad (4.36) \end{aligned}$$

Здесь на первом этапе 1-я и 3-я строки были разделены на (2) и (-1) соответственно; на втором этапе из 2-й и 3-й строк вычли 1-ю; на третьем этапе 2-ю и 3-ю строки разделили на (-2) и (-2,5) соответственно; на четвертом этапе было произведено вычитание 2-й строки из 1-й, а также из 3-й при умножении 2-й на (1,25); на пятом этапе 3-я строка была разделена на (0,1); на последнем (шестом) этапе из 1-й и 2-й строк была вычтена 3-я, умноженная на (0,75). Правильность полученного решения можно проверить перемножением исходной и обратной матриц, или наоборот (в любом случае будет получена единичная матрица).

С учетом изложенного ранее представляется несложным вывод матричного уравнения для определения коэффициентов регрессии стохастических зависимостей вида (4.25). В простейшем случае, для однофакторной линейной зависимости, истинное (неизвестное) зна-

чение  $Y$  может быть выражено с помощью его расчетного значения  $\hat{Y}$  и ошибки  $\varepsilon$ :

$$Y_{\text{ист}} = \hat{Y} + \varepsilon, \quad (4.37)$$

где  $\varepsilon$  — остаток, иногда называемый возмущающей переменной.

В свою очередь, расчетное линейное уравнение регрессии можно представить в виде

$$\hat{Y} = B_0^* X_0 + B_1^* X_1, \quad (4.38)$$

где  $B_0^*$  — статистическая оценка коэффициента при фиктивной переменной (оценка свободного члена);  $X_0$  — фиктивная переменная, равная единице;  $B_1^*$  — статистическая оценка коэффициента при независимой переменной;  $X_1$  — заданное значение независимой переменной;

Для определения неизвестных коэффициентов расчетного уравнения регрессии  $B_0^*$  и  $B_1^*$  в соответствии с изложенными ранее принципами составляется выражение возмущающей переменной  $\varepsilon$  и выдвигается требование минимума суммы квадратов указанной переменной по всем опытам:

$$\sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (B_0^* X_{0i} + B_1^* X_{1i} - Y_i)^2 \rightarrow \min. \quad (4.39)$$

Дифференцируя выражение (4.39) по неизвестным  $B_0^*$  и  $B_1^*$ , приравнявая полученные производные нулю и раскрывая скобки, получим следующую систему двух нормальных уравнений:

$$B_0^* \sum_{i=1}^n X_{0i} X_{0i} + B_1^* \sum_{i=1}^n X_{0i} X_{1i} = \sum_{i=1}^n X_{0i} Y_i; \quad (4.40)$$

$$B_0^* \sum_{i=1}^n X_{1i} X_{0i} + B_1^* \sum_{i=1}^n X_{1i} X_{1i} = \sum_{i=1}^n X_{1i} Y_i.$$

Систему уравнений (4.40) можно представить в матрицах вида (4.31), из которых можно увидеть, что в соответствии с правилами (4.26) и (4.27) коэффициенты при неизвестных являются результатом перемножения двухстрочной транспонированной матрицы  $X^T$  на ее исходное значение  $X$ , учитывая, что исходное значение матрицы  $X$  целесообразно записать в виде двух столбцов (см. табл. 3.4), в которых подразумевается очевидная запись  $X_{0i} = 1$  для первого столбца. Также матрица свободных членов будет являться результатом перемножения транспонированной матрицы  $X^T$  (из двух столбцов) на однострочную матрицу опытных значений  $Y_{\text{опытн}}$ , как это также можно видеть из табл. 3.4. Таким образом, получим следующую матричную форму записи для системы (4.40):

$$(X^T X) B^* = X^T Y. \quad (4.41)$$

Применяя метод математической индукции, можно показать, что данное матричное уравнение будет справедливым и для системы из трех, четырех и большего числа нормальных уравнений, а также для линейного и нелинейного многофакторного уравнения регрессии вида (4.25). Отсюда для получения значений коэффициентов регрессии необходимо обе части матричного уравнения (4.41) умножить на обратную матрицу  $(X^T X)^{-1}$ . В результате окончательно получим [2]:

$$B^* = (X^T X)^{-1} X^T Y. \quad (4.42)$$

Модуль Multiple Regression (множественная регрессия) программы STATISTICA производит автоматическое вычисление коэффициентов линейной многофакторной модели по матричному уравнению (4.42) при условии ввода исходных данных по  $X_{ik}$  и  $Y_i$  в виде столбца. В то же время данный модуль можно использовать и для получения коэффициентов нелинейных многофакторных уравнений регрессии вида (4.25) при условии предварительной «ручной» подготовки исходных данных по  $X_{ik}^n$  и комбинации  $X$  первой степени, что не представляет особой сложности, хотя и требует внимания. Разработанная ранее в МАДИ (ГТУ) DOS-я программа выполняет эти расчеты автоматически, что значительно сокращает процедуру поиска рационального решения.

#### 4.4. Оценка точности и адекватности многофакторной регрессионной модели линейного и нелинейного (степенного) видов

Для многофакторной модели поиск рационального решения представляет определенные методические сложности, поскольку одновременно необходимо решать несколько задач и для начала исключить из рассмотрения независимые переменные (факторы влияния), между которыми существует сильная линейная зависимость, называемая мультиколлинеарностью (считается в случае, если коэффициент парной корреляции  $r_{X_i X_j}^* > 0,8$  [2]).

Это свидетельствует о дублировании параметров, которое должно быть исключено, и усложняет процедуру расчета стандартных ошибок, о которой будет сказано далее. Данные ошибки становятся очень большими («взрываются»), поскольку определитель (в силу его свойств) при вычислении обратной матрицы приближается к нулю и уравнение регрессии становится непригодным для практического использования.

Обнаружить мультиколлинеарность можно на основе просмотра матрицы парных корреляций, автоматически вычисляемой программой STATISTICA. Рекомендуется исключить тот параметр, для которого наблюдается большее снижение коэффициента множественной детерминации (последний, как отмечалось ранее, автоматически

вычисляется по выражению (4.22) для регрессионных зависимостей любых видов), однако такой формализованный подход лучше продублировать интуицией исследователя [2], оставив за ним окончательное решение.

Более сложными являются установление вида зависимостей и оценка точности модели в отношении ее прогнозных свойств. Как уже отмечалось ранее, линейная модель не всегда может быть использована, поскольку часто искажает физический смысл исследуемого явления. Например, при обработке данных табл. 3.4 для старых предприятий по линейной зависимости получается уравнение вида

$$Y = -2,71 + 1,12X,$$

из которого следует, что производительность труда при отсутствии дополнительной фондовооруженности может стать отрицательной, что абсурдно всегда (например, ТО и ремонт автомобилей в основном осуществляется на основе комплекта водительского инструмента).

В то же время нелинейная модель вида

$$Y = 13,6 + 0,00191X^2$$

дает более приемлемую зависимость, удобную для прогнозных расчетов. Поэтому для многофакторной модели предварительно рекомендуется сначала отработать вид однофакторной зависимости по каждому из факторов (это удобно делать на основе программы STATISTICA), а затем переходить к полной зависимости с использованием полученных моделей.

Для оценки точности модели в соответствии с общими требованиями необходимо знать ошибки оценки ее коэффициентов, определяющих линию условных средних  $\hat{Y}$ . Большие значения этих ошибок могут привести к тому, что вычисленные по матричному уравнению (4.42) коэффициенты могут оказаться статистически незначимыми и не пригодными для прогнозных расчетов. Оценку их значимости принято осуществлять, как уже отмечалось для однофакторной линейной модели, на основе  $t$ -критерия (Стьюдента) [2], однако процедура оценки значительно усложняется. В основу процедуры положен общий принцип определения доверительного интервала для статистической оценки центра распределения по выражению (2.43), однако в данном случае он модернизирован, аналогично операции о проверке существенности меры расхождения между средними двух выборок (3.3), непосредственным вычислением отношения величины коэффициента  $B_K^*$  к его ошибке  $S_{B_K}$ :

$$t_{q, n-m} = \frac{B_K^*}{S_{B_K}}, \quad (4.43)$$

где  $n$  — число опытных данных;  $m$  — количество определяемых коэффициентов модели (с учетом свободного члена).

В специальной литературе [2] доказывается, что дисперсия ошибок для всех определяемых коэффициентов будет представлять собой диагональные элементы ковариационной матрицы  $K_{B^*}$ , которая является произведением обратной матрицы  $(X^T X)^{-1}$  на величину остаточной дисперсии  $S_e^2$ , выступающей в данном случае как масштабный множитель (наглядное доказательство соответствия данного вывода общему принципу оценки стандартной ошибки по выражению (2.38) будет дано в подразд. 4.5, посвященному активному многофакторному эксперименту):

$$K_{B^*} = (X^T X)^{-1} S_e^2. \quad (4.44)$$

При этом оценка остаточной дисперсии будет равна [2]:

$$S_e^2 = \frac{Q_{\text{ост}}}{n-m}, \quad (4.45)$$

где  $Q_{\text{ост}}$  представляет собой сумму квадратов отклонений фактических данных  $Y_i$  от их сглаженных (предсказанных) значений  $\hat{Y}_i$  и программно вычисляется автоматически в процессе обработки данных.

Обратную матрицу  $(X^T X)^{-1}$  в силу зависимости (4.44) принято называть дисперсионной матрицей, или матрицей ошибок, а исходную матрицу  $(X^T X)$  — информационной матрицей, содержащей все сведения о значениях факторов  $X$ . Соответственно правую часть матричного уравнения (4.41) называют вектором моментов, поскольку произведение  $Y_i$  (сила) на  $X_i$  (плечо) эквивалентно моментам различного порядка [2].

Если уровень значимости  $q$  при оценке  $t$ -критерия по выражению (4.43) будет достаточно весомым (не более 5%, хотя для целей технической эксплуатации следует считать приемлемым и до 20%), то коэффициент считается значимым и его оставляют в уравнении регрессии; в противном случае он принимается равным нулю и соответствующий член уравнения исключается из рассмотрения. Однако определение значимых коэффициентов считается недостаточным для оценки качества многофакторной модели, поскольку не позволяет в полной мере выявлять важность отдельных факторов по их вкладу в раскрытие вариации показателя  $Y$ . Это связано с тремя обстоятельствами [2]:

- 1) величина  $B_K^*$  зависит от масштаба представлений фактора  $X_K$  в таблице данных;
- 2) все коэффициенты  $B_K^*$  имеют различные размерности;
- 3) при рассмотрении коэффициента  $B_K^*$  не учитывается диапазон, в пределах которого изменяется  $X_K$  в таблице исходных данных.

Получение более полной информации о мере влияния отдельных факторов становится возможным после стандартизации переменных, основой которой является нахождение так называемых

бетта-коэффициентов при стандартизированных переменных по формуле

$$\beta_K = B_K^* \frac{S_{X_K}}{S_Y}, \quad (4.46)$$

где  $S_{X_K}$  — стандартное (среднеквадратическое) отклонение для  $K$ -го фактора влияния;  $S_Y$  — стандартное отклонение для параметра отклика  $Y$ .

Бетта-коэффициенты являются безразмерными, поэтому сразу «снимаются» причины 1 и 2. Физически они показывают, на какую долю своей единицы изменения  $S_Y$  изменится показатель  $Y$ , если фактор  $X_K$  увеличится на свою единицу изменения  $S_{X_K}$  (при условии, что все остальные факторы останутся неизменными). При этом  $S_{X_K}$  и  $S_Y$  означают просто меру вариации и никак не связаны с обязательностью нормального распределения системы  $\{Y, X\}$ . Таким образом, бетта-коэффициенты учитывают влияние факторов в сопоставимых единицах. Стандартизированное уравнение с этими коэффициентами не имеет свободного члена.

Например, для двухфакторной линейной зависимости было получено следующее уравнение со значащими коэффициентами:

$$\hat{Y} = 0,365 + 0,0383X_1 + 0,0925X_2, \quad (4.47)$$

из которого можно сделать предварительное заключение о более сильном влиянии параметра  $X_2$  на показатель  $Y$ . Однако после перехода к бетта-коэффициентам

$$\hat{Y}_{\text{станд}} = 0,843X_1 + 0,277X_2 \quad (4.48)$$

видно, что влияние фактора  $X_1$  втрое выше по сравнению с фактором  $X_2$ .

Последнее заключение о качестве полученной многофакторной нелинейной модели делается, как уже отмечалось ранее, на основе ее проверки на адекватность по критерию Фишера и величинам коэффициентов множественной детерминации и корреляции (последнее осуществляется для линейных моделей при условии нормальных законов распределения для функции отклика  $Y$  и независимых параметров  $X$ ). Критерий Фишера в любом случае подсчитывается как отношение объясненной дисперсии параметра  $Y$  к ее остаточной части по выражению (4.23), а в случае линейных зависимостей через множественный коэффициент детерминации  $R_{yX}^2$  — по формуле

$$F_{q, m-1, n-m} = \frac{R_{yX}^2(n-m)}{(m-1)(1-R_{yX}^2)}. \quad (4.49)$$

В первом приближении соотношение (4.49), как и связанные с ним расчеты, могут быть использованы и для нелинейных моделей любых видов. Для линейных моделей значение  $R_{yX}^2$ , как отмечалось ранее,

может непосредственно рассчитываться по выражению (4.22) как отношение суммы квадратов объясненных отклонений к общим для параметра отклика  $Y$ . Рекомендуется использовать для прогнозных целей уравнения регрессии при значении  $R_{yX}^2 > 0,75$  [2], однако такие показатели на практике будут встречаться редко и по интуиции исследователя можно допускать более низкие величины, ориентируясь в основном на достаточность оценки адекватности модели по критерию Фишера и возможность получения приемлемых прогнозных показателей, о чем будет сказано далее.

Для линейных моделей имеет место следующая связь коэффициента множественной детерминации  $R_{yX}^2$  с коэффициентами парной корреляции  $r_{YX_K}^*$  [2]:

$$R_{yX}^2 = \beta_1 r_{YX_1}^* + \beta_2 r_{YX_2}^* + \dots + \beta_K r_{YX_K}^*, \quad (4.50)$$

которую можно использовать для проверки правильности подсчета  $R_{yX}^2$ , в том числе и для нелинейных моделей.

Программа STATISTICA обеспечивает получение всех отмеченных ранее показателей из условия расчета линейных моделей, хотя при этом могут заданы и нелинейные по выражению (4.25). При этом, в первую очередь, на экран выводятся: значения коэффициента множественной корреляции  $R_{yX}$ ; значения множественной детерминации  $R_{yX}^2$ ; его скорректированное значение, поскольку считается, что при небольшом числе данных для последнего получается завышенная оценка [2]; значение критерия Фишера с оценкой уровня его значимости; величины бетта-коэффициентов, которые выделяются красным цветом при уровне значимости ниже 5%.

При необходимости дополнительно можно получить таблицы со значениями всех коэффициентов модели  $B_K$  и оценками их значимости по критерию Стьюдента, а также корреляционные и ковариационные матрицы, по которым можно осуществить «доводку» модели по принципам, изложенным ранее.

#### 4.5. Осуществление прогноза по разработанным регрессионным моделям и выявление аномальных исходных данных

Однако программа STATISTICA (как и другие) имеет ограниченные возможности по вычислению прогнозных значений, которые позволяют оценить поведение исследуемого объекта в некоторых ситуациях, отличных от тех, для которых разрабатывалась модель, т.е. в точках пространства факторов  $X_L$ , выходящих за пределы поля задания  $X$  таблицей исходных данных. Фактически для линейной однофакторной зависимости программа STATISTICA обеспечивает только графическое изображение 95%-ного доверительного растроба

для показателя  $Y$  в пределах диапазона изменения независимой переменной  $X$ , как это схематично показано на рис. 4.4.

Для сравнения отметим, что DOS-я программа МАДИ (ГТУ) позволяет осуществлять расчет доверительного интервала за пределами диапазона  $X$ , хотя и без построения графика, для более сложной однофакторной модели, выражаемой полиномом 3-й степени. Однако важность указанной задачи может обуславливать необходимость применения и «ручных» методов расчета для многофакторных моделей, которые возможны на основе правильного понимания теоретических и методических предпосылок. Согласно им [2] точку в пространстве  $X_L$ , в которой предполагается использовать модель, в матричной форме можно обозначить следующим образом:  $f_L^T = \|1, X_{L1}, \dots, X_{Lm}\|$ , где  $f_L^T$  повторяет по форме уже подобранное линейное или нелинейное уравнение многофакторной регрессии (без коэффициентов). Исследователю в точке  $X_L$  надо знать:

- 1) значение показателя  $\hat{Y}_L$ ;
- 2) точность модели, оцениваемой дисперсией  $S_{\hat{Y}_L}^2$ ;
- 3) интервал (диапазон изменения) возможных значений показателя  $Y_L$ .

Первое требование выполняется при помощи очевидного матричного произведения:

$$\hat{Y}_L = f_L^T B^* \quad (4.51)$$

Согласно [2] дисперсия ошибки модели  $S_{\hat{Y}_L}^2$ , которая является суммарным отражением ошибок всех коэффициентов модели, будет определяться через произведение матриц  $f_L^T$ ,  $(X^T X)^{-1}$  и  $f_L$  с общим сомножителем  $S_e^2$  по выражению

$$S_{\hat{Y}_L}^2 = S_e^2 f_L^T (X^T X)^{-1} f_L \quad (4.52)$$

Выражение (4.52) несложно рассчитывать «вручную», используя готовые программы перемножения матриц (например, имеющиеся в MATLAB; при этом при вводе данных не надо забывать о том, что клетки первого столбца матрицы  $\|X\|$  для  $X_{0i}$  должны быть заполнены единицами), а также программы нахождения обратных матриц (последнюю операцию можно выполнять «вручную» по алгоритму (4.36)).

Из выражения (4.52) можно видеть, что ошибка модели зависит от координаты той точки, куда экспериментатор переносит модель, и растет пропорционально квадрату роста относительного значения факторов. Это значит, что за пределами диапазона задания исходных данных таблицы (матрицы)  $X$  дисперсия ошибки модели растет очень значительно и может привести к неудовлетворительному прогнозу, скорректировав тем самым последующие практические решения.

Именно рост ошибки по мере удаления от центра двумерного распределения для линейной однофакторной регрессии обуславливает графический раструб, показанный на рис. 4.4, хотя считается [2], что для линейных моделей и при предположении о нормальности закона распределения случайных величин  $\{Y, X\}$  точность модели не зависит от  $f_L$  и постоянна на всем возможном наборе  $X_L$  (для однофакторной зависимости она будет определяться в соответствии с выражением (4.16)). Таким образом, предположение о нормальности распределения дает большие выгоды в отношении получения благоприятного прогноза и оно должно быть убедительно доказано в тех случаях, когда по физическим соображениям возможно использование линейных моделей. Однако более надежным следует считать получение прогноза с использованием зависимости (4.52).

Оценка интервала  $Y_{L\min} \dots Y_{L\max}$ , в пределах которого с заданным уровнем доверительной вероятности (обычно 95 %) будет находиться опытное значение показателя  $Y_L$ , производится на основе следующих соображений. В пределах табличных данных оценки ошибок  $\hat{Y}_L$  имеют дисперсию  $S_e^2$ . По аналогии можно полагать, что ошибки измерения показателя  $Y$  будут иметь такие же свойства и за пределами измерения таблицы данных.

Следовательно, если модель верна, то фактическое значение  $Y_L$  может отличаться от модельного значения на ошибку, определяемую дисперсией  $S_e^2$ . Кроме того, само модельное значение имеет ошибку, характеризующуюся дисперсией  $S_{\hat{Y}_L}^2$ . Поскольку ошибки модели и измерения  $Y_L$  независимы, общая ошибка прогноза  $S_{\text{прогн}}^2$  будет равна их сумме. Отсюда с учетом нормального закона распределения могут быть определены верхняя  $Y_{L\max}$  и нижняя  $Y_{L\min}$  границы доверительного интервала для  $\hat{Y}_L$  по выражениям:

$$S_{\text{прогн}}^2 = S_e^2 + S_{\hat{Y}_L}^2 \quad (4.53)$$

$$Y_{L\min} = \hat{Y}_L - t_{\beta} \cdot S_{\text{прогн}}; \quad Y_{L\max} = \hat{Y}_L + t_{\beta} \cdot S_{\text{прогн}} \quad (4.54)$$

где  $t_{\beta}$  — значение параметра нормированной функции Лапласа (см. Приложение 1), соответствующего двухсторонней доверительной вероятности  $\beta^*$  (для  $\beta^* = 0,95$   $t = 1,96$ ).

Повышению точности прогноза помимо изложенной ранее методики будет способствовать исключение из статистической обработки (при определении коэффициентов модели) аномальных исходных данных, которые также приводят к дополнительным ошибкам. Если для однофакторных моделей эти данные могут быть установлены другими путями, например на основе подбора закона распределения при ограничении диапазона статистических данных (см. рис. 3.1), позволившего выявить несовместимость данных по предприятию № 17 для группы старых предприятий (см. табл. 3.4), то для многофакторных моделей подобный подход может даже способствовать

увеличению ошибки, особенно при ограниченном числе данных. (В общем случае аномальные данные необходимо выявлять и исключать всеми возможными способами.)

Поэтому интерес представляет способ непосредственного решения данной задачи на основе использования так называемой проективной матрицы, не реализованной в программе STATISTICA, хотя и широко применяемый в специализированных программах (например, разработанных в МАДИ (ГТУ) [2]). Однако данный подход несложно реализовать и на основе «ручного» расчета, как это рекомендовалось делать при использовании выражения (4.52) для оценки ошибки модели.

В основе методических предпосылок использования проективной матрицы лежит очевидная зависимость, показывающая, что вектор модельных (сглаженных) значений показателя  $Y$ , отраженный матрицей  $\|\hat{Y}\|$ , выражается через произведение матриц значений параметров  $X$  и коэффициентов модели  $B^*$  [2]:

$$\hat{Y} = XB^* \quad (4.55)$$

Подставив в формулу (4.55) значение  $B^*$  из (4.42), получим:

$$\hat{Y} = X(X^T X)^{-1} X^T Y = HY, \quad (4.56)$$

где

$$H = X(X^T X)^{-1} X^T \quad (4.57)$$

есть проективная матрица, проектирующая множество  $Y$  в множество сглаженных значений  $\hat{Y}$  (ее несложно вычислить «вручную»).

Проективная матрица формируется на точно известных табличных значениях факторов и не содержит (по определению) ошибок измерений. Она является симметрической и имеет размерность  $(n \cdot n)$ , где  $n$  — число независимых наблюдений. Для всех ее элементов справедливо условие:

$$|h_{ij}| < 1; \quad 0 < |h_{ij}| < 1.$$

Согласно [2] элементы проективной матрицы являются весовыми коэффициентами при проектировании  $Y$  в  $\hat{Y}$ , поскольку в соответствии с формулой (4.55) любое сглаженное значение  $\hat{Y}_i$  есть взвешенная сумма всех опытных значений  $Y_j$  (здесь индекс  $j$  взамен смыслового  $i$  для строк однострочковой матрицы  $\|Y\|$  введен для отражения процедуры перемножения матриц в соответствии с выражениями (4.26) и (4.27));

$$\hat{Y}_i = \sum_{j=1}^n h_{ij} Y_j; \quad (4.58)$$

при этом само значение  $Y_i$  входит в свое сглаженное значение с весом  $h_{ii}$ , являющимся диагональным элементом матрицы  $H$ . Суммирование в выражении (4.58) всех  $Y_j$  весами  $0 < |h_{ij}| < 1$  приводит к тому, что ошибки отдельных  $Y_j$  частично компенсируются и взвешенное значение  $\hat{Y}_i$  имеет в среднем меньшую ошибку. В этом и заключается сущность сглаживания.

Сумма всех диагональных элементов матрицы  $H$  (след матрицы  $tr$ ) равна числу оцениваемых коэффициентов модели  $m$ , включая свободный член, и не зависит от количества опытных наблюдений (размера квадратной матрицы)  $n$ :

$$trH = m.$$

Проективная матрица обладает рядом свойств, учитываемых при анализе исходных данных.

1. Проективная матрица идемпотентна, что выражается условием

$$H \times H = H, \quad (4.59)$$

которое легко доказывается (перемножение приводит к образованию единичной матрицы  $E$ ). Из этого свойства следует, что нельзя повторно сгладить уже однажды сглаженную последовательность.

2. Диагональный элемент  $h_{ii}$  равен сумме квадратов всех элементов  $i$ -й строки (или  $j$ -го столбца, поскольку матрица  $H$  симметрическая), включая самого себя:

$$h_{ii} = \sum_{j=1}^n h_{ij}^2. \quad (4.60)$$

Отсюда следует, что дисперсия любого  $i$ -го сглаженного значения показателя пропорциональна величине  $i$ -го диагонального элемента проективной матрицы:

$$S_{\hat{Y}_i}^2 = h_{ii} S_e^2. \quad (4.61)$$

Следовательно, наименее сглаженными оказываются те опытные данные  $Y_i$ , для которых  $h_{ii}$  близко к максимуму, равному единице. Таким образом, большое значение  $h_{ii}$ , в том числе и резко выделяющееся по сравнению с основной массой диагональных элементов, может являться признаком аномальности набора факторов  $\{X_{i1}, X_{i2}, \dots, X_{iA}\}$ . В целом с позиций сглаживания оптимальным будет такое условие, когда все диагональные элементы равны между собой (что имеет место при ортогональном планировании активного многофакторного эксперимента, см. гл. 5).

3. Пусть  $Y_i$  содержит аномальное отклонение (ошибку) и пусть для этого опыта  $h_{ii}$  велико ( $h_{ii} \approx 1$ ).

Рассмотрим ошибку сглаживания  $\varepsilon_{\hat{Y}_i}$  для этой  $i$ -й строки данных. При этом будем иметь:  $\varepsilon_{\hat{Y}_i} = Y_i - \hat{Y}_i$ .

Раскроем  $\hat{Y}_i$  и выделим из него член с  $h_{ii} \cdot Y_i$ . Тогда получим:

$$\varepsilon_{\hat{Y}_i} = (1 - h_{ii})Y_i - \sum_{j=1, (6 \neq j=i)}^n h_{ij}Y_j. \quad (4.62)$$

Из выражения (4.62) видно, что если  $h_{ii} \approx 1$ , то в составе ошибки сглаживания  $\varepsilon_{\hat{Y}_i}$  само аномальное значение  $Y_i$  практически не будет представлено в связи с тем, что  $(h_{ii} - 1) \approx 0$ . Отсюда следует, что в общем случае, даже у простейших однофакторных моделей, аномалия  $Y_i$  может и не проявиться в аномалии  $\varepsilon_{\hat{Y}_i}$ , получаемой после разработки модели, а скажется на других сглаженных значениях. Таким образом, анализ таблицы исходных данных на предмет выявления аномальных результатов наблюдений должен заключаться в параллельном анализе как вектора остатков  $\varepsilon_{\hat{Y}_i}$ , что предусмотрено программой STATISTICA в разнообразных вариантах, так и (главным образом) вектора диагональных элементов проективной матрицы  $H$  (вектор  $n$ -мерного пространства есть упорядоченная совокупность  $n$  действительных чисел  $(x_1, x_2, \dots, x_n)$ ).

Если выделяющиеся значения будут обнаружены, хотя их может быть немного из-за ограниченного размера следа матрицы  $H$ , то надо их или исключить, или сформировать из них отдельную группу, построив для нее свою модель. Так, анализ диагональных элементов матрицы  $H$  регрессионной модели для старых предприятий ( $Y = 13,6 + 0,00191X^2$ , см. табл. 3.4), составивших значения  $0,34 - 0,14 - 0,20 - 0,72 - 0,22 - 0,23 - 0,14$ , подтвердил ранее сделанное заключение об аномальности предприятия № 17 для этой группы (значение  $0,72$  более чем втрое выше основной массы показателей), и после его исключения скорректированные параметры модели составили немного лучшие значения ( $Y = 14,2 + 0,00183X^2$ ) при диагональных элементах:  $0,47 - 0,23 - 0,23 - 0,58 - 0,24 - 0,26$ .

Дальнейшее исключение двух предприятий уже дает практически одинаковые диагональные элементы и возвращение к параметрам предыдущей модели, однако при этом резко ухудшаются оценки по Стьюденту и Фишеру, т.е. прогнозные свойства модели. Таким образом, первую корректировку следует считать существенной. С другой стороны, «включение» предприятия № 17 в группу модернизированных, целесообразное с позиций распределения по параметру  $Y$ , показало еще большую его аномальность по диагональным элементам матрицы  $H$  (в данном случае из-за значения параметра  $X$ ), а полученное для модернизированных предприятий уравнение регрессии  $Y = 27,2 + 0,007X^2$ , хотя и имело «на пределе» оценку по Стьюденту, однако оказалось неадекватным по Фишеру.

Таким образом, можно сделать очевидное заключение, что хотя модернизация и повысила средний уровень показателя производительности труда, однако уже не дает отдачи средств, вложенных в механизацию, поэтому более целесообразным (при сопоставимых эконо-

Таблица 4.2

Исходные данные о работе предприятий

	Y	X <sub>1</sub>	X <sub>2</sub>	i	Y	X <sub>1</sub>	X <sub>2</sub>	i	Y	X <sub>1</sub>	X <sub>2</sub>
1	230	135	107	10	166	116	86	19	120	104	82
2	81	93	67	11	146	125	78	20	198	116	76
3	100	108	81	12	196	114	89	21	189	112	80
4	212	138	93	13	202	128	84	22	184	109	85
5	216	123	81	14	203	129	80	23	174	113	75
6	156	116	86	15	201	125	99	24	168	113	87
7	201	119	86	16	195	120	86	25	143	104	69
8	194	112	96	17	180	126	92	26	131	103	65
9	164	128	80	18	174	136	95	27	130	125	84

номических затратах) для старых предприятий следует считать повышение уровня их механизации.

В заключение приведем пример построения модели многофакторной регрессии [2], способствующий освоению работы с модулем Multiple Regression (множественная регрессия) программы STATISTICA. Исходные данные, отражающие некоторые экономические показатели работы 27 предприятий, приведены в табл. 4.2.

Здесь при обработке в первом приближении была получена полиномиальная линейная регрессионная модель

$$Y = -104 + 1,50X_1 + 1,18X_2 \quad (4.63)$$

при  $R_{yx}^2 = 0,476$ ,  $r_{yx}^* = 0,69$ ,  $t_1 = 2,34$ ,  $t_2 = 1,59$  и  $F_{2,24} = 10,9$ , которые в целом можно признать удовлетворительными. Коэффициент корреляции между независимыми параметрами  $X_1$  и  $X_2$  оказался равен  $r_{X_1X_2}^* = 0,64$ , что указывает на отсутствие мультиколлинеарности. Анализ диагональных элементов проективной матрицы также не выявил аномальные исходные данные.

Однако данная модель не соответствует физическому смыслу изучаемого явления, поскольку свободный член имеет отрицательное значение. Для ее уточнения проводилась подборка моделей для парных связей параметра отклика отдельно с факторами  $X_1$  и  $X_2$ , в ре-



зультате чего была принята новая (второй степени) полиномиальная регрессионная модель:

$$\hat{Y} = 35,1 + 0,0061X_1^2 + 0,0072X_2^2. \quad (4.64)$$

Для этой модели получены такие же статистические оценки ( $R_{yx}^2 = 0,46$ ,  $r_{yx}^* = 0,68$ ,  $t_1 = 2,24$ ,  $t_2 = 1,65$  и  $F_{2,24} = 10,0$ ), однако именно физическая достоверность обеспечивает возможность ее применения для целей прогноза и принятия практических решений.

### КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Дайте определение понятия сглаживания стохастических зависимостей при построении математической модели экспериментального исследования.
2. В чем состоит идея корреляционного анализа и какова область его применения?
3. В чем состоит идея регрессионного анализа и какова область его применения?
4. В чем состоит идея метода наименьших квадратов? Приведите формулы для определения коэффициентов модели.
5. Каким образом осуществляется проверка значимости коэффициентов регрессии?
6. Дайте определение понятия коэффициента детерминации. Как он используется для оценки точности и адекватности линейной модели?
7. В чем заключается преимущество матричных способов оценки коэффициентов полиномиальной модели  $n$ -й степени?
8. В чем сущность метода оценки точности и адекватности полиномиальных регрессионных моделей (линейных и степенных)?
9. Каким образом может быть осуществлено выявление аномальных данных для расчета коэффициентов однофакторной линейной модели?
10. В чем сущность методики прогноза по разработанным регрессионным моделям?

## ГЛАВА 5

### Применение активных многофакторных экспериментов при решении задач технической эксплуатации автомобилей

#### 5.1. Простейший случай статистического планирования активного однофакторного эксперимента

В гл. 1 в первом приближении рассмотрены широкие и зачастую единственные возможности получения представления об интересующих свойствах исследуемого объекта типа «черный ящик» на основе активного статистического эксперимента. Идея этого эксперимента (Experimental Design) принадлежит Рональду Фишеру (конец 20-х годов XX в.), который впервые показал целесообразность однофакторного варьирования всеми факторами в противовес широко распространенному (и в настоящее время) однофакторному эксперименту.

В процессе однофакторного эксперимента для случая многофакторной зависимости поочередно варьировался каждый фактор, при этом все остальные факторы стабилизировались на некотором постоянном уровне. Помимо необходимости реализации большого числа опытов (и, следовательно, возрастания трудоемкости исследований) такой эксперимент приводил к сложностям при обобщении полученных результатов; при этом их окончательная точность и правдоподобность обычно оказывались невысокими ввиду отсутствия методически проработанной формализации таких обобщений, а также из-за неизбежности ошибок учета фактического влияния на параметр отклика непредсказуемой совокупности изменений управляемых параметров, что характерно для практики.

Основные сложности такого планирования — стремление к минимизации общего числа опытов, которое возрастает в степенной зависимости по мере увеличения числа факторов, без потери качества конечных интерполяционных полиномиальных моделей.

Понадобилось еще несколько десятилетий, чтобы в начале 1930-х годов появилось новое направление в планировании, связанное с оптимизацией многофакторных процессов — планирование

да экспериментального планирования, поскольку название метода Уилсона (по имени разработавших его исследователей), остается актуальной, и он рекомендуется для широкого применения.

В ходе активного эксперимента могут решаться следующие типы задач:

- поиск модели, описывающей функционирование системы;
- выявление и отбор существенных факторов, влияющих на функцию отклика;
- оценка и уточнение констант теоретической модели;

- поиск оптимальных условий функционирования системы.

В определенном смысле активный эксперимент можно рассматривать как специфический случай многофакторного регрессионного анализа, в котором исходные данные не отбираются случайным образом из практики, а целенаправленно задаются исследователем по определенному алгоритму с использованием принципа кодирования (-1, +1) для наименьшего и наибольшего значений диапазона варьирования влияющих факторов, за счет которого достигается резкое упрощение как самих процедур планирования, так и обработки результатов.

Дальнейшая обработка полученных данных фактически производится рассмотренным ранее матричным методом по формуле (4.42), с сохранением необходимости использования критериев Стьюдента и Фишера для оценки точности и значимости модели, т.е. полностью используются возможности модуля Multiple Regression (множественная регрессия) программы STATISTICA при условии правильности ввода исходных данных. Это особенно важно, так как специфический модуль Experimental Design (планирование эксперимента) в обычной «рыночной» программе STATISTICA не присутствует, хотя его (в русифицированном варианте) в случае острой необходимости можно приобрести у специализированных фирм. Данный модуль обладает широкими возможностями по составлению планов многофакторного эксперимента (до 65 факторов) с дробными откликами, сокращающими число опытов; рандомизацией опытов; алгоритмам обработки различных вариантов отношения сигнал—шум и др.

В то же время особенности уже выполненных работ с использованием планирования эксперимента в области технической эксплуатации автомобилей (главным образом связанных с разработкой

ном числе факторов (до пяти—семи, для исследования которых может быть достаточно проведения восьми опытов) достаточно полно представлены в литературе [1, 2], а дополнительные необходимые дробные отклики и их рандомизация могут быть получены самим исследователем на основе сведений, приведенных далее.

Несложно учесть также ненужность бета-коэффициентов для полиномиальной модели регрессии, полученной в результате активного эксперимента, поскольку в силу особенностей планирования

рделении проективной матрицы, поскольку ее диагональные элементы при проведении активного эксперимента равнозначны. Самое главное, модуль Multiple Regression без всяких затруднений производит полную обработку (с критериями Стьюдента и Фишера и их непонятно как определяемыми степенями свободы и уровнями значимости) аномальных данных (с позиций классического активного планирования), когда, например, имеет место неодинаковое число статистических параллельных опытов для ортогонального линейного планирования.

Также значительное упрощение достигается и при работе со специальными планами второго порядка, не являющимися ортогональными и ротатабельными, для которых повторные опыты в основном могут отсутствовать (соответственно становится невозможным проводить расчет дисперсии воспроизводимости и оценку однородности дисперсий по критерию Кохрана) в сочетании с их многократным повторением (до четырех—пяти раз) для одной «точки» — центра плана.

И последнее обстоятельство, связанное со спецификой расчета критерия Фишера при проверке адекватности модели для активного эксперимента, в процессе которого по специфическим подходам, которые будут представлены далее, вычисляется отношение дисперсии адекватности [1] (по другой терминологии дисперсии неадекватности [2]) к дисперсии воспроизводимости. По мнению авторов, данный подход принципиально не меняет физической сущности оценки адекватности по сравнению с классическим методом, применяемым в модуле Multiple Regression, а в основном отражает специфику «ручного» расчета, используемого до появления вычислительной техники, который стал уже ненужным.





















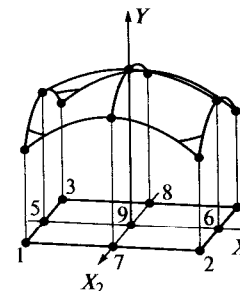
Меньшую критичность данной оценки можно видеть на примере однофакторной линейной модели вида, представленного на рис. 5.1. Для удобства рассмотрим условные результаты трех экспериментов, у которых на нижнем уровне получены одинаковые значения параметра отклика — 1—2—3, а на верхнем — 2—4—6, 3—5—7 и 6—8—10 соответственно.

Уравнения регрессии при этом имеют вид:  $Y_1 = 3 + X_1$ ;  $Y_2 = 3,5 + 1,5X_1$ ;  $Y_3 = 5 + 3X_1$ . Дисперсии воспроизводимости и неадекватности при расчетах по выражениям (5.17) и (5.21) для всех трех моделей одинаковы; соответственно будет одинаково и значение критерия (5.22) Фишера (для данных примеров равно нулю), в соответствии с которым все уравнения должны быть признаны адекватными в равной степени. Различие здесь проявляется только в статистических оценках для коэффициентов регрессии, которые составили:  $t_1 = 1,55$ , что соответствует 20%-ному процентному уровню значимости, для первой модели (как отмечалось ранее, для целей технической эксплуатации такой уровень может считаться допустимым);  $t_2 = 2,3$ , что соответствует 8%-ному уровню значимости, для второй модели;  $t_3 = 4,7$ , что соответствует 1%-ному уровню значимости, для третьей модели. Однако при использовании модуля Multiple Regression оценки критерия Фишера (при сохранении оценок для коэффициентов модели) оказались разными: 2,4, 5,4 и 21,4 при уровнях значимости, соответствующих полученным для критерия Стьюдента; при этом первая модель безусловно является неадекватной и существенно различие в оценке адекватности между второй и третьей моделями.

Указанное ранее различие в критичности двух оценок адекватности модели можно видеть только на простых задачах; для сложных многофакторных и особенно нелинейных моделей оно незаметно, поэтому в любых случаях по возможности необходимо пользоваться общим методом (4.23) оценки адекватности для регрессионных моделей, вплоть до выполнения «ручных» расчетов, хотя это значительно сложнее по сравнению со схемой, приведенной в табл. 5.8. Для рассматриваемого примера двухфакторной модели при расчетах данных табл. 5.9 на адекватность было получено значение  $F = 65$ , соответствующее уровню значимости 0,004 %, т. е. модель с высокой точностью может использоваться для целей прогноза.

Если двухфакторная линейная модель признается неадекватной, то необходимо переходить к нелинейной, обычно квадратичной, интерполяции (более высокие степени полинома в активном эксперименте не рекомендуются [1, 2]). Графическая интерпретация поверхности отклика для этой модели показана на рис. 5.4, на котором отражен идеальный случай, когда вершина «холма» совпадает с центром плана (обычно она находится в стороне, иногда совпадая с периферийными точками). Для полнофакторного планирования эксперимента под указанную зависимость обычно достраивают план первого порядка (см. табл. 5.8) путем добавления основных опытов

Рис. 5.4. Графическая интерпретация планирования активного эксперимента и поверхность отклика для двухфакторной квадратичной модели



«нулевыми» значениями для двух факторов планирования, соответствующих центру интервала планирования, с сохранением свойств ортогональности, симметричности и ротатабельности, как это показано на схеме табл. 5.10. Из данной таблицы, отражающей обязательные этапы «ручных» расчетов для проверки однородности дисперсий параллельных опытов, затем можно получить табл. 5.11 и ввода данных в модуль Multiple Regression для высокоточных расчетов.

В то же время можно видеть значительное усложнение экспериментального планирования по сравнению с линейной моделью (количество основных опытов увеличивается до девяти, а их общее число с учетом трех параллельных — до 27). Осуществить подобный эксперимент зачастую становится нереально, поэтому для этой цели рекомендуется использовать специальные менее трудоемкие планы, которые описаны далее.

Таблица 5.10

Матрица ортогонального планирования активного эксперимента для двухфакторной квадратичной модели с таблицами для расчета дисперсий основных опытов

Основные опыты $N$	Уровни факторов в кодированных значениях		Параллельные опыты ( $r = 3$ )			$\bar{Y}_j$	$S^2(Y_j)$
	$X_1$	$X_2$	$Y_{j1}$	$Y_{j2}$	$Y_{j3}$		
1	—	—	$Y_{11}$	$Y_{12}$	$Y_{13}$	$\bar{Y}_1$	$S^2(Y_1)$
2	+	—	$Y_{21}$	$Y_{22}$	$Y_{23}$	$\bar{Y}_2$	$S^2(Y_2)$
3	—	+	$Y_{31}$	$Y_{32}$	$Y_{33}$	$\bar{Y}_3$	$S^2(Y_3)$
4	+	+	$Y_{41}$	$Y_{42}$	$Y_{43}$	$\bar{Y}_4$	$S^2(Y_4)$
5	—	0	$Y_{51}$	$Y_{52}$	$Y_{53}$	$\bar{Y}_5$	$S^2(Y_5)$
6	+	0	$Y_{61}$	$Y_{62}$	$Y_{63}$	$\bar{Y}_6$	$S^2(Y_6)$
7	0	—	$Y_{71}$	$Y_{72}$	$Y_{73}$	$\bar{Y}_7$	$S^2(Y_7)$
8	0	+	$Y_{81}$	$Y_{82}$	$Y_{83}$	$\bar{Y}_8$	$S^2(Y_8)$
9	0	0	$Y_{91}$	$Y_{92}$	$Y_{93}$	$\bar{Y}_9$	$S^2(Y_9)$

Таблица 5.11

Схема ввода статистических данных полнофакторного активного эксперимента для двухфакторной квадратичной модели при использовании модуля Multiple Regression

Опыты	$X_{1i}$	$X_{2i}$	$X_{1i}^2$	$X_{2i}^2$	$Y_{ji}$	Опыты	$X_{1i}$	$X_{2i}$	$X_{1i}^2$	$X_{2i}^2$	$Y_{ji}$
1	-1	-1	1	1	$Y_{11}$	15	-1	0	1	0	$Y_{53}$
2	-1	-1	1	1	$Y_{12}$	16	1	0	1	0	$Y_{61}$
3	-1	-1	1	1	$Y_{13}$	17	1	0	1	0	$Y_{62}$
4	1	-1	1	1	$Y_{21}$	18	1	0	1	0	$Y_{63}$
5	1	-1	1	1	$Y_{22}$	19	0	-1	0	1	$Y_{71}$
6	1	-1	1	1	$Y_{23}$	20	0	-1	0	1	$Y_{72}$
7	-1	1	1	1	$Y_{31}$	21	0	-1	0	1	$Y_{73}$
8	-1	1	1	1	$Y_{32}$	22	0	1	0	1	$Y_{81}$
9	-1	1	1	1	$Y_{33}$	23	0	1	0	1	$Y_{82}$
10	1	1	1	1	$Y_{41}$	24	0	1	0	1	$Y_{83}$
11	1	1	1	1	$Y_{42}$	25	0	0	0	0	$Y_{91}$
12	1	1	1	1	$Y_{43}$	26	0	0	0	0	$Y_{92}$
13	-1	0	1	0	$Y_{51}$	27	0	0	0	0	$Y_{93}$
14	-1	0	1	0	$Y_{52}$						

**5.3. Ортогональное планирование активного эксперимента для линейной модели с количеством факторов больше двух и возможность сокращения числа основных опытов за счет использования реплик различной дробности**

Увеличение количества исследуемых факторов при ортогональном полнофакторном планировании активного эксперимента приводит к непропорциональному росту основных опытов, число которых в общем случае определяется зависимостью

$$N = a^k, \tag{5.24}$$

где  $a$  — количество задаваемых уровней для факторов;  $k$  — количество исследуемых факторов.

Отсюда для линейных молей при  $a = 2$  общее количество основных опытов для трех факторов будет равно восьми; для четырех факторов — 16. Для квадратичных моделей при  $a = 3$  уже получаем соответственно 27 и 81.

Основой оптимального планирования активного эксперимента (при количестве факторов больше двух) является сокращение числа опытов на основе использования эффекта незначимости коэффициентов неполной линейной модели для взаимодействий между двумя и более факторами. Так, для двухфакторной линейной модели, ортогональное планирование которой выполняется на основе доопределения к матрице табл. 5.7 столбца-вектора  $X_1X_2$  для учета эффекта взаимодействия, неполное квадратное уравнение будет иметь вид

$$\hat{Y} = B_0^* X_0 + B_1^* X_1 + B_2^* X_2 + B_{12}^* X_1 X_2, \tag{5.25}$$

где при линейном приближении  $B_{12}^* \rightarrow 0$  вектор-столбец  $X_1X_2$  можно использовать для нового фактора  $X_3$ . Отразим этот план матрицей табл. 5.12.

Здесь четыре опыта являются достаточными для определения четырех коэффициентов трехфакторной линейной модели вида

$$\hat{Y} = B_0^* X_0 + B_1^* X_1 + B_2^* X_2 + B_3^* X_3. \tag{5.26}$$

Если в дополнение к столбцам матрицы табл. 5.12 вычислить еще столбцы для произведений  $X_1X_3$  и  $X_2X_3$ , то можно увидеть, что элементы столбца  $X_1X_3$  совпадут с элементами столбца  $X_2$ , а элементы столбца  $X_2X_3$  — с элементами столбца  $X_1$ . Таким образом, найденные по матрице планирования табл. 5.12 коэффициенты линейной модели (5.26) будут оценками для совместных эффектов:

$$B_1^* \rightarrow \beta_1 + \beta_{23}; B_2^* \rightarrow \beta_2 + \beta_{13}; B_3^* \rightarrow \beta_3 + \beta_{12}. \tag{5.27}$$

Так как постулируется линейная модель, то предполагается, что эффекты взаимодействия близки к нулю или незначительны по сравнению с основными эффектами.

Таблица 5.12

Матрица ортогонального планирования полуреплики активного эксперимента для трехфакторной линейной модели

Основные опыты $N$	$X_{0i}$	$X_{1i}$	$X_{2i}$	$X_{3i} (X_1 X_2)$
1	+	-	-	+
2	+	+	-	-
3	+	-	+	-
4	+	+	+	+

нению с основными эффектами, поэтому соотношение (5.27) можно представить в следующем виде:

$$B_1^* \cong \beta_1; B_2^* \cong \beta_2; B_3^* \cong \beta_3. \quad (5.28)$$

На основе соотношений (5.27) и (5.28) количество основных опытов было уменьшено вдвое (с восьми до четырех), т. е. была использована только половина полного факторного эксперимента вида  $2^3$ , которая в теории [1, 2] получила название «полуреплики» и обозначение  $2^{3-1}$ . Другая полуреплика для трехфакторного эксперимента может быть получена, если  $X_3$  приравнять к  $(-X_1 \cdot X_2)$ ; в этом случае

$$B_1^* \rightarrow \beta_1 - \beta_{23}; B_2^* \rightarrow \beta_2 - \beta_{13}; B_3^* \rightarrow \beta_3 - \beta_{12}, \quad (5.29)$$

т. е. эффекты парных взаимодействий будут иметь противоположный знак и при усреднении результатов экспериментов по двум полурепликам могут быть получены отдельные оценки для основных эффектов, как в полном факторном эксперименте. Обе полуреплики в сочетании будут представлять два различных блока [1] полнофакторного активного эксперимента  $2^3$ , требующего восемь опытов. (В общем случае разбиение матрицы сложного полнофакторного эксперимента на блоки используется для получения плана, который может повысить точность оценок по факторам, подверженным сильному влиянию неуправляемых внешних воздействий [1]. Например, таким образом могут быть учтены ошибки, связанные с колебаниями давления, температуры и влажности атмосферного воздуха и качества топлива, являющимися существенными при оценках экономичности и экологической безопасности автомобильных двигателей.) В данном случае (для трехфакторной модели) обе полуреплики по отдельности будут отражать так называемый насыщенный эксперимент, при котором количество основных опытов равно числу коэффициентов модели.

Изложенное ранее позволяет добавить следующие удобные для последующего использования в более сложных случаях понятия. Так, при построении полуреплики  $2^{3-1}$  существует всего две возможности: приравнять  $X_3$  к  $X_1 \cdot X_2$  или к  $(-X_1 \cdot X_2)$ . Поэтому есть только две полуреплики  $2^{3-1}$ , первая из которых реализована матрицей планирования табл. 5.12. При произведении трех столбцов этой матрицы выполняется соотношение  $+1 = X_1 X_2 X_3$ ; для столбцов матрицы другой полуреплики это же соотношение (равное единице) будет иметь противоположный знак:  $-1 = X_1 X_2 X_3$ .

Символическое обозначение произведения столбцов, равного (+1) или (-1), называется *определяющим контрастом*. Контраст помогает производить оценку смешанных эффектов. Для этого нужно помножить обе части определяющего контраста на столбец, соответствующий данному эффекту. Так, для первого контраста для  $X_1$  получим

$$X_1 = X_1^2 X_2 X_3 = X_2 X_3, \quad (5.30)$$

поскольку  $X_i^2$  всегда равно единице. Соответственно для  $X_2$  и  $X_3$  имеем:

$$X_2 = X_1 X_2^2 X_3 = X_1 X_3; X_3 = X_1 X_2 X_3^2 = X_1 X_2. \quad (5.31)$$

Это значит, что коэффициенты линейного уравнения будут оценены, уже отраженными выражениями (5.27). Соотношения вида (5.30) и (5.31), показывающие, с каким из эффектов смешан основной эффект, называются *генерирующими соотношениями*.

Дальнейшее увеличение количества факторов приводит к непропорциональному росту числа оптимальных вариантов активного эксперимента (для четырех факторов их уже восемь вместо двух для трех) и требует от исследователя определенного искусства планирования в целом. При этом, в первую очередь, должно быть обращено внимание на правильное формирование целевой функции (обобщенного параметра отклика), являющейся неограниченной характеристикой цели исследования, и обоснованный выбор факторов влияния для этой целевой функции.

Целевые функции могут быть различными и отражать экономический, технологический, технико-экономический, эргономический, эстетический, организационный и другие аспекты исследуемой системы.

Целевую функцию выбирают с учетом ряда требований:

- она должна быть обобщенным результатом функционирования системы (в то же время ее нельзя делать слишком глобальной);
- она должна быть количественно измеряемой;
- она должна иметь определенный физический смысл;
- она должна реагировать на изменение факторов, т. е. обладать определенной чувствительностью;
- она должна быть статистически эффективной, т. е. определяться с достаточной точностью.

Факторы выбираются с учетом следующих требований, частично уже обоснованных ранее:

- они должны быть управляемыми, т. е. в ходе эксперимента значения факторов должны находиться на определенном уровне или изменяться по определенной программе;
- они должны быть линейно и функционально независимыми, т. е. должна проследиваться возможность задавать факторам любые значения независимо от уровней других факторов. Данное требование не могло быть выполнено в рассмотренном ранее примере двухфакторного эксперимента с исследованием компрессионных колец, однако при важности таких зависимостей необходимо искать иные пути задания приемлемых комбинаций уровней факторов, подобные отмеченному ранее методу поиска из статистической выборки;

- они должны быть совместимыми, т. е. должна прослеживаться возможность и безопасность реализации всех запланированных комбинаций уровней факторов;

- они должны быть измерены с достаточной точностью;

- в рассмотрении должны участвовать все существенные факторы;

- они должны быть количественно оценены (в рассмотрении могут участвовать как количественные, так и качественные факторы).

Что касается четырехфакторного эксперимента, то для его реализации необходимо, как минимум, восемь основных опытов (может быть использована только полуреплика  $2^{4-1}$ ), а основой минимизации будет являться матрица планирования и расчета коэффициентов модели полного факторного эксперимента  $2^3$ , представленная в табл. 5.13 (без очевидного столбца для  $X_{0i} = 1$ ).

В общем случае сочетание уровней факторов при их числе три и более может осуществляться с помощью следующих двух эквивалентных правил:

- при добавлении нового фактора каждая комбинация уровней исходного плана встречается дважды, в сочетании с нижним и верхним уровнями фактора; при этом частота смены уровней варьирования каждого последующего фактора вдвое меньше предыдущей (по данному наиболее простому правилу построены три первых столбца матрицы планирования табл. 5.13);

- уровень каждого последующего фактора определяется перемножением уровней предыдущих факторов, а затем исходный план повторяется с противоположными знаками (первая часть этого правила очевидно использовалась для построения с 4-го по 7-го включительно столбцов матрицы планирования табл. 5.13).

Таблица 5.13

**Матрица ортогонального планирования и расчета коэффициентов активного полного факторного эксперимента для трехфакторной неполной степенной модели**

Основные опыты $N$	$X_1$	$X_2$	$X_3$	$X_1X_2$	$X_1X_3$	$X_2X_3$	$X_1X_2X_3$
1	-	-	-	+	+	+	-
2	+	-	-	-	-	+	+
3	-	+	-	-	+	-	+
4	+	+	-	+	-	-	-
5	-	-	+	+	-	-	+
6	+	-	+	-	+	-	-
7	-	+	+	-	-	+	-
8	+	+	+	+	+	+	+

Важность данной матрицы заключается в том, что на основе ее восьми основных опытов могут быть построены оптимальные планы активных экспериментов с числом факторов от четырех до семи, достаточных для выполнения самых сложных исследований в области технической эксплуатации автомобилей (при семи факторах и восьми опытах будет иметь место насыщенный план). Так, четвертый фактор  $X_4$  здесь легко вводится взамен седьмого столбца, отражающего эффект тройного взаимодействия (или взаимодействия второго порядка [1]), который для линейных моделей еще ниже, чем мало значащие эффекты парных взаимодействий первого порядка, а столбцы с парными взаимодействиями, которые для линейной модели обычно исключаются, при необходимости дополнительно могут быть использованы для проверки значимости парных взаимодействий, если имеются какие-либо априорные сведения в этом отношении.

Таким образом, основным генерирующим соотношением здесь будет являться  $X_4 = X_1X_2X_3$  (или  $X_4 = -X_1X_2X_3$ , приводящее к другой матрице планирования, которую можно рассматривать как второй блок [1]) полнофакторного эксперимента  $2^4$ , требующего проведения 16 опытов), поскольку другими возможными соотношениями с эффектами парных взаимодействий, отраженных четвертым—шестым столбцами матрицы табл. 5.13, пользоваться в данном случае неразумно.

Определяющим контрастом для оптимального планирования при этом будет являться соотношение

$$I = X_1X_2X_3X_4, \quad (5.32)$$

из которого можно видеть, что такой тип смешивания даст возможность оценивать линейные эффекты совместно с эффектами второго порядка, а взаимодействия первого порядка — совместно друг с другом [1].

При пяти факторах в теории [1] рассматриваются уже два варианта сокращения количества опытов: 1) использование полуреплики  $2^{5-1}$ , требующей проведения 16 опытов; 2) использование 1/4 реплики (четверть-реплики)  $2^{5-2}$ , требующей проведения всего восьми опытов. Наиболее просто осуществить по изложенным ранее принципам планирование полуреплики на основе матрицы для полнофакторного эксперимента  $2^4$  с использованием генерирующего соотношения  $X_5 = X_1X_2X_3X_4$ , однако уже здесь будет наблюдаться большое количество избыточных опытов для линейной модели. Поэтому при количестве факторов от пяти до семи рекомендуется осуществлять планирование на основе 1/4, 1/8 и 1/16 реплик, приравнивая, как правило, последний фактор ( $X_7$ ,  $X_6$  или  $X_5$ ) тройному взаимодействию матрицы планирования табл. 5.13, а остальные — двойному (при пяти или шести факторах здесь очевидны определенные варианты, в том числе с приравниванием последнего фактора двойному взаимодействию в тех случаях, когда значимость тройного взаимодействия

может быть значительна и требовать экспериментальной проверки). В общем случае реплики, которые используются для сокращения опытов в  $2^m$  раз, где  $m = 1, 2, 3, \dots$ , называются регулярными и их популярность связана с сохранением простоты расчета коэффициентов уравнения, как и в случае полного факторного эксперимента.

Для пяти факторов возможно 12 вариантов планирования при обычном использовании четверть-реплики, из которых первый может быть задан следующими очевидными генерирующими соотношениями:  $X_4 = X_1X_2$  и  $X_5 = X_1X_2X_3$ . Тогда определяющими контрастами будут являться:  $1 = X_1X_2X_4$  и  $1 = X_1X_2X_3X_5$ . Если перемножить эти определяющие контрасты, то получится третье соотношение, задающее элементы столбца  $1 = X_3X_4X_5$ . Чтобы полностью охарактеризовать разрешающую способность реплики, необходимо записать так называемый обобщающий определяющий контраст:  $1 = X_1X_2X_4 = X_3X_4X_5 = X_1X_2X_3X_5$ . Система смешивания определяется умножением обобщающего определяющего контраста последовательно на  $X_1, X_2, X_3$  и т. д.:

$$\begin{aligned} X_1 &= X_2X_4 = X_1X_3X_4X_5 = X_2X_3X_5; \\ X_2 &= X_1X_4 = X_2X_3X_4X_5 = X_1X_3X_5; \\ X_3 &= X_1X_2X_3X_4 = X_4X_5 = X_1X_2X_5; \\ X_4 &= X_1X_2 = X_3X_5 = X_1X_2X_3X_4X_5; \\ X_5 &= X_1X_2X_4X_5 = X_3X_4 = X_1X_2X_3; \\ X_1X_3 &= X_2X_3X_4 = X_1X_4X_5 = X_2X_5; \\ X_1X_5 &= X_2X_4X_5 = X_1X_3X_4 = X_2X_3. \end{aligned} \quad (5.33)$$

Можно видеть, что в результате получается довольно сложная система смешивания линейных эффектов с эффектами взаимодействия первого, второго, третьего и даже четвертого порядков. Поэтому при проверочных расчетах если, например, коэффициенты  $B_{13}^* \rightarrow \beta_{13} + \beta_{234} + \beta_{145} + \beta_{25}$  и  $B_{15}^* \rightarrow \beta_{15} + \beta_{245} + \beta_{134} + \beta_{23}$  окажутся значимыми, то могут возникнуть сомнения, можно ли пренебрегать другими парными взаимодействиями, с которыми связаны линейные эффекты. Тогда рекомендуется [1] поставить вторую серию опытов, выбрав нужным образом другую четверть-реплику. При этом рекомендуется воспользоваться методом «перевала», который практически осуществляется путем изменения всех знаков первой матрицы на обратные. Тогда в обобщающем определяющем контрасте тройные произведения имеют знак, противоположный их знаку в первой четверть-реплике. Поскольку они задают парные взаимодействия в совместных оценках для линейных эффектов, при усреднении результатов обеих четверть-реплик можно получить линейные эффекты, не смешанные с парными взаимодействиями.

В этом легко убедиться в рассматриваемом нами примере планирования  $2^{5-2}$ , если выбрать вторую четверть-реплику по генериру-

ющим соотношениям  $X_4 = -X_1X_2$  и  $X_5 = X_1X_2X_3$ , соответствующим методу «перевала». При этом обобщающий контраст принимает вид  $1 = X_1X_2X_4 = -X_3X_4X_5 = X_1X_2X_3X_5$ , при котором все нечетные взаимодействия (первого и третьего порядков) в соотношениях (5.33) меняют свои знаки на противоположные и при сложении двух четверть-реплик происходит освобождение линейных эффектов от парных взаимодействий. Таким образом, описанная ранее процедура может служить основой построения плана с более высокой разрешающей способностью (если в этом имеется необходимость) путем объединения двух реплик одинаковой дробности и обладает гибкостью в отношении целенаправленного исследования определенных (двойных или тройных) взаимодействий.

При шести факторах при сохранении восьми опытов матрицы табл. 5.13 требуется  $1/8$  реплика  $2^{6-3}$ . Здесь возможно 24 варианта выбора реплики, однако их количество резко сокращается на основе следующих принципов: из рассмотрения исключаются варианты без использования тройного взаимодействия и варианты с тем двойным взаимодействием, по которому априорно имеются сведения о его предполагаемой значимости, а из факторов  $X_4, X_5$  и  $X_6$  к тройному взаимодействию приравнивается наиболее существенный. Соответственно необходимо предусматривать рациональный вариант до тройки плана до четверть-реплики в целях освобождения линейных эффектов от эффектов взаимодействия первого порядка.

Что касается  $1/16$  реплики плана  $2^{7-4}$  для семи факторов, то здесь генерирующие соотношения непосредственно следуют из матрицы табл. 5.13 и имеют малую разрешающую способность [1], поскольку каждый линейный эффект определяется совместно с тремя парными взаимодействиями. Такой репликой можно пользоваться только в том случае, если все парные взаимодействия равны нулю. Поскольку до начала исследования это трудно предсказать (при решении задач в области технической эксплуатации некоторые парные эффекты оказывались значимыми), в случае сомнения необходимо предусмотреть последующее расширение плана за счет специальным образом выбранной второй реплики  $2^{7-4}$  на основе метода «перевала». Здесь должны быть использованы генерирующие соотношения:

$$X_4 = -X_1X_2; X_5 = -X_1X_3; X_6 = -X_2X_3; X_7 = X_1X_2X_3. \quad (5.34)$$

В обобщающем определяющем контрасте все тройные произведения окажутся со знаком «минус», и в совместных оценках для линейных эффектов (по принципам схемы (5.33)) не будет парных взаимодействий со знаком «плюс». Усредняя результаты вычислений для двух таких реплик, можно получить отдельные оценки для всех линейных эффектов. При отсутствии чрезмерных сложностей в осуществлении такого расширенного плана его целесообразно применять непосредственно на первом этапе.



На последнем примере можно видеть целесообразность избежания насыщенных планов за счет обоснованного снижения количества факторов, чтобы обеспечить возможность исследования парных взаимосвязей без удвоения числа опытов. Для этого, как отмечалось ранее, необходимо тщательно проанализировать всю имеющуюся информацию об исследуемых факторах или даже провести некоторые предварительные более простые исследования на двух или трех факторах, используя их результаты для получения менее насыщенного плана с восемью основными опытами. Шестнадцать опытов следует использовать только при числе факторов от восьми и более, если не удается избежать таких сложных планов.

Усложнение плана также может происходить в результате необходимости интерполяции одного, иногда двух, факторов квадратичными зависимостями, когда для остальных достаточны линейные. Такие модели имеют место в практике технической эксплуатации (например, при исследовании топливной экономичности и экологической безопасности), поскольку априорно известно о влиянии на них степени отклонения как в одну, так и в другую сторону углов опережения зажигания или впрыскивания дизельного топлива от их оптимальных значений (для последних моделей бензиновых двигателей с инжекторами эти факторы действуют уже одновременно). Авторам представляется достаточным получение подобных несложных планов для одного фактора на основе описанных ранее реплик для четырех и восьми опытов. Так, для трех факторов матрица планирования, представляющая собой расширение полуреплики табл. 5.12, без нарушения принципов ортогональности (5.11) и симметричности (5.12) может быть дополнена двумя опытами с «нулевыми» точками для фактора  $X_3$ , задаваемого на основе эффекта смешивания, т. е. применено так называемое композиционное планирование (табл. 5.14).

Таблица 5.14

Композиционная матрица ортогонального планирования и расчета коэффициентов модели на основе полуреплики активного эксперимента для трехфакторной зависимости с одним квадратичным членом

Основные опыты $N$	$X_{1i}$	$X_{2i}$	$X_{3i} (X_{1i} \cdot X_{2i})$	$X_{3i}^2$
1	-	-	+	+
2	+	-	-	+
3	-	+	-	+
4	+	+	+	+
5	+	-	0	0
6	-	+	0	0

Таблица 5.15

Композиционная матрица ортогонального планирования и расчета коэффициентов модели на основе 1/16 реплики активного эксперимента для семифакторной зависимости с одним квадратичным членом

Основные опыты $N$	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_7^2$
1	-	-	-	+	+	+	-	+
2	+	-	-	-	-	+	+	+
3	-	+	-	-	+	-	+	+
4	+	+	-	+	-	-	-	+
5	-	-	+	+	-	-	+	+
6	+	-	+	-	+	-	-	+
7	-	+	+	-	-	+	-	+
8	+	+	+	+	+	+	+	+
9	+	-	-	+	-	+	0	0
10	-	+	-	-	+	-	0	0
11	+	-	+	-	+	-	0	0
12	-	+	+	+	-	+	0	0

Для четырех, пяти, шести и семи факторов композиционное планирование потребует дополнить матрицу табл. 5.13 четырьмя опытами с «нулевыми» точками для фактора, обусловленного тройным смешиванием, хотя в последнем случае для этих четырех опытов фактор  $X_4$  целесообразно задать генерирующим соотношением  $X_4 = X_2 X_3$  для выполнения условий (5.11) и (5.12), как это сделано в табл. 5.15. Эксперимент, выполненный по полученным планам, несложно обработать описанными ранее методами, что приведет к его значительному упрощению по сравнению с другими подходами. Что касается планов с одновременным заданием квадратичных зависимостей для двух, трех и более факторов, то указанных планов, приводящих к сложным экспериментам, даже в случае их острой необходимости, желательно избегать.

Приведем некоторые интерполяционные модели, отражающие ригидное изменение факторов, оказывающих влияние на топливную экономичность и экологическую безопасность двигателей, которые могут дать полезную информацию при планировании новых исследований в этом актуальном направлении.

Так, для карбюраторного двигателя показатель экономичности определялся по величине циклового расхода топлива  $G_{цикл}$ , см<sup>3</sup> на 100 оборотов коленчатого вала двигателя, на прямой передаче при различных скоростях движения автомобиля.

При этом исследовалось девять факторов:  $X_1$  — нагрузка автомобиля;  $X_2$  — уровень топлива в поплавковой камере;  $X_3$  — пропускная способность главного жиклера первой смесительной камеры;  $X_4$  —

момент включения экономайзера;  $X_5$  — качество регулировки холостого хода;  $X_6$  — заедание привода воздушной заслонки, приводящее к ее неполному открытию;  $X_7$  — зазор в контактах прерывателя;  $X_8$  — работоспособность вакуумного регулятора;  $X_9$  — работоспособность свечей. Были получены следующие результаты:

$$\begin{aligned} G_{\text{цикл}} (40 \text{ км/ч}) &= 6,62 + 0,46X_1 + 0,53X_3 + 0,41X_4 + \\ &+ 0,35X_8 - 0,49X_9; \\ G_{\text{цикл}} (50 \text{ км/ч}) &= 7,72 + 0,66X_3 + 0,25X_4 - 0,28X_5 - \\ &- 0,21X_6 - 0,57X_9; \\ G_{\text{цикл}} (60 \text{ км/ч}) &= 8,17 + 0,88X_3 + 0,27X_4 - 0,31X_5 - \\ &- 0,25X_6 - 0,94X_9; \\ G_{\text{цикл}} (80 \text{ км/ч}) &= 10,4 + 0,44X_2 + 0,92X_3 + 0,27X_4 - \\ &- 0,28X_5 - 0,45X_6 - 0,14X_9 - 0,45X_1X_4 + 0,32X_3X_4. \end{aligned} \quad (5.35)$$

Можно видеть, что хотя на трех первых режимах оказалась незначима почти половина линейных факторов, при повышенной скорости проявили себя и некоторые парные взаимодействия. Поэтому выбор реплики  $2^{9-5}$  с 16 основными опытами в целом оказался целесообразным.

Более сложным оказалось планирование для определения ранжирования влияния небольшого количества факторов на показатели экологической безопасности дизеля (на максимальной частоте вращения коленчатого вала  $G_x$ ,  $\text{м}^{-1}$ ) с одновременными оценками по изменению значения такого диагностического параметра, как ускорение свободного разгона коленчатого вала  $\epsilon$ ,  $\text{рад}/\text{с}^2$ . Исследовались факторы:  $X_1$  — величина максимальной цикловой подачи;  $X_2$  — изменение угла опережения впрыска топлива от номинала;  $X_3$  — давление начала подъема иглы распылителя;  $X_4$  — герметичность компрессионных колец;  $X_5$  — отклонение от номинала теплового зазора впускного клапана;  $X_6$  — отклонение от номинала теплового зазора выпускного клапана. Поскольку для трех факторов ( $X_2$ ,  $X_5$  и  $X_6$ ) исследования необходимо было проводить с «нулевой» точкой, то это можно было реализовать только на основе симметричного композиционного плана (табл. 5.16), при котором все факторы приходится задавать на трех уровнях.

(Коэффициенты для квадратичных членов и парных эффектов для данного плана затем получаются, как неоднократно отмечалось ранее, при обработке расширенной матрицы с включением в нее необходимых столбцов.) Хотя основой композиции являлась четверть-реплика  $2^{6-2}$  с 16 основными опытами, их общее количество возросло до 29, т.е. примерно вдвое по сравнению с вариантом, при котором на основе более простого предварительного трехфакторного эксперимента можно было бы исключить как некоторые факторы, так и необходимость их исследования на трех уровнях.

Симметричный композиционный план второго порядка для шестифакторной зависимости на основе четверть-реплики активного эксперимента для линейной модели

Опыты N	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	Опыты N	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
1	+	+	+	+	+	+	16	-	-	-	-	-	-
2	-	+	+	+	-	-	17	+	0	0	0	0	0
3	+	-	+	+	-	+	18	-	0	0	0	0	0
4	-	-	+	+	+	-	19	0	+	0	0	0	0
5	+	+	-	+	-	-	20	0	-	0	0	0	0
6	-	+	-	+	+	+	21	0	0	+	0	0	0
7	+	-	-	+	+	-	22	0	0	-	0	0	0
8	-	-	-	+	-	+	23	0	0	0	+	0	0
9	+	+	+	-	+	-	24	0	0	0	-	0	0
10	-	+	+	-	-	+	25	0	0	0	0	+	0
11	+	-	+	-	-	-	26	0	0	0	0	-	0
12	-	-	+	-	+	+	27	0	0	0	0	0	+
13	+	+	-	-	-	+	28	0	0	0	0	0	-
14	-	+	-	-	+	-	29	0	0	0	0	0	0
15	+	-	-	-	+	+							

Такую возможность можно видеть по полученным конечным интерполяционным моделям:

$$\begin{aligned} G_x &= 0,047 - 0,013X_1 + 0,012X_4 - 0,013X_2X_6 + 0,033X_6^2; \\ \epsilon &= 120 + 7,3X_1 + 1,2X_2 - 2X_4 + 1,3X_6. \end{aligned} \quad (5.36)$$

Информационная ценность полученных моделей заключается в окончательном обосновании и ранжировании факторов, влияющих на дымность дизеля, в целях создания для них соответствующей системы диагностирования, а также в возможности оценки состояния одного из факторов — максимальной цикловой подачи — по параметру ускорения свободного разгона, удобного с метрологических и технологических показателей для процессов технического обслуживания автомобилей в стационарных и полевых условиях.

Из приведенных ранее примеров записей матриц планирования основных экспериментов (в сочетании со столбцами для получения коэффициентов моделей) можно видеть достаточную громоздкость подобного «графического» представления. Для ее сокращения и решения других задач используется метод условных обозначений строк непосредственной матрицы планирования буквами латинского алфавита [1].

Для этого порядковый номер фактора ставится в соответствие строчной букве латинского алфавита ( $X_1 - a$ ,  $X_2 - b$ ,  $X_3 - c$  и т.д.),

которые выписываются для соответствующих строк матрицы только при значениях факторов на верхнем уровне; опыт со всеми факторами на нижних уровнях принято обозначать (1). Тогда четыре опыта плана табл. 5.7 однозначно запишутся комбинацией: (1),  $a$ ,  $b$ ,  $ab$ ; соответственно восемь опытов плана табл. 5.13 будут отражены комбинацией: (1),  $a$ ,  $b$ ,  $ab$ ,  $c$ ,  $ac$ ,  $bc$ ,  $abc$ . Буквенная запись обеспечивает удобный алгоритм разбиения плана полного факторного линейного эксперимента на блоки [1], при использовании которых в качестве реплик могут быть учтены ошибки от влияния внешней среды (при этом фактор, подверженный влиянию, обычно приравнивается к взаимодействию самого высокого порядка или не имеющему физико-химического смысла). Однако она не может быть применена для планов второго порядка и обладает меньшей наглядностью по сравнению с матричным представлением.

#### 5.4. Планирование эксперимента при поиске оптимальных условий

Как уже отмечалось ранее, в области технической эксплуатации автомобилей данное направление активного эксперимента для целенаправленного поиска условий для получения наилучших значений функции отклика практически не реализовывалось, а в литературе по исследованиям проблем автомобильно-дорожной техники [2] давались только общие сведения о методических возможностях таких задач. Поэтому наиболее правильно рассмотреть особенности данного направления на примерах, заимствованных из работ, посвященных указанному вопросу [1], по аналогии с которыми затем правильно проводить исследования в области технической эксплуатации.

При этом целесообразно ограничиться рассмотрением наиболее простого метода планирования — Бокса—Уилсона, о котором уже упоминалось ранее. Его идея заключается в последовательном проведении небольшой серии многофакторных линейных опытов, которые организуются таким образом, чтобы после обработки результатов предыдущей серии можно было выбрать условия проведения следующей серии, т.е. ее спланировать. Так последовательно достигается область оптимума.

Данный подход, носящий название градиентного метода, или крутого *восхождения* по поверхности отклика (с постоянным или переменным шагом), является одной из разновидностей итеративных методов [1], включающих в себя метод слепого перебора, метод направленного поиска, метод конфигураций, а также метод наискорейшего спуска (при решении задач на минимум) при переменном (оптимальном) шаге. Однако все указанные методы преимущественно применяются уже при квадратичных, а не чисто линейных моделях,

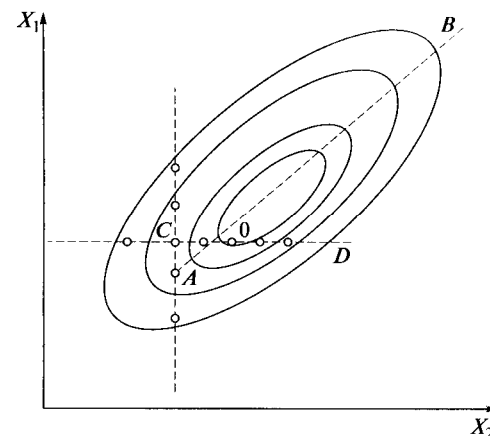


Рис. 5.5. Движение по поверхности отклика методами однофакторного эксперимента и градиента

на которых базируется подход Бокса—Уилсона, и в этом заключается его основное достоинство.

Идея градиентного метода легко поясняется на примере двухфакторной зависимости вида рис. 5.4, которую можно представить проекциями сечений поверхности отклика на плоскость факторов  $X_1$ — $X_2$  линиями равного отклика, подобно линиям равной высоты на географических картах (рис. 5.5).

При этом становится «видна» область оптимума и можно реализовать простейший экспериментальный метод определения значений параметров  $X_1$  и  $X_2$  для точки 0, находящейся внутри этой области, свидетельствующей с определенной точностью о достижении области оптимума.

Указанный метод, называемый классическим, или методом Гаусса—Зейделя, заключается в том, что сначала последовательно изменяют значения одного фактора ( $X_1$  на рис. 5.5), а после нахождения его наилучшего значения (перехода из точки  $A$  в точку  $C$ ) производят эксперименты по изменению и нахождению наилучшего значения для второго фактора (по линии  $CD$ ); затем процедура повторяется для первого фактора (и т.д.), пока не будет экспериментального подтверждения достижения области оптимума (вершины «холма»). При этом постулируется (предполагается), что поверхность отклика является непрерывной, гладкой и имеет единственный оптимум, который может быть и на границе уровней определения факторов (данный постулат в основном подтверждается исследованиями при решении практических задач).

Можно видеть, что даже в случае двухфакторной зависимости такой метод требует большого числа опытов, а при добавлении факторов количество опытов становится чрезмерным.

Резкое сокращение количества опытов произойдет в случае целенаправленного движения к вершине «холма» по направлению градиента функции отклика, показанному на рис. 5.5 прямой  $AB$ , перпендикулярной линиям уровня. В общем случае градиент  $\nabla\varphi$  непрерывной однозначной функции  $\varphi$  есть вектор

$$\nabla\varphi = \frac{\partial\varphi}{\partial X_1}i + \frac{\partial\varphi}{\partial X_2}j + \dots + \frac{\partial\varphi}{\partial X_k}k, \quad (5.37)$$

где  $\partial\varphi/\partial X_i$  — частная производная функции по  $i$ -му фактору;  $i, j, k$  — единичные векторы в направлении координатных осей.

Для рассматриваемых линейных моделей составляющие градиента представляют собой частные производные функции отклика, оценками которых являются коэффициенты регрессии. Для примера (см. рис. 5.5) градиент по выражению (5.37) можно получить, проводя соответствующим образом активный двухфакторный эксперимент с центром (нулевой точкой) в точке  $A$ , при котором границы варьирования факторов находятся на склоне «холма» и не перекрывают область оптимума, как это показано на рис. 5.6. Тогда, изменяя независимые переменные пропорционально величинам коэффициентов регрессии, можно двигаться в направлении градиента функции отклика по самому крутому пути. Процедура движения к почти стационарной области называется крутым восхождением.

Расчет крутого восхождения базируется как на формализованных, так и на неформализованных подходах. Основой его является шаг движения, задаваемый одновременным умножением абсолютных (не нормированных) величин коэффициентов модели на любое положительное число. При этом получаются точки, заданные физическими координатами факторов, лежащие на том же градиенте, но с другим шагом. Процедура движения заключается в том, чтобы ста-

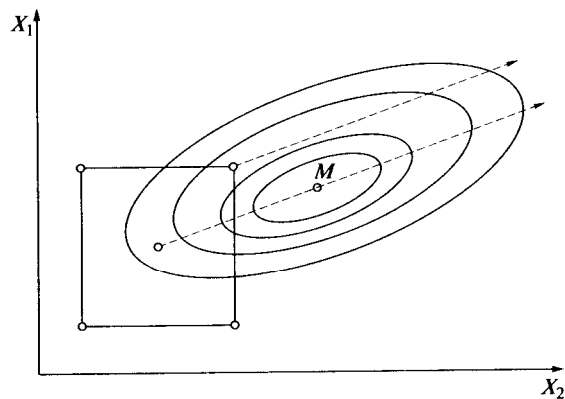


Рис. 5.6. Движение по градиенту из нулевой и из наилучших точек плана

вить новые опыты в точках, которые находятся путем последовательного алгебраического прибавления к нулевому уровню величин, пропорциональных составляющим градиента. Составляющие градиента по каждому фактору определяются перемножением коэффициентов модели на величину интервала варьирования. Из рис. 5.6 можно видеть, что если движение по градиенту начинать от наилучшего опыта (а не нулевого уровня), то велика вероятность провести серию опытов в стороне от оптимальных условий.

Очевидно, что небольшой шаг потребует значительного числа опытов при движении к оптимуму, большой шаг увеличивает вероятность проскока области оптимума. Поэтому рекомендуется [1], по аналогии с выбором интервалов варьирования, нижнюю границу шага задавать возможностью фиксирования двух соседних опытов, а верхнюю — областью определения фактора. Для облегчения работы шаги по факторам обычно округляют для получения экспериментальных точек с удобными параметрами для постановки опытов.

На расчет градиента не оказывает влияние значение коэффициента  $B_0^*$ . Для качественных факторов на двух уровнях либо фиксируется лучший уровень, либо градиент реализуется дважды для каждого уровня в отдельности. Незначимые факторы, которые могут стать такими ввиду их нахождения в зоне, оказывающей несущественное влияние на функцию отклика или при малом интервале варьирования, стабилизируются на любом уровне в интервале  $\pm 1$ . Если нет специальных соображений, то для них выбирают нулевой уровень. Если же по экономическим соображениям, например, выгодно поддерживать нижний уровень, то выбирают его. При движении по градиенту указанные факторы не участвуют.

Поясним указанные положения на примере двухфакторного эксперимента, заимствованного из [1]. Здесь ставилась задача получить максимальное разделение смеси редкоземельных элементов (в процентах) за счет подбора оптимальных значений параметров технологического процесса  $X_1$  и  $X_2$ . В результате полного факторного эксперимента, все этапы планирования и проведения которого отражены в табл. 5.17, была получена модель

$$\hat{Y} = 88 - 2X_1 - 4,5X_2, \quad (5.38)$$

которая явилась основой для проведения последующих этапов расчета крутого восхождения.

1. Расчет составляющих градиента:

$$B_1^*I_1 = -2 \cdot 0,5 = -1,0; \quad B_2^*I_2 = -4,5 \cdot 1,0 = -4,5, \quad (5.39)$$

где  $I_1$  и  $I_2$  — интервалы варьирования для 1-го и 2-го факторов.

Теперь необходимо прибавлять составляющие градиента к основному уровню факторов. При этом условия 5-го опыта в физических величинах составят:  $X_1 = 0,5$  и  $X_2 = 2,5$ , т. е. уже существенно выходят за пределы ранее выполненного эксперимента; для 6-го опыта они

Таблица 5.17

**Матрица планирования, результаты эксперимента  
и расчет крутого восхождения при поиске оптимальных условий  
для двухфакторной зависимости**

Физический уровень факторов	$\tilde{X}_1$	$\tilde{X}_2$	
Основной	1,5	7,0	
Интервал варьирования	0,5	1,0	
Верхний	2,0	8,0	
Нижний	1,0	6,0	
Опыты в кодированных значениях	$X_1$	$X_2$	$\bar{Y}_i, \%$
1	-1	-1	95,0
2	+1	-1	90,0
3	-1	+1	85,0
4	+1	+1	82,0
$B_j$	-2,0	-4,5	
$B_j I_j$	-1,0	-4,5	
Шаг при изменении $\tilde{X}_2$ на 0,5	-0,11	-0,5	
Округление	-0,1	-0,5	
Мысленные опыты	$\tilde{X}_1$	$\tilde{X}_2$	$\hat{Y}_i, \%$
5	1,4	6,5	
6	1,3	6,0	
7	1,2	5,5	95,95
8	1,1	5,0	98,6
9	1,0	4,5	101,25

уже имеют нереальные значения. Отсюда можно сделать вывод, что шаг движения велик.

2. Для выбора более удобного шага используется условие: умножение составляющих градиента на любое положительное число дает точки, также лежащие на градиенте. Для данной задачи удобно изменять физическое значение фактора более важного фактора  $X_2$  на 0,5, поскольку при движении от основного уровня уже на втором опыте достигается граница (или возможен выход за ее пределы) ранее выполненного эксперимента (см. табл. 5.17).

В общем случае рекомендуется [1] выполнять один или два опыта крутого восхождения в области эксперимента для неадекватной модели; для адекватной модели такие опыты не являются необходимыми, поскольку их результаты можно получить простым расчетом по уже полученной зависимости, а уже проводить эксперименты для значений, выходящих хотя бы по одному фактору за пределы области эксперимента.

В указанном примере неформальный выбор диапазона изменения фактора  $X_2$  на 0,5 уменьшает составляющую градиента в девять раз, соответственно во столько же раз необходимо уменьшить составляющую градиента и по фактору  $X_1$  (до -0,11). Однако поскольку метод анализа позволяли задавать значения фактора  $X_1$  с точностью до одного знака после запятой, произведено округление шага до значения -0,1.

3. Последовательное прибавление составляющих градиента к основному уровню. Получается серия опытов крутого восхождения (опыты 5—9 в табл. 5.17), условия которых заданы физическими величинами. Эти опыты часто называют мысленными. Как уже отмечалось ранее, указанную серию целесообразно реализовывать начиная с 7-го опыта.

Считается полезным и необходимым оценивать ожидаемые значения параметра оптимизации для некоторых мысленных опытов до начала их реализации. Для рассматриваемого примера такими опытами являются 7, 8, 9-й. Для расчета  $\hat{Y}_i$  по модели (5.38) для этого надо перевести натуральные (физические) значения факторов  $X_1$  и  $X_2$  в их кодовые величины по выражению (5.1), что дает:  $X_1 = -0,6$  и  $X_2 = -1,5$  для 7-го опыта;  $X_1 = -0,8$  и  $X_2 = -2,0$  для 8-го опыта;  $X_1 = -1,0$  и  $X_2 = -2,5$  для 9-го опыта. Отсюда параметры отклика будут равны 95,95; 98,6 и 101,25 соответственно.

Последний результат хотя и не имеет физического смысла (разделение вещества не может быть более чем на 100%), однако подсказывает возможность попадания в область оптимума и целесообразность реализации опыта при данных физических значениях факторов. Таким образом, анализ результатов мысленных опытов «подсказывает» ненужность реализации 7-го опыта и возможность достижения оптимума при технологических условиях всего двух дополнительных опытов: 8-го и 9-го. Очевидно, что полученные в этих опытах значения параметра оптимизации могут существенно отличаться от расчетных.

В общем случае реализация выбранных мысленных опытов, общее рекомендуемое количество которых — не более 10, может быть осуществлена как последовательно, так и одновременно (если опыты требуют много времени). При последовательной стратегии наиболее целесообразно реализовывать опыты по одному и после каждого анализировать результаты. Другой вариант — ставится одновременно два три опыта и потом принимается решение.

При этом возможен метод «ножниц», при котором реализуются два крайних мысленных опыта, а затем прощупывается пространство внутри этого интервала. Крутое восхождение считается эффективным, если хотя бы один из реализованных опытов даст лучший результат по сравнению с наилучшим опытом серии. Если не удастся достигнуть оптимума в реализованных опытах, то выбирается наилучший из них и его условия принимаются за основной уровень факторов

в следующей серии опытов. Если в одном из реализованных опытов достигаются оптимальные условия, то эксперимент заканчивается. Если при эффективном крутом восхождении достигается граница области определения одного из факторов, то рекомендуется зафиксировать его значение и дальше двигаться по остальным факторам, продолжая расчет мысленных опытов и выбирая стратегию их реализации.

В подавляющем большинстве случаев ситуация при поиске оптимальных условий будет значительно сложнее по сравнению с примером табл. 5.17, особенно при увеличении числа факторов, поэтому полезно знать общие рекомендации получения конечного решения, которые выработаны на практике [1]. Если область оптимума достигнута, то согласно концепции Бокса—Уилсона, можно прекращать исследования (если полученные значения параметров считаются достаточными для практической реализации) или описать область оптимума методами нелинейного планирования (полиномом 2-й степени), с использованием которого решить задачу о самой лучшей точке — экстремуме.

Если область оптимума не достигнута, хотя крутое восхождение привело к заметному росту значения параметра оптимизации и имеется целесообразность его дальнейшего повышения, то рекомендуется или исследовать область вблизи наилучшего значения методами нелинейного планирования, или перейти ко второму этапу крутого восхождения на основе линейной модели с центром нового плана в той части факторного пространства, которая соответствует условиям наилучшего опыта при крутом восхождении. Последний вариант, как правило, связан с меньшим числом опытов.

Если крутое восхождение оказалось неэффективным, а область оптимума близка, то рекомендуется или заканчивать исследования (при этом его результатом будет являться наилучший опыт), или строить план второго порядка для описания области оптимума. Возможны также редкие ситуации, когда область оптимума далека при адекватной линейной модели. Это может быть связано с провалами функции отклика, при которых ее поверхность теряет постулированные выше условия гладкости и одноэкстремальности. В таких случаях целесообразно передвинуться в другую область факторного пространства и построить линейный план второго цикла крутого восхождения.

Наиболее сложной является ситуация, когда область оптимума далека при неадекватной линейной модели. Здесь возможно единственное решение — возвратиться к исходной точке и выяснить причины неадекватности линейной модели. Основными из них могут быть следующие:

- интервалы варьирования выбраны неудачно. Признаком этого является асимметричность модели, при которой ее коэффициенты на порядок отличаются друг от друга. Увеличение интервалов варь-

ирования для мало значащих факторов может привести к устранению этого недостатка, делая модель удобной для расчета крутого восхождения;

- исходная модель строилась на полуреплике. Нужно достроить полуреплику до полного факторного эксперимента, получить раздельные коэффициенты для всех коэффициентов регрессии и совершить новое крутое восхождение;

- исходная модель строилась на дробной реплике. Целесообразно использовать метод «перевала», т. е. построить матрицу второй серии опытов, изменив все знаки на обратные. Это даст возможность отделить линейные эффекты от совместных оценок с парными взаимодействиями. Однако положение не улучшится, если значимыми будут являться взаимодействия более высокого порядка, хотя последнее в сфере технической эксплуатации автомобилей практически исключено.

Если нет информации о положении оптимума и на стадии крутого восхождения не удалось улучшить значение параметра оптимизации, то можно рекомендовать дополнительно поставить опыты в центре эксперимента, с тем чтобы оценить вклад квадратичных членов. При их значительной сумме можно приступить к достройке линейного плана до плана второго порядка, так как наличие квадратичных членов свидетельствует о близости к почти стационарной области. Однако надо иметь в виду, что при незначительной сумме в силу разности знаков коэффициентов квадратичных членов может иметь место два оптимума и выбрать нужное решение довольно трудно. Если же есть основание полагать, что оптимум один, то при незначительной сумме квадратичных членов можно приступить ко второму циклу крутого восхождения.

Покажем некоторые из отмеченных ранее особенностей поиска оптимальных условий на более сложном примере, где ставилась задача обеспечения максимального выхода фармацевтического продукта (в процентах) в зависимости от трех технологических факторов. Для ее решения было выполнено две серии опытов по крутому восхождению. В результате первой серии (табл. 5.18) было получено увеличение выхода реакции с 46,8 до 72,5 %. При этом запланированные шесть мысленных опытов (их количество обуславливалось достижением границы по первому фактору) реализовались избирательно.

Десятый опыт не реализовался, поскольку находился в пределах уже выполненной серии, а 10-й опыт по мнению исследователей по своим параметрам еще не очень далеко выходил за границы второго и третьего факторов, которые в основном обуславливали увеличение выхода реакции.

Поэтому сначала был реализован 11-й опыт, а по его результатам — 13-й и 14-й опыты. При этом 13-й опыт оказался в «вилке» между 11-м и 14-м опытами, что свидетельствовало о близости оптимальной области.

Таблица 5.18

Матрица планирования, результаты эксперимента и расчеты первой серии крутого восхождения при поиске оптимальных условий для трехфакторной зависимости

Физический уровень факторов первой серии опытов	$\tilde{X}_1$	$\tilde{X}_2$	$\tilde{X}_3$	
Основной	0,7	135	30	
Интервал варьирования	0,2	5	15	
Верхний	0,9	140	45	
Нижний	0,5	130	15	
Опыты в кодированных значениях	$X_1$	$X_2$	$X_3$	$\bar{Y}_i, \%$
1	+1	+1	+1	46,80
2	+1	-1	+1	20,47
3	-1	-1	+1	16,80
4	-1	-1	-1	5,08
5	+1	+1	-1	24,15
6	+1	-1	-1	8,89
7	-1	+1	-1	16,63
8	-1	+1	+1	46,45
$B_j$	1,92	10,35	9,47	
$B_j I_j$	0,384	51,51	140,4	
Шаг при изменении $\tilde{X}_2$ на 5	0,0373	5	14,3	
Округление	0,03	5	14,0	
Мысленные опыты крутого восхождения первой серии	$\tilde{X}_1$	$\tilde{X}_2$	$\tilde{X}_3$	$\bar{Y}_i, \%$
9	0,73	140	44	—
10	0,76	145	58	—
11	0,79	150	72	66,70
12	0,82	155	86	—
13	0,85	160	100	72,50
14	0,88	165	114	68,40

Параметры наилучшего 13-го опыта и выявленные особенности протекания процесса использовались при выборе локальной области факторного пространства на втором цикле планирования (табл. 5.19).

Здесь прежде всего учитывалось технологическое обстоятельство о нежелательности повышения значений показателя первого фактора, который в первой серии крутого восхождения хотя и изменялся медленно, однако достиг критической верхней границы, поэтому при выборе второй серии опытов значение фактора  $X_1$  было уменьшено в предположении того, что также будет наблюдаться его рост.

Таблица 5.19

Матрица планирования, результаты эксперимента и расчеты второй серии крутого восхождения при поиске оптимальных условий для трехфакторной зависимости

Физический уровень факторов второй серии опытов	$\tilde{X}_1$	$\tilde{X}_2$	$\tilde{X}_3$	
Основной	0,5	160	55	
Интервал варьирования	0,2	6	15	
Верхний	0,7	166	70	
Нижний	0,3	154	40	
Опыты в кодированных значениях	$X_1$	$X_2$	$X_3$	$\bar{Y}_i, \%$
1	+1	+1	+1	76,00
2	+1	-1	+1	74,05
3	-1	-1	+1	80,90
4	-1	-1	-1	73,00
5	+1	+1	-1	76,81
6	+1	-1	-1	62,65
7	-1	+1	-1	81,40
8	-1	+1	+1	82,40
$B_j$	-3,52	3,25	2,44	
$B_j I_j$	-0,705	19,51	36,54	
Шаг при изменении $\tilde{X}_1$ на 0,1	-0,1	2,76	5,17	
Округление	-0,1	3	5	
Мысленные опыты крутого восхождения второй серии	$\tilde{X}_1$	$\tilde{X}_2$	$\tilde{X}_3$	$\bar{Y}_i, \%$
9 (нулевой уровень)	0,5	140	55	77,3
10	0,4	163	60	82,2
11	0,3	166	65	—
12	0,2	169	70	83,6
13	0,1	172	75	84,8
14	0	175	80	36,9

Повышение значений фактора  $X_2$  являлось технологически желательным, и для нулевого уровня второй серии было взято его значение для наилучшего опыта первой серии. Чрезмерное повышение фактора  $X_3$ , который на стадии крутого восхождения первой серии имел наибольший относительный рост (почти в 4 раза от нулевого уровня и в 2,5 раза от верхнего уровня), также являлось нежелательным, поэтому исследователями было принято компромиссное решение о менее существенном повышении нулевого уровня (примерно в 2 раза) для второй серии.

Проведение полнофакторного эксперимента из восьми опытов второй серии показало возможность дальнейшего увеличения выхода реакции (до 82,4%), хотя предполагалось, что при удачном подборе условий выход может быть повышен до 95...97%. Новое уравнение регрессии, данные которого приведены в табл. 5.19, имело симметричные (равнозначные) коэффициенты регрессии, однако для первого фактора знак изменился на противоположный, что свидетельствовало о необходимости пересмотра старых представлений о его технологическом значении.

Линейная модель оказалась неадекватной (критерий Фишера даже не вышел за пределы менее жестких 5%-ных ограничений), что свидетельствовало о нахождении показателей факторного пространства по функции отклика вблизи области оптимума.

Для ее качественного исследования в целях решения задачи о дальнейшем повышении выхода реакции фармацевтического продукта необходимо было переходить к нелинейному планированию для получения адекватных квадратичных зависимостей, однако это было связано со значительным количеством дорогостоящих опытов.

Поэтому исследователи исходя из своей квалификации приняли некорректное решение провести дополнительно менее сложные эксперименты крутого восхождения, хотя вероятность успеха в случае неадекватности линейного приближения и близости области оптимума была очень мала. В данном случае рассматривалась возможность дополнительного снижения показателей параметра  $X_1$  (знак коэффициента регрессии минусовой), выгодного с технологической точки зрения.

Результаты крутого восхождения приведены в нижней части табл. 5.19.

Поскольку модель была неадекватной, то два опыта (9-й и 10-й) были реализованы внутри факторного пространства; 11-й опыт находился на границах факторного пространства, поэтому его воспроизведение было нецелесообразным. Реализация оставшихся трех мысленных опытов показала незначительное увеличение выхода реакции по сравнению с лучшим результатом в матрице (всего на 2,4%), однако при этом технологические показатели первого фактора оказались в 3 раза лучше.

Кроме того, было более наглядно подтверждено достижение оптимума, делающее нецелесообразным проведение сложных исследований по нелинейному планированию. Не менее важным следует считать изменение общих представлений о характере процесса, поскольку ранее считалось, что она может протекать при физическом значении первого параметра более 0,7, а выход реакции свыше 70% был неизвестен.

Данные примеры могут оказаться методически полезными при решении задач в области техники.

## 5.5. Нелинейное планирование активного эксперимента для получения моделей многофакторных зависимостей второго порядка и поиска экстремальных значений функции отклика

Модели второго порядка, которые зачастую приходится использовать для решения задач в области технической эксплуатации, при ортогональном планировании требуют постановки значительного числа экспериментов, поэтому подобное планирование используется только для двухфакторных моделей (см. табл. 5.11), если необходимые для исследования девяти основных опытов технологически несложно осуществить. В других ситуациях, а также при увеличении количества факторов рекомендуется применять другие (рассмотренные далее) планы, подобные симметричному композиционному плану табл. 5.16, где для исследования шести квадратичных факторов требовалась постановка 29 опытов (вместо 729).

Из приведенного примера можно видеть популярность так называемых композиционных планов, которые имеют и другие разновидности. Так, на рис. 5.7 показан композиционный план для двухфакторной зависимости, также требующий постановки девяти основных опытов, но при большем числе факторов планирование по указанному принципу приводит к уже более существенному снижению объемов эксперимента (табл. 5.20) [2]. Матрица плана (рис. 5.7) состоит из композиции линейного ядра (1—4-й опыты), звездных точек (5—8-й опыты) и точек в центре плана. Таким образом, для данного

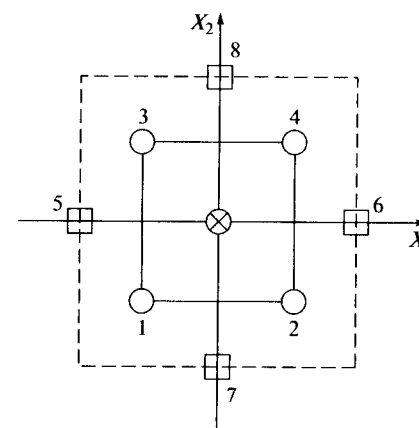


Рис. 5.7. Точки факторного пространства, соответствующие центральному композиционному планированию, для двухфакторной квадратичной модели (точки линейного ядра плана обозначены кружочками; звездные точки — квадратами; точки в центре плана — крестиками)



Таблица 5.20

## Сравнительные характеристики планов ОЦКП и РЦКП

Параметры	Обозначения	Планы ОЦКП				Планы РЦКП			
		Число факторов				Число факторов			
		2	3	4	5	2	3	4	5
Линейное ядро	$N = 2^K$	4	8	16	32	4	8	16	32
Число звездных точек	$m$	4	6	8	10	4	6	8	10
Величины звездных плеч	$p$	1	1,2	1,41	1,68	1,41	1,68	2	2,36
Число точек в центре плана	$n(0)$	1	1	1	1	5	6	7	8
Общее число основных опытов	$N_{\text{общ}}$	9	15	25	43	13	20	31	50

композиционного плана каждый фактор уже берется на пяти уровнях (а не трех, как в симметричном плане табл. 5.16) или трех сферах, что повышает точность модели.

Из табл. 5.20 можно видеть, что при пятиуровневом задании факторов различают планы ОЦКП (ортогональное центральное композиционное планирование) и планы РЦКП (ротатабельное центральное композиционное планирование).

Разница указанных планов состоит в том, что плечи звездных точек для них имеют различные значения. Кроме того, в центре плана ставится различное число опытов и обработка результатов (вычисление коэффициентов регрессии, оценка точности и достоверности модели) может производиться различными способами. В частности, для планов ОЦКП после определенной корректировки квадратичных переменных возможно осуществлять вычисление их коэффициентов, как для линейных моделей, по простейшим алгоритмам вида (5.6), когда для планов РЦКП для этого требуется матричное уравнение.

Поскольку модуль Multiple Regression осуществляет обработку данных только на основе универсального матричного уравнения и непосредственно произведет расчеты результатов эксперимента по плану ОЦКП без корректировки квадратичных переменных, данное различие в планах становится несущественным. При этом наиболее важное значение приобретают другие показатели, такие как снижение общего количества опытов для планов РЦКП за счет исключения повторных экспериментов для линейного ядра и звездных точек (общее количество опытов для планов ОЦКП за счет повторов, как минимум, возрастет вдвое по сравнению с данными табл. 5.20) и ро-

тативность, которой обычно отдается предпочтение и которой не обладают планы ОЦКП, но приобретают планы РЦКП при одноразовом воспроизведении звездных точек и линейного ядра и нескольких точек центра плана, необходимых для статистической оценки значимости коэффициентов регрессии и проверки полученной модели на адекватность.

Естественно, что при реализации подобных подходов, положенных в основу и некоторых других приведенных далее планов, невозможно осуществить важную для оценки качества эксперимента проверку на однородность дисперсий воспроизведения, а общая точность полученной модели неизбежно падает за счет резкого сокращения статистики. В то же время именно сокращение общего количества опытов может иметь решающее значение для постановки сложного эксперимента для исследования нелинейной зависимости и весь комплекс указанных выше обстоятельств, обобщенных далее в табл. 5.25, должен учитываться исследователем. Поскольку ротативность и сокращение общего числа опытов являются наиболее важными обстоятельствами, на практике предпочтение отдается планам РЦКП.

В табл. 5.21 приведен пример матрицы планирования, расчета коэффициентов модели и результатов экспериментов по плану РЦКП для трехфакторной модели, по которой можно непосредственно осуществить ввод данных в модуль Multiple Regression (заменяя минусы на 1, а плюсы — на -1). В результате расчетов, которые полезно провести для контроля за правильностью работой с программой, получается следующая модель исследуемого явления:

Таблица 5.21

**Пример матрицы планирования, расчета коэффициентов модели и результатов активного эксперимента согласно плану РЦКП для трехфакторной полной квадратичной зависимости, соответствующих требованиям ввода данных в модуль Multiple Regression (опыты 1...8 соответствуют ядру плана; 9...14 — звездным точкам; 15...20 — центру плана)**

Номер опыта	$X_1$	$X_2$	$X_3$	$X_1^2$	$X_2^2$	$X_3^2$	$X_1X_2$	$X_1X_3$	$X_2X_3$	$Y_{ji}$
1	-	-	-	+	+	+	+	+	+	17
2	+	-	-	+	+	+	-	+	+	50
3	-	+	-	+	+	+	-	+	-	6
4	+	+	-	+	+	+	+	-	-	17
5	-	-	+	+	+	+	+	-	-	20
6	+	-	+	+	+	+	-	+	-	58
7	-	+	+	+	+	+	-	-	+	7
8	+	+	+	+	+	+	+	+	+	20

Номер опыта	$X_1$	$X_2$	$X_3$	$X_1^2$	$X_2^2$	$X_3^2$	$X_1X_2$	$X_1X_3$	$X_2X_3$	$Y_{ij}$
9	-1,68	0	0	2,82	0	0	0	0	0	9
10	1,68	0	0	2,82	0	0	0	0	0	81
11	0	-1,68	0	0	2,82	0	0	0	0	40
12	0	1,68	0	0	2,82	0	0	0	0	7
13	0	0	-1,68	0	0	2,82	0	0	0	14
14	0	0	1,68	0	0	2,82	0	0	0	18
15	0	0	0	0	0	0	0	0	0	16
16	0	0	0	0	0	0	0	0	0	15
17	0	0	0	0	0	0	0	0	0	18
18	0	0	0	0	0	0	0	0	0	13
19	0	0	0	0	0	0	0	0	0	12
20	0	0	0	0	0	0	0	0	0	15

$$\hat{Y} = 15 + 16X_1 - 11,2X_2 + 9,8X_1^2 + 2,1X_2^2 - 6,1X_1X_2, \quad (5.40)$$

а коэффициенты при  $X_3$ ,  $X_3^2$ ,  $X_1X_3$  и  $X_2X_3$  оказываются незначимыми.

Дальнейшее уменьшение числа опытов может быть достигнуто на основе некомпозиционного ротатабельного плана второго порядка [2], при котором для двухфакторной зависимости экспериментальные точки располагаются в вершинах пяти- или шестиугольника (рис. 5.8) и соответствующим образом корректируется матрица планирования и расчета коэффициентов модели (табл. 5.22).

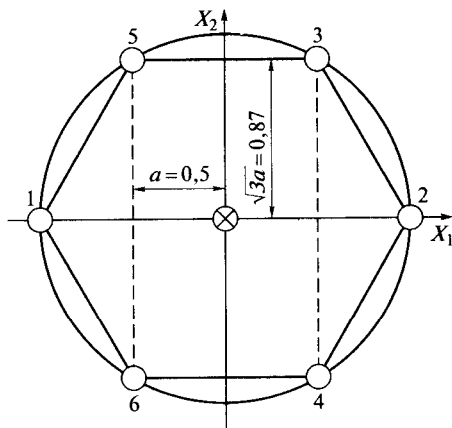


Рис. 5.8. Схема наименьшего некомпозиционного ротатабельного плана для двухфакторной зависимости второго порядка

Матрица планирования и расчета коэффициентов модели согласно наименьшего некомпозиционного ротатабельного плана для определения двухфакторной полной квадратичной зависимости, соответствующая требованиям ввода данных в модуль Multiple Regression

Наименование	Номер опыта	$X_1$	$X_2$	$X_1X_2$	$X_1^2$	$X_2^2$
Ядро плана	1	-	0	0	+	0
	2	+	0	0	+	0
	3	+0,5	+0,87	+0,43	0,25	0,75
	4	+0,5	-0,87	-0,43	0,25	0,75
	5	-0,5	+0,87	-0,43	0,25	0,75
	6	-0,5	-0,87	+0,43	0,25	0,75
Точки в центре плана	7	0	0	0	0	0
	8	0	0	0	0	0
	9	0	0	0	0	0
	10	0	0	0	0	0

По сравнению с матрицей полнофакторного планирования (табл. 5.11) здесь общее число опытов может быть сокращено в два-три раза. Подобным образом составляются матрицы некомпозиционных планов для трех и большего числа факторов, однако при их окончательном выборе целесообразно воспользоваться специально составленным каталогом матриц второго порядка [2].

Кроме рассмотренных ранее планов на практике применяются и другие виды планов: планы Бокса,  $D$ -оптимальные планы, планы Хартли, которые для двухфакторных моделей представлены далее, и др. Указанные планы не являются ортогональными и ротатабельными, однако дают возможность осуществлять проверку однородности дисперсии и имеют приемлемые показатели по точности модели (см. табл. 5.25) при достаточно существенном снижении количества основных и общих опытов.

Так, удобство планов Бокса заключается в расположении звездных точек на  $n$ -мерных гранях; при этом их радиус равен единице, а в центре плана точки не ставятся (рис. 5.9). В этом отношении наиболее сложными являются  $D$ -оптимальные планы (рис. 5.10, табл. 5.23) из-за сложности расположения звездных точек, однако они имеют наименьшее количество основных опытов — всего шесть для двухфакторной модели.

Планы Хартли (рис. 5.11, табл. 5.24), которые можно считать определенной трансформацией планов Бокса, резко снижающей количество основных опытов при увеличении числа факторов,

Матрица планирования и расчета коэффициентов модели, соответствующая  $D$ -оптимальному плану для двухфакторной полной квадратичной модели и требованиям ввода данных в модуль Multiple Regression

Номер опыта	$X_1$	$X_2$	$X_1 X_2$	$X_1^2$	$X_2^2$
1	+	+	+	+	+
2	-	+	-	+	+
3	-	-	+	+	+
4	+	-0,4	-0,4	+	0,16
5	+0,4	-	-0,4	0,16	+
6	-0,1	+ 0,1	-0,01	0,01	0,01

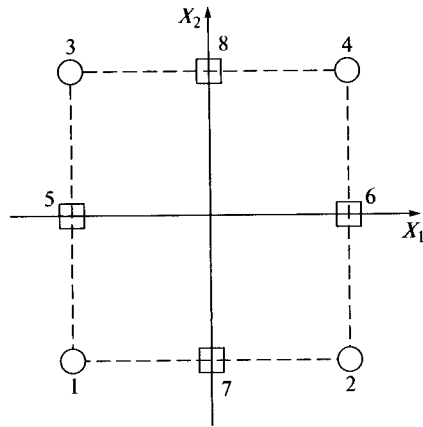


Рис. 5.9. Схема расположения опытных точек линейного ядра (кружки) и звездных точек (квадраты) плана Бокса для двухфакторной квадратичной модели

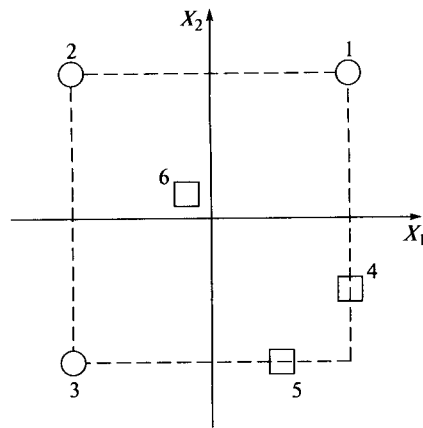


Рис. 5.10. Схема расположения опытных точек  $D$ -оптимального плана для двухфакторной квадратичной модели

хотя и приводящей к более низкой точности (табл. 5.25), являются также предпочтительными и ввиду простоты задания опытных точек. Однако для получения  $D$ -оптимальных планов, планов Бокса и Хартли при числе факторов больше двух необходимо, как и при некомпозиционном планировании, использовать каталог планов второго порядка [8] или модуль Experimental Design (планирование эксперимента), который обычно отсутствует в «рыночной» программе STATISTICA.

На этом фоне привлекательным является симметричное композиционное планирование на основе реплик различной дробности, вариант которого для шести факторов представлен в табл. 5.16. Здесь количество основных опытов соответствует плану Хартли при сохранении ортогональности и можно предполагать не худшую точность модели за счет параллельных опытов. Если схему матрицы табл. 5.16 распространить на семь факторов, то можно видеть, что количество основных опытов (если пренебречь некоторыми парными взаимодействиями, которые всегда имеют место) увеличится незначительно — с 29 до 31, а для восьми факторов — до 33.

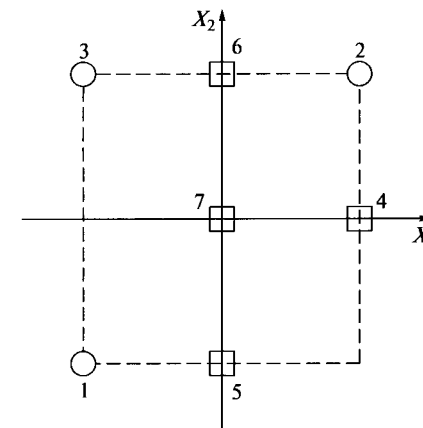


Рис. 5.11. Расположение опытных точек плана Хартли для двухфакторной квадратичной модели

Таблица 5.24

Матрица планирования плана Хартли для двухфакторной квадратичной модели

Номер опыта	$X_1$	$X_2$
1	-	-
2	+	+
3	-	+
4	+	0
5	0	-
6	0	+
7	0	0

Таким образом, планирование эксперимента для получения моделей второго порядка является достаточно сложной задачей, требующей по сравнению с линейными планами учета факторов потери

Таблица 5.25

Сравнительные оценки показателей некоторых планов активного эксперимента для исследования моделей второго порядка [2]

Число факторов	Виды планов для моделей второго порядка	Требуемое число опытов (ядро, звездные точки, точки в центре плана)	Критерии оценки точности модели*	
			$D$ -оптимальность	$G$ -оптимальность
2	Ортогональный	$3^2 = 9$	1,47	1,35
	План ОЦКП	$4 + 4 + 1 = 9$	1,47	1,35
	План РЦКП	$4 + 4 + 5 = 13$	2,1	1,65
	План Бокса-2	$4 + 4 = 8$	1,48	1,65
	$D$ -оптимальный	6	1,54	1,5
3	Ортогональный	$3^3 = 27$	1,5	1,42
	План ОЦКП	$8 + 6 + 1 = 15$	1,65	1,4
	План РЦКП	$8 + 6 + 6 = 20$	2,78	2,27
	План Бокса-3	$8 + 6 = 14$	1,47	1,41
	$D$ -оптимальный	10	1,58	1,77
4	Ортогональный	$3^4 = 81$	1,52	1,46
	План ОЦКП	$16 + 8 + 1 = 25$	2,28	1,86
	План РЦКП	$16 + 8 + 7 = 31$	3,61	3,1
	План Бокса-4	$16 + 8 = 27$	1,48	1,5
	$D$ -оптимальный	15	1,53	1,93
	План Хартли	17	1,85	2,12

Окончание табл. 5.25

Число факторов	Виды планов для моделей второго порядка	Требуемое число опытов (ядро, звездные точки, точки в центре плана)	Критерии оценки точности модели*	
			$D$ -оптимальность	$G$ -оптимальность
5	Ортогональный	$3^5 = 243$	1,52	1,48
	План ОЦКП	$32 + 10 + 1 = 43$	2,47	2,3
	План РЦКП	$32 + 10 + 8 = 50$	4,79	4,32
	План Бокса-5	42	1,48	1,72
	$D$ -оптимальный	15	1,48	1,95
	План Хартли	17	1,53	1,51
6	Ортогональный	$3^6 = 729$	1,52	1,5
	План ОЦКП	$64 + 12 + 1 = 77$	3,25	2,86
	План РЦКП	$64 + 12 + 15 = 91$	6,24	6,07
	План Бокса-6	76	1,53	2,08
	$D$ -оптимальный	35	1,51	2,18
	План Хартли	29	1,98	2,91
7	Ортогональный	$3^7 = 2187$	1,53	1,57
	План РЦКП	163	8,58	8,42
	$D$ -оптимальный	116	1,37	1,75

Для планов с параллельными опытами оценки точности даны при условии проведения двух повторных опытов.

ортогональности, рототабельности и точности. Поскольку модуль Multiple Regression может обрабатывать любые результаты, даже не соответствующие описанным ранее схемам планирования, выбор того или иного плана в конечном счете может сводиться к предполагаемой точности получаемой математической модели. Исходя из этой теории [1, 2] дает две группы критериев, оценивающих оптимальность избранного плана.

Первую группу составляют критерии, связанные с объемом эллипсоида рассеяния дисперсий коэффициентов модели, который является многомерным аналогом обычного доверительного интервала. При этом различают  $D$ -,  $A$ - и  $E$ -оптимальность планов.

Наиболее популярной является  $D$ -оптимальность, за основу оценки которой берется базовая составляющая ковариационной матрицы (4.44) без постоянного сомножителя  $S_e^2$ , отражающего ошибку модели (экспериментальную величину остаточной дисперсии).

Получаемую при этом матрицу применительно к задачам активного многофакторного эксперимента принято называть матрицей дисперсий-ковариаций [1]:

$$M^{-1} = (X^T X)^{-1}. \quad (5.41)$$

Указанная матрица при ортогональном линейном планировании, как было показано на примере выражения (5.7), содержит только одинаковые диагональные элементы, отражающие процедуру планирования (матрицу исходных данных  $X$ ) и общее количество опытов. Применительно к двухфакторной неполной степенной модели производится независимое определение трех ее коэффициентов для двух линейных и одного смешанного членов, ошибки оценки которых, представляемые средними квадратическими отклонениями  $\sigma_{B_1}$ ,  $\sigma_{B_2}$ ,  $\sigma_{B_3}$ , имеют одинаковые величины.

Ввиду ортогональности планирования и независимости этих ошибок графически их можно представить в виде отрезков на трех перпендикулярных друг к другу координатных осях, а все пространство стандартных ошибок — в виде шара объема

$$V = \frac{4}{3} \pi \sigma_{B_1}^3. \quad (5.42)$$

Для квадратичных моделей дисперсии коэффициентов уже будут отличаться друг от друга, поэтому пространство стандартных ошибок должно быть представлено в виде эллипсоида объема:

$$V = \frac{4}{3} \pi \sigma_{B_1} \sigma_{B_2} \sigma_{B_3}.$$

Чем меньше объем эллипсоида, тем точнее модель. Аналогично можно представить эллипсоид дисперсий стандартных ошибок, который уже не имеет геометрической интерпретации, однако вычисление произведений дисперсий для него аналогично вычислению определителя (детерминанта  $\det$ ) матрицы дисперсий-ковариаций (5.41) при ортогональном планировании (берется произведение диагональных элементов с последующим умножением на общую величину  $S_e^2$ ).

Таким образом, критерием точности модели может являться условный объем эллипсоида рассеяния дисперсий ошибок коэффициентов модели, количественно оцениваемый через значение определителя (детерминанта) матрицы дисперсий-ковариаций  $\det(M^{-1})$ . Вычисление детерминанта несложно осуществить и в случае неортогонального планирования, которое характерно для большинства приведенных ранее планов. Таким образом, для любого планирования условие  $D$ -оптимальности можно записать в виде [1]:

$$\det(M^{-1}) \rightarrow \min. \quad (5.43)$$

Особенности других критериев первой группы:

- при  $A$ -оптимальности уровни факторов модели должны располагаться так, чтобы обеспечивался минимум среднего значения дисперсий коэффициентов модели;
- при  $E$ -оптимальности уровни факторов должны минимизировать наибольшее собственное значение матрицы дисперсий-ковариаций

(эллипсоид должен приближаться к шару, поскольку всегда существует).

**Вторая группа** критериев для оценки оптимальности связана с дисперсией предсказанного значения функции отклика, определяемой по выражению (5.46). Она включает в себя  $G$ - и  $Q$ -критерии. Наиболее популярной является  $G$ -оптимальность, по которой уровни факторов должны располагаться так, чтобы максимальная дисперсия предсказанного значения функции отклика стремилась к минимуму. Из выражения (4.52) можно видеть, что она также, как и  $D$ -оптимальность, может быть оценена по показателям матрицы дисперсий-ковариаций  $M^{-1}$  по выражению

$$\max \sigma_Y^2 = \max[X_i^T M^{-1} X_i] \rightarrow \min, \quad (5.44)$$

где  $X_i^T$  в матричной форме повторяет задаваемую модель второго порядка (в +1, -1 или в других десятичных цифрах со знаками «плюс» или «минус», по абсолютным значениям больших или меньших единице, в соответствии со схемой планирования) для всех основных опытов (см. матрицы табл. 5.21 ... 5.24).

Если план  $G$ -оптимален, то экспериментатор имеет гарантию, что в области планирования не окажется точек, в которых точность оценки поверхности отклика будет слишком низкой. При  $Q$ -оптимальности уровни факторов должны располагаться так, чтобы обеспечить минимум среднего значения дисперсии предсказанного значения функции отклика.

Оценка свойства *ротатабельности*, т.е. обеспечения постоянства дисперсий предсказания поверхности отклика на равных расстояниях от центра эксперимента, также связана со значениями матрицы дисперсий-ковариаций [1]:

$$X_i^T M^{-1} X_i = f(\rho), \quad (5.45)$$

где  $f(\rho)$  — функция, зависящая от величины так называемого осевого расстояния  $\rho = \sqrt{X_1^2 + X_2^2 + \dots + X_n^2}$  точки плана многофакторной модели от центра эксперимента.

Выражение (5.45) в последнее время стали называть дисперсионной функцией [2].

Для линейных и квадратичных ортогональных планов свойство ротатабельности выполняется автоматически, однако для описанных ранее специальных планов, направленных на сокращение количества опытов, указанное свойство может нарушаться. Из анализа табл. 5.20 можно видеть, что при центральных композиционных планах повышению ротатабельности способствует как увеличение расстояний до звездных точек, так и рост количества точек в центре плана. К сожалению, в литературе не даются простые алгоритмы по повышению ротатабельности планов, кроме упомянутых ранее общих рекомендаций и характеристик по данному показателю для уже раз-

работанных планов, поэтому в особых случаях следует его предварительно просчитывать по выражению (5.45) для всех точек плана и анализировать получаемые результаты.

В табл. 5.25 приведены сравнительные оценки показателей некоторых планов второго порядка (по количеству основных опытов и точности получаемых моделей по критериям *D*- и *G*-оптимальности), которые могут быть полезными при планировании исследований в конкретных условиях.

Полученная на основе запланированных тем или иным способом активных экспериментов квадратичная модель обычно подвергается дальнейшим исследованиям в целях определения экстремальных точек (наибольших или наименьших значений функции отклика), по которым уже даются непосредственные практические рекомендации по оптимизации исследуемого процесса. В литературе приводится ряд рекомендаций по проведению этого процесса, сложность которого возрастает по мере увеличения количества факторов, хотя интенсивное внедрение и использование соответствующих компьютерных программ приводит к его упрощению.

Наиболее просто данная задача решается для двух факторов, где поверхность отклика имеет геометрическую интерпретацию и ее можно непосредственно анализировать по внешнему виду объемного графика, получение которого предусмотрено определенными модулями программы STATISTICA.

В идеальном случае в пределах факторного пространства экстремальная точка имеет выпуклость вниз, или минимум, или выпуклость вверх, или максимум, как на рис. 5.4, и ее приблизительные можно найти при помощи зон различной окраски и линий равного отклика, а затем вычислить значение функции отклика. Более точно параметры экстремальной точки получаются на основе решения системы двух линейных уравнений, получаемых при дифференцировании и приравнении к нулю функции отклика по независимым переменным  $X_1$  и  $X_2$ .

В более сложных случаях экстремальное значение приходится на так называемую седловую точку, а поверхность отклика при этом соответствует форме седла для верховой езды. Визуальный анализ затрудняется, когда экстремальная точка находится на достаточном расстоянии за пределами факторного пространства, о чем могут свидетельствовать результаты решения системы из двух линейных уравнений, и для определения характера экстремума необходимо проводить дополнительный анализ.

Для двухфакторной модели он достаточно простой, поскольку знак «+» при коэффициентах для обоих квадратичных членов ( $X_1^2$  и  $X_2^2$ ) указывает на минимум (выпуклость вниз); знак «-» — на максимум (выпуклость вверх); чередование знаков — на наличие седловой точки. Данная информация поможет сократить количество расчетов при определении максимального или минимального значе-

ния функции отклика в пределах факторного пространства или технологически допустимом диапазоне за его пределами, которые необходимо знать для решения задачи оптимизации, а зона расчета в последнем случае может оказаться не очевидной даже при использовании визуального анализа объемного графика.

При увеличении количества факторов графический анализ становится невозможным, а сложность процесса возрастает. Его проведение становится особенно трудным при наличии нескольких экстремумов, на которые могут приходиться выпуклости, вогнутости и седловые точки многомерного пространства, которые экспериментально затрудняют поиски оптимума по методу Бокса—Уилсона, и вложенного ранее, однако сглаживаются за счет проведения активного эксперимента на моделях второго порядка (подобное сглаживание в определенных случаях может являться основным источником ошибочности модели).

В то же время квадратичные модели всегда обеспечивают наличие только одной экстремальной точки, координаты которой определяются на основе решения системы линейных уравнений, получаемых при дифференцировании и приравнении к нулю полученной функции отклика (без чисто линейных составляющих, которые могут появиться после обнуления незначущих коэффициентов модели) по всем независимым переменным. Указанная точка может находиться как в пределах факторного пространства, так и вне его по всем или нескольким факторам, а также иметь в *n*-мерном измерении выпуклость вниз, выпуклость вверх или седловую точку.

Характер экстремума определяется с помощью так называемой матрицы Гессе [2], элементы которой рассчитываются на основе двойного дифференцирования полученной функции отклика  $F(X)$  по схеме

$$H(x) = \begin{vmatrix} \frac{\partial^2 F(X)}{\partial x_1 \partial x_1} & \frac{\partial^2 F(X)}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 F(X)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 F(X)}{\partial x_2 \partial x_1} & \frac{\partial^2 F(X)}{\partial x_2 \partial x_2} & \dots & \frac{\partial^2 F(X)}{\partial x_2 \partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 F(X)}{\partial x_n \partial x_1} & \frac{\partial^2 F(X)}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 F(X)}{\partial x_n \partial x_n} \end{vmatrix}. \quad (5.46)$$

При этом:

- если все главные миноры, получаемые путем вычеркивания из матрицы Гессе строк и столбцов на пересечении диагональных элементов, не отрицательны или равны нулю, то в экстремальной (стационарной) точке функция  $F(X)$  имеет минимум, т.е. имеет место многомерная выпуклость вниз;

- если все главные миноры (их определители) отрицательны (меньше нуля), то в стационарной точке имеет место максимум, т.е. многомерная выпуклость вверх;

- если определители главных миноров имеют как положительные или нулевые, так и отрицательные значения, то в стационарной точке функция отклика не имеет ни максимума, ни минимума, а имеет многомерную седловую точку.

Рассмотрим объективность использования матрицы Гессе на примере графически интерпретируемых двухфакторных зависимостей. Так, для модели

$$\hat{Y}_1 = 14,8 + X_1 - 0,5X_2 - 2,3X_1^2 - 9,8X_2^2 \quad (5.47)$$

матрица будет иметь вид

$$H_1(x) = \begin{vmatrix} -4,6 & 0 \\ 0 & -19,6 \end{vmatrix}, \quad (5.48)$$

а определители ее главных миноров соответственно составят величины  $(-19,6)$  и  $(-4,6)$ , т.е. подтверждается наличие выпуклости вверх. С другой стороны, для модели

$$\hat{Y}_2 = 5,44 + X_1 - 0,5X_2 + 3,3X_1^2 - 4,2X_2^2 \quad (5.49)$$

матрица Гессе представляется зависимостью

$$H_2(x) = \begin{vmatrix} 6,6 & 0 \\ 0 & -8,4 \end{vmatrix} \quad (5.50)$$

при определителях главных миноров  $(-8,4)$  и  $(+6,6)$ , т.е. подтверждается наличие седловой точки.

Приведем пример анализа с использованием матрицы Гессе для трехфакторной квадратичной модели вида

$$\hat{Y}_3 = 20,4 + 15,2X_1 - 10,5X_2 + 7,4X_3 + 8,1X_1^2 + 4,4X_2^2 - 5,3X_3^2 - 6,7X_1X_2 + 4,8X_1X_3. \quad (5.51)$$

Система линейных уравнений для определения координат экстремальной точки будет иметь вид:

$$\begin{aligned} 15,2 + 8,1 \cdot 2X_1 - 6,7X_2 + 4,8X_3 &= 0; \\ -10,5 + 4,4 \cdot 2X_2 - 6,7X_1 &= 0; \\ 7,4 - 5,3 \cdot 2X_3 + 4,8X_1 &= 0. \end{aligned} \quad (5.52)$$

Отсюда координаты экстремальной точки будут равны:  $X_1 = -0,8$ ,  $X_2 = 0,59$ ,  $X_3 = 0,34$ , т.е. она находится в пределах факторного пространства.

Матрица Гессе имеет вид

$$H_3(x) = \begin{vmatrix} 16,2 & -6,7 & 4,8 \\ -6,7 & 8,8 & 0 \\ 4,8 & 0 & -10,6 \end{vmatrix}, \quad (5.53)$$

а определители ее главных миноров равны  $-93,3$ ,  $-172$  и  $+188$ . Таким образом, имеет место многомерная седловая точка и в пределах факторного пространства относительно ее находятся как меньшие, так и большие значения.

Рассчитаем минимальные и максимальные значения функции отклика  $F(X_1; X_2; X_3)$  в пределах факторного пространства, которые могут быть использованы для оптимизации исследуемого процесса. Для экстремальной (седловой) точки функция отклика будет равна:  $F_{\text{седл}}(-0,8; 0,59; 0,34) = 12,5$ . Координаты факторного пространства для максимального и минимального значения функции отклика легко найти по следующему правилу: они должны находиться на пересечении следов двух многомерных осей, проходящих через экстремальную точку, с границами факторного пространства, расположенных на максимальном удалении от экстремальной точки.

На основе полученных значений знаков для определителей главных миноров можно видеть, что такими точками являются:

$$F_{\text{max}}(1; -1; 0,34) = 73,7; F_{\text{min}}(-0,8; 0,59; -1) = 3,0. \quad (5.54)$$

Для данных оценок должны быть найдены доверительные границы по выражениям (4.53) и (4.54) и проведено экспериментальное подтверждение путем постановки одного основного опыта (с двух или трехкратным повтором) в соответствующей точке факторного пространства.

Хотя возможны и другие методы анализа квадратичных многофакторных моделей [2], изложенный ранее простейший подход способен обеспечить решение любых задач в области технической эксплуатации автомобилей. Он может быть также распространен на смешанные модели, в которых имеются как линейные, так и квадратичные зависимости от факторов. Главное — это как можно шире базируя практику исследования на активных экспериментах.

## КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Назовите основные задачи, которые решаются с помощью активного эксперимента.
2. Дайте определение понятиям «целевая функция» и «фактор», перечислите предъявляемые к ним требования.
3. Назовите последовательность проведения активного эксперимента.
4. Назовите и кратко охарактеризуйте методы выбора математической модели при эксперименте.

5. Дайте краткую характеристику полнофакторного активного эксперимента.
6. Дайте характеристику методов варьирования уровнями факторов, приведите примеры.
7. Дайте определение понятия «план эксперимента», приведите примеры.
8. Назовите основные свойства матрицы планирования полнофакторного активного эксперимента первого порядка.
9. Дайте определение понятия «рандомизация плана эксперимента», приведите примеры.
10. Каким образом осуществляется проверка адекватности математической модели?
11. Каким образом осуществляется проверка случайности и независимости наблюдений?
12. Каким образом исключаются грубые ошибки наблюдений?
13. Каким образом проводится проверка однородности дисперсий опытов?
14. Назовите основные преимущества и недостатки полнофакторного активного эксперимента первого порядка.
15. Дайте краткую характеристику дробно-факторного эксперимента, приведите примеры плана эксперимента.

## ГЛАВА 6

### **Особенности компонентного анализа и основные предпосылки для его применения при управлении процессами технической эксплуатацией автомобилей**

#### **6.1. Основные принципиальные подходы при оценке влияющих факторов с использованием многошагового регрессионного и компонентного анализов**

Как известно из анализа литературных источников [3, 8], необходимость применения системного подхода при решении задач управления инженерно-технической службой (ИТС) автотранспорта общего пользования (в частности, в отношении минимизации расходов топлива, запасных частей, трудовых ресурсов и др.) применительно к эксплуатации городских автобусов муниципального (государственного) подчинения требует одновременной оценки степени влияния большого количества факторов. Количественную оценку такого влияния удобно представить в виде полиномиальных моделей многофакторной регрессии, методы нахождения которых с использованием корреляционно-регрессионного анализа изложены ранее (в подразд. 4.3...4.5).

Недостатком такого подхода является то, что при построении многофакторной регрессии взаимосвязанные (мультиколлинеарные) факторы должны выводиться из модели, т. е. модель должна строиться на независимых переменных. В действительности на функцию отклика действует все многообразие факторов, в том числе и взаимосвязанных между собой, и игнорирование этого обстоятельства может привести к существенным ошибкам, что очень нежелательно для практической деятельности предприятий. Поэтому для устранения явления мультиколлинеарности без потери полезной информации в многомерной статистике используют специальный математический аппарат компонентного анализа, эффективность которого можно увидеть далее на конкретных примерах.

Несмотря на сложность теоретических предпосылок компонентного (факторного) анализа его применение для решения практических задач не встречает особых затруднений при использовании специально разработанных компьютерных программ, без которых



вообще невозможно получить необходимые решения, в частности модуля Factor Analysis (Анализ фактора) программы STATISTICA для среды WINDOWS [4] и др. Приведенные далее практические примеры ориентированы на указанные программы, и для освоения всего комплекса процесса применения факторного анализа целесообразно одновременно изучать особенности работы данных модулей.

Анализ литературных источников показал, что влияющие факторы являются факторами «дерева систем» технической эксплуатации автомобилей [8]. Оценка их влияния на параметры технико-эксплуатационных свойств автомобилей и, в частности, городских автобусов может сводиться к нескольким вариантам.

Суть *первого варианта* заключается в следующем (рис. 6.1). Имеется функция отклика (параметр технико-эксплуатационных свойств), например маршрутная норма расхода топлива городского автобуса. Она определяется рядом факторов сложности маршрута движения. Это влияние, как правило, отражается многофакторной моделью следующего вида:

$$H_m = f(X_1, X_2, \dots, X_i, \dots, X_n), \quad (6.1)$$

где  $H_m$  — маршрутная норма расхода топлива автобусов, л/100 км;  $X_i$  ( $i = 1, 2, \dots, n$ ) — факторы (исходные признаки), влияющие на величину маршрутной нормы расхода топлива автобусов.

Исходная математическая модель (6.1) для городских автобусов, построенная с использованием корреляционно-регрессионного анализа, должна быть скорректирована для независимых переменных; при этом может быть получена регрессия на главных компонентах:

$$H_m = f(V_1, V_2, \dots, V_i, \dots, V_n), \quad (6.2)$$

где  $V_i$  ( $i = 1, 2, \dots, n$ ) — главные компоненты.

Поскольку главные компоненты являются линейными комбинациями влияющих факторов (6.3), от регрессии на главных компонентах можно перейти к регрессии на исходных признаках (факторах), т. е. к выражению (6.1), но уже в их независимом сочетании (6.4) [3, 9]:

$$V_i = f(X_1, X_2, \dots, X_i, \dots, X_n). \quad (6.3)$$

При этом число главных компонент равно числу исходных признаков (факторов). Это *второй вариант* оценки и выбора влияющих факторов (его принципиальная схема представлена на рис. 6.2):

$$\begin{aligned} V_1 &= f(X_1, X_2, \dots, X_i, \dots, X_n); \\ V_2 &= f(X_1, X_2, \dots, X_i, \dots, X_n); \\ &\dots \\ V_i &= f(X_1, X_2, \dots, X_i, \dots, X_n); \\ &\dots \\ V_n &= f(X_1, X_2, \dots, X_i, \dots, X_n). \end{aligned} \quad (6.4)$$

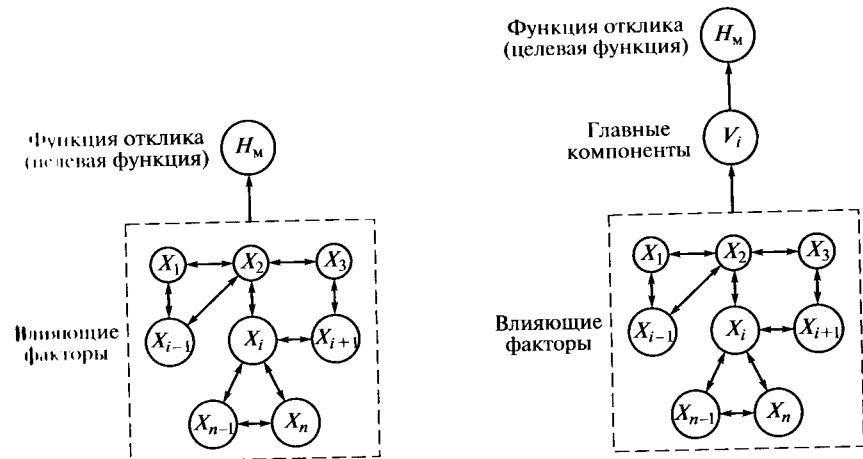


Рис. 6.1. Принципиальная схема взаимодействия влияющих на расход топлива факторов «дерева систем» ТЭА [3] при регрессионном анализе (вариант 1)

Рис. 6.2. Принципиальная схема взаимодействия влияющих на расход топлива факторов «дерева систем» ТЭА [3] при компонентном анализе (без вращения компонент — вариант 2)

Если для оценки и выбора влияющих факторов применить процедуру вращения главных компонент (более подробно изложенную выше), то можно получить еще один вариант решения поставленной проблемы (третий). Принципиальная схема *третьего варианта* представлена на рис. 6.3.

Суть его сводится к следующему. На первом этапе проводится корреляционный анализ и по его результатам формируется совокупность взаимосвязанных (коррелируемых между собой) факторов.

Далее проводится компонентный анализ, на основании которого определяются система главных компонент и осуществляется процедура отбора наиболее значимых главных компонент и осуществляется переход от главных компонент к обобщенным факторам [3, 4, 5, 9]. Это достигается путем вращения наиболее значимых главных компонент (как правило, в целях достижения максимального значения дисперсии для стандартизованных по нормальному закону влияющих факторов). При этом суммарная дисперсия отобранных наиболее значимых главных компонент ( $D_V$ ) равна дисперсии обобщенных факторов ( $D_V'$ ):

$$D_V = D_V'. \quad (6.5)$$

Затем на обобщенных факторах строится регрессия вида

$$H_m = f(V'_1, V'_2, \dots, V'_i, \dots, V'_n), \quad (6.6)$$

от которой осуществляется переход к исходным признакам (выражение (6.1)), поскольку обобщенные факторы также являются линейными комбинациями исходных признаков [3, 4, 5, 9]:

$$V'_i = f(X_1, X_2, \dots, X_i, \dots, X_n). \quad (6.7)$$

При решении ряда задач в области технической эксплуатации автомобилей (ТЭА) (классификация условий эксплуатации, оценка одновременного влияния на показатели ТЭА или автомобиля «возраста» подвижного состава, измеряемого разными единицами измерения (тыс. км пробега, годами), технико-экономический анализ производственной деятельности АТП и др.) применяется простой компонентный анализ, имеющий две разновидности (варианты 4 и 5).

Блок-схема реализации *четвертого варианта* представлена на рис. 6.4. Суть его сводится к следующему. Здесь также на первом этапе проводится корреляционный анализ и по его результатам формируется совокупность взаимосвязанных факторов. Далее проводится компонентный анализ и производится определение главных компонент, а затем выполняется процедура отбора наиболее значимых главных компонент. Выбранные главные компоненты или на стандартизо-

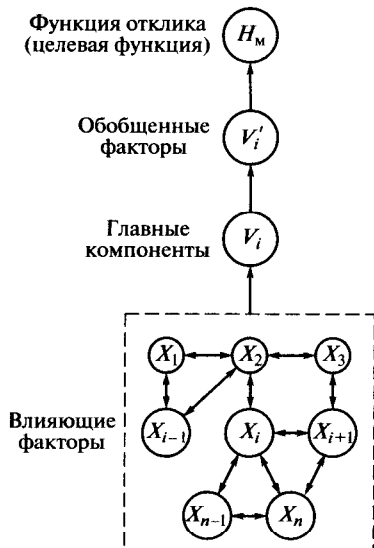


Рис. 6.3. Принципиальная схема взаимодействия влияющих на расход топлива факторов «дерева систем» ТЭА [3] при компонентном анализе (с вращением компонент — вариант 3)

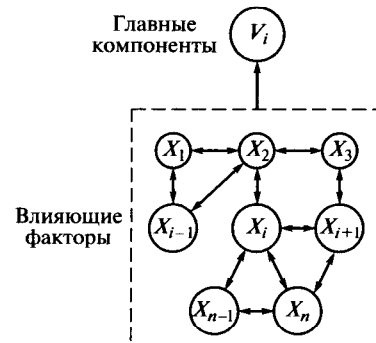
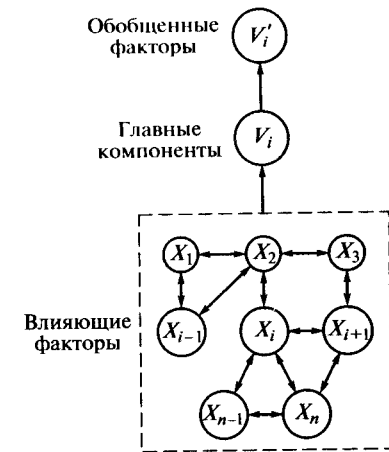


Рис. 6.4. Принципиальная схема взаимодействия влияющих факторов «дерева систем» ТЭА [3] при простом компонентном анализе (без вращение компонент — вариант 4)

Рис. 6.5. Принципиальная схема взаимодействия влияющих факторов «дерева систем» ТЭА [3] при простом компонентном анализе (с вращением компонент — вариант 5)



ванных переменных (выражение (6.8)) или на исходных признаках (выражение (6.3)) используются для дальнейшего анализа:

$$V_i = f(Z_1, Z_2, \dots, Z_i, \dots, Z_n),$$

где  $Z_i$  — стандартизованные значения влияющих факторов.

Стандартизованные значения влияющих факторов «дерева систем» ТЭА [3] определяются по формуле

$$Z_i = \frac{X_i - \bar{X}_i}{\sigma}, \quad (6.8)$$

где  $X_i$  — текущее значение  $i$ -го влияющего фактора;  $\bar{X}_i$  — среднее значение  $i$ -го влияющего фактора;  $\sigma$  — среднеквадратическое отклонение  $i$ -го влияющего фактора.

Если от полученных главных компонент путем их вращения перейти к обобщенным факторам, то можно получить упрощенную факторную структуру. Это вторая разновидность компонентного анализа (*пятый вариант*). Блок-схема его реализации представлена на рис. 6.5.

## 6.2. Метод главных компонент

### 6.2.1. Общая характеристика метода главных компонент

Во многих задачах обработки многомерных наблюдений, в частности в задачах классификации, исследователя интересуют, в первую очередь, лишь те признаки, которые обнаруживают наибольшую изменчивость (наибольший разброс) при переходе от одного объекта к другому.

Следуя общей оптимизационной постановке задачи снижения размерности и полагая анализируемый признак  $X$   $p$ -мерной случайной величиной с вектором средних значений  $a = (a^{(1)}, \dots, a^{(p)})$  и ковариационной матрицей  $\Sigma = (\sigma_{ji})$  ( $i, j = 1, 2, \dots, p$ ) неизвестными, определим в качестве класса  $V(X)$  допустимых преобразований исследуемых признаков  $x^{(1)}, x^{(2)}, \dots, x^{(p)}$  их всевозможные линейные ортогональные нормированные комбинации, т. е.

$$V = \left\{ Z : z^{(j)} = \sum_{v=1}^p c_{jv} (x^{(v)} - a^{(v)}), j=1, 2, \dots, p \right\}, \quad (6.9)$$

где

$$\sum_{v=1}^p c_{jv}^2 = 1 \text{ и } \sum_{v=1}^p c_{jv} c_{kv} = 0, \quad (6.10)$$

для  $j = 1, 2, \dots, p$  и  $k = 1, 2, \dots, p$ , но  $j \neq k$ , а в качестве критерия (меры) информативности  $p'$ -мерной системы показателей

$$Z(X) = (z^{(1)}(X), z^{(2)}(X), \dots, z^{(p')}(X))$$

применяется выражение

$$I_{p'}(Z(X)) = \frac{\text{var } z^{(1)} + \dots + \text{var } z^{(p')}}{\text{var } x^{(1)} + \dots + \text{var } x^{(p)}}. \quad (6.11)$$

Тогда при любом фиксированном  $p' = 1, 2, \dots, p$  вектор искомых вспомогательных переменных  $\tilde{Z}(X) = (\tilde{z}^{(1)}(X), \dots, \tilde{z}^{(p')}(X))^T$  определяется как линейная комбинация:

$$\tilde{Z} = XL^T, \quad (6.12)$$

где  $L = \begin{pmatrix} l_{11} & \dots & l_{1p} \\ \dots & \dots & \dots \\ l_{p'} & \dots & l_{p'p} \end{pmatrix}$  — матрица, строки которой удовлетворяют

условию ортогональности, что дает:

$$I_{p'}(\tilde{z}^{(1)}(X), \dots, \tilde{z}^{(p')}(X)) = \max_{Z(X) \in F} I_{p'}(Z(X)). \quad (6.13)$$

Полученные таким образом переменные  $\tilde{z}^{(1)}(X), \dots, \tilde{z}^{(p')}(X)$  называются главными компонентами вектора  $X$ .

**Замечание 1 (переход к центрированным переменным).** Поскольку, как увидим далее, решение задачи (получение матрицы линейного преобразования  $L$ ) зависит только от элементов ковариационной матрицы  $\Sigma$ , которые, в свою очередь, не изменяются при замене исходных переменных  $x^{(j)}$  переменными  $x^{(j)} - c^{(j)}$  ( $c^{(j)}$  —

произвольные постоянные числа), то в дальнейшем будем считать, что исходная система показателей уже центрирована, т. е.  $E x^{(j)} = 0$ ,  $j = 1, 2, \dots, p$ .

В статистической практике этого добиваются, переходя к наблюдениям:

$$\tilde{x}_i^{(j)} = x_i^{(j)} - \bar{x}^{(j)},$$

где  $\bar{x}^{(j)} = \sum_{i=1}^n x_i^{(j)} / n$  (для упрощения обозначений знак «тильда» над

центрированной переменной и над главной компонентой далее ставить не будем).

**Замечание 2 (переход к выборочному варианту).** Поскольку в реальных статистических задачах располагаем лишь оценками  $\hat{a}$  и  $\hat{\Sigma}$  соответственно вектора средних  $a$  и ковариационной матрицы  $\Sigma$ , то во всех дальнейших рассуждениях под  $\hat{a}^{(j)}$  понимается  $\bar{x}^{(j)}$ , а под  $\hat{\sigma}_{kj}$  — выборочная ковариация  $\hat{\sigma}_{kj} = \sum_{i=1}^n (x_i^{(k)} - \bar{x}^{(k)})(x_i^{(j)} - \bar{x}^{(j)}) / n$  ( $j, k = 1, 2, \dots, p$ ).

**Замечание 3.** Использование главных компонент оказывается наиболее естественным и плодотворным в ситуациях, в которых все компоненты  $x^{(1)}, x^{(2)}, \dots, x^{(p)}$  исследуемого вектора  $X$  имеют общую физическую природу и соответственно измерены в одних и тех же единицах. К таким примерам можно отнести исследование структуры бюджета времени индивидуумов (все  $x^{(i)}$  измеряются в единицах времени), исследование структуры потребления семей (все  $x^{(i)}$  измеряются в денежных единицах), исследование общего развития и умственных способностей индивидуумов с помощью специальных тестов (все  $x^{(i)}$  измеряются в баллах) и т. д.

Если же признаки  $x^{(1)}, x^{(2)}, \dots, x^{(p)}$  измеряются в различных единицах, то результаты исследования с помощью главных компонент будут существенно зависеть от выбора масштаба и природы единиц измерения. Поэтому в подобных ситуациях исследователь предварительно переходит к вспомогательным безразмерным признакам  $x^{*(i)}$ , например с помощью нормирующего преобразования:

$$x_v^{*(i)} = \frac{x_v^{(i)}}{\sqrt{\hat{\sigma}_{ii}}} \quad \left( \begin{matrix} i=1, 2, \dots, p \\ v=1, 2, \dots, n \end{matrix} \right) \quad (6.14)$$

(где  $\hat{\sigma}_{ii}$  соответствует ранее введенным обозначениям), а затем строит главные компоненты относительно этих вспомогательных признаков  $X^*$  и их ковариационной матрицы  $\hat{\Sigma}_{X^*}$ , которая, как легко видеть, является одновременно выборочной корреляционной матрицей  $\hat{R}$  исходных наблюдений  $X_i$ .

### 6.2.2. Вычисление главных компонент

Из определения главных компонент следует, что для вычисления первой главной компоненты необходимо решить оптимизационную задачу вида [5]

$$\begin{aligned} \text{var}(l_1 X) \rightarrow \max_{l_1}; \\ l_1 l_1^T = 1, \end{aligned} \quad (6.15)$$

где  $l_1$  — первая строка матрицы  $L$  (6.12). Учитывая центрированность переменной  $X$  (т.е.  $EX = 0$ ), а также то, что  $E(XX^T) = \Sigma$ , имеем:

$$\text{var}(l_1 X) = E(l_1 X)^2 = E(l_1 X X^T l_1^T) = l_1 \Sigma l_1^T.$$

Следовательно, задача оптимизации может быть записана следующим образом:

$$\begin{aligned} l_1 \Sigma l_1^T \rightarrow \max_{l_1}; \\ l_1 l_1^T = 1. \end{aligned} \quad (6.16)$$

Вводя функцию Лагранжа  $\varphi(l_1, \lambda) = l_1 \Sigma l_1^T - \lambda(l_1 l_1^T - 1)$  и дифференцируя ее по компонентам вектор-столбца  $l_1^T$ , имеем:

$$\frac{\partial \varphi}{\partial l_1^T} = 2 \Sigma l_1^T - 2 \lambda l_1^T,$$

что дает систему для определения  $l_1$ :

$$(\Sigma - \lambda I) l_1^T = 0, \quad (6.17)$$

где  $0 = (0, 0, \dots, 0)^T$  —  $p$ -мерный вектор-столбец из нулей.

Для того чтобы существовало ненулевое решение системы (6.17) (а оно должно быть ненулевым, так как  $l_1 l_1^T = 1$ ), матрица  $|\Sigma - \lambda I|$  должна быть вырожденной, т.е.

$$|\Sigma - \lambda I| = 0.$$

Этого добиваются подбором соответствующего значения  $\lambda$ . Уравнение (6.17) (относительно  $\lambda$ ) называется характеристическим для матрицы  $\Sigma$ . Известно, что при симметричности и неотрицательной определенности матрицы  $\Sigma$  (каковой она и является, как всякая ковариационная матрица) это уравнение имеет  $p$  вещественных неотрицательных корней  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ , называемых характеристическими (или собственными) значениями матрицы  $\Sigma$ .

Учитывая, что  $\text{var} \tilde{z}^{(1)} = \text{var}(l_1 X) = l_1 \Sigma l_1^T$  и  $l_1 \Sigma l_1^T = \lambda$  (последнее соотношение следует из соотношения (6.3) после его умножения слева на  $l_1$  с учетом  $l_1 l_1^T = 1$ ), получаем:

$$\text{var} \tilde{z}^{(1)}(X) = \lambda.$$

Поэтому для обеспечения максимальной величины дисперсии переменной  $\tilde{z}^{(1)}$  нужно выбрать из  $p$  собственных значений матрицы  $\Sigma$  наибольшее, т.е.

$$\text{var} \tilde{z}^{(1)}(X) = \lambda_1. \quad (6.18)$$

Подставляя  $\lambda_1$  в систему уравнений (6.17) и решая ее относительно  $l_{11}, \dots, l_{1p}$ , определяем компоненты вектора  $l_1$ . Таким образом, первая главная компонента получается как линейная комбинация

$$\tilde{z}^{(1)}(X) = l_1 X,$$

где  $l_1$  — собственный вектор матрицы  $\Sigma$ , соответствующий наибольшему собственному числу этой матрицы.

Далее аналогично можно показать, что  $\tilde{z}^{(k)}(X) = l_k X$ , где  $l_k$  — собственный вектор матрицы  $\Sigma$ , соответствующий  $k$ -му по величине собственному значению  $\lambda_k$  этой матрицы.

Таким образом, соотношения для определения всех  $p$  главных компонент вектора  $X$  могут быть представлены в виде (6.13), где  $Z = (\tilde{z}^{(1)}, \dots, \tilde{z}^{(p)})^T$ ,  $X = (x^{(1)}, \dots, x^{(p)})^T$ , а матрица  $L$  состоит из строк  $l_j = (l_{j1}, \dots, l_{jp})$ ,  $j = \overline{1, p}$ , являющихся собственными векторами матрицы  $\Sigma$ , соответствующих собственным числам  $\lambda_j$ . При этом сама матрица  $L$  в соответствии с условиями (6.13) является ортогональной, т.е.

$$LL^T = L^T L = 1. \quad (6.19)$$

В дальнейшем в целях упрощения обозначений мы будем опускать знак «тильда» над переменными главных компонент, т.е. обозначать главные компоненты просто  $Z = (z^{(1)}, z^{(2)}, \dots, z^{(p)})$ .

### 6.2.3. Основные числовые характеристики главных компонент

Определим основные числовые характеристики (средние значения, дисперсии, ковариации) главных компонент в терминах основных числовых характеристик исходных переменных и собственных значений матрицы  $\Sigma$ :

$$\text{а) } EZ = E(LX) = LEX = 0;$$

б) ковариационная матрица вектора главных компонент:

$$\Sigma_Z = E(ZZ^T) = E((LX)(LX)^T) = E(LXX^T L^T) = LE(XX^T)L^T = L\Sigma L^T.$$

Умножая слева соотношение  $(\Sigma - \lambda_k I) l_k^T = 0$  ( $k = \overline{1, p}$ ) на  $l_j$  ( $j = \overline{1, p}$ ), получаем:

$$L\Sigma L^T = \Sigma_Z = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \dots \\ & & & \lambda_p \end{pmatrix}. \quad (6.20)$$

Из выражения (6.20), в частности, следует подтверждение взаимной некоррелированности главных компонент, а также  $\text{var } z^{(k)} = \lambda_k$  ( $k = \overline{1, p}$ );

в) сумма дисперсий исходных признаков равна сумме дисперсий всех главных компонент. Действительно,

$$\begin{aligned} \sum_{k=1}^p \text{var } z^{(k)} &= \text{tr } \Sigma_Z = \text{tr}(L\Sigma L^T) = \text{tr}((L\Sigma)L^T) = \\ &= \text{tr}(L^T(L\Sigma)) = \text{tr}((L^T L)\Sigma) = \text{tr } \Sigma = \sum_{k=1}^p \text{var } x^{(k)}. \end{aligned} \quad (6.21)$$

**Следствие 1.** Из «б» и «в», в частности, следует, что критерий информативности метода главных компонент (6.21) может быть представлен в виде

$$I_{p'}(Z(X)) = \frac{\lambda_1 + \dots + \lambda_{p'}}{\lambda_1 + \dots + \lambda_p}, \quad (6.22)$$

где  $\lambda_1, \lambda_2, \dots, \lambda_p$  — собственные числа ковариационной матрицы  $\Sigma$  вектора  $X$ , расположенные в порядке убывания.

Представление  $I_{p'}(Z(X))$  в виде (6.22) дает исследователю некоторую основу, опорную точку зрения, при вынесении решения о том, сколько последних главных компонент можно без особого ущерба изъять из рассмотрения, сократив тем самым размерность исследуемого пространства.

Действительно, анализируя с помощью выражения (6.22) изменение относительно доли дисперсии, вносимой первыми  $p'$  главными компонентами, в зависимости от числа этих компонент, можно определить число компонент, которое целесообразно оставить в рассмотрении.

**Следствие 2.** Если  $X^*$  — вектор нормированных признаков  $x^{*(1)}, \dots, x^{*(p)}$ , т.е.  $\text{E}x^{*(j)} = 0$  и  $\text{var } x^{*(j)} = 1$  для  $j = \overline{1, p}$ , то согласно замечанию 3 ковариационная и корреляционная матрицы совпадают (т.е.  $\Sigma_{X^*} = R$ ) и из «б» и «в» следует:

$$\text{tr } \Sigma_Z = \text{tr } \Sigma_{X^*} = \text{tr } R = p,$$

или

$$\sum_{k=1}^p \lambda_k = p.$$

Тогда критерий информативности может быть представлен в виде

$$I_{p'}(Z(X)) = \frac{\lambda_1 + \dots + \lambda_{p'}}{p}. \quad (6.23)$$

Следует иметь в виду, что свойства оценок коэффициентов регрессии зависят от свойств случайного члена в регрессионной модели.

#### 6.2.4. Выбор главных компонент и переход к обобщенным факторам

Существует несколько часто употребляемых критериев определения числа главных компонент. Некоторые из них являются альтернативными по отношению к другим, а часть этих критериев можно использовать вместе, чтобы один дополнял другой. В области управления технической эксплуатацией автомобилей наиболее часто применяются:

- различные правила, формулируемые в терминах собственных чисел;
- критерий, основанный на величине долей дисперсий факторов;
- критерий отсеивания;
- критерий интерпретируемости и инвариантности.

**I. Критерии, основанные на собственных числах.** Выбор главных компонент осуществляется по величине собственного числа компоненты:

$$\lambda \geq 1. \quad (6.24)$$

Значимой считается любая главная компонента, имеющая собственное число больше единицы (критерий Кайзера). При этом используется корреляционная (нередуцированная) матрица. Этот простой критерий хорошо себя зарекомендовал, так как обычно дает результаты, совпадающие с теми, которые ожидает получить исследователь [9]. Указанный критерий был применен при нормировании расхода топлива городскими автобусами.

Для корреляционной матрицы, относящейся к генеральной совокупности, рассматриваемый критерий всегда дает нижнюю оценку числа главных компонент.

**II. Критерий, основанный на доле воспроизводимой дисперсии.** Критерии значимости оперируют с выборочной изменчивостью данных. Критерии, основанные на собственных числах, формулируются в терминах абстрактных характеристик матрицы.

Возможен другой подход — для каждой главной компоненты оценивается доля дисперсии, воспроизводимая ею. Данный критерий

становится особенно наглядным, когда выделение первоначальных главных компонент производится с помощью нередуцированной корреляционной матрицы. Тогда в качестве статистики этого критерия выступает доля дисперсии, воспроизводимой последней главной компонентой по отношению к полной дисперсии, равной числу параметров. Критерий определяется уровнем для минимальной доли воспроизводимой дисперсии [9].

Основной недостаток критерия, основанного на величине доли воспроизводимой дисперсии, состоит в определенной его субъективности. Однако он основан на легко поддающейся интерпретации статистике, и в этом преимущество данного метода [9].

**III. Критерий отсеивания.** Рассматривается графическое изображение (рис. 6.6) собственных чисел корреляционной матрицы, которые наносятся на график в порядке их убывания. Выделение заканчивается на том факторе (точке), после которого исследуемая зависимость близка к почти горизонтальной прямой линии. Эту прямую Кэттель и предложил использовать для выделения факторов [9]. Данный критерий принято называть критерием каменистой осыпи.

Предполагается, что справа от этой точки находится только факториальная осыпь. «Осыпь» является геологическим термином, обозначающим обломки горных пород, скапливающиеся в нижней части скалистого склона. График осыпи всегда показывается модулем «Анализ фактора» программы STATISTICA.

**IV. Критерий интерпретируемости и инвариантности.** Для исключения сомнительных результатов можно применить к одним и тем же данным комбинацию различных независимых критериев и принимать только те результаты, которые подходят ко всем критериям. Окончательное решение должно базироваться на его приемлемости с точки зрения научных представлений в данной области [9]. Для оптимизации полученных главных компонент применяется проце-

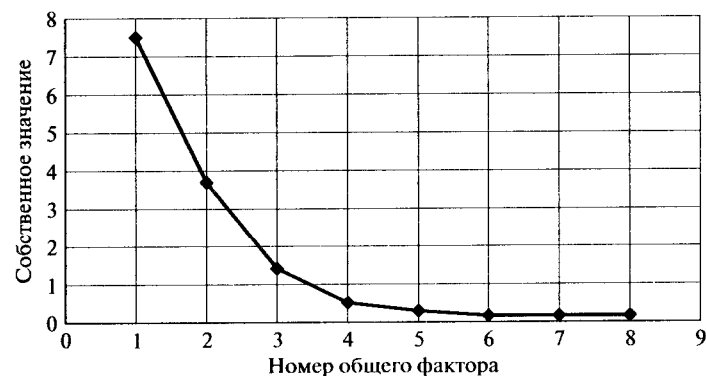


Рис. 6.6. График «каменистой осыпи»

ду их вращения [5]. При этом при вращении используется принцип простой структуры.

В математической статистике принято, что главные компоненты после вращения называются обобщенными факторами, а простая структура должна удовлетворять следующим требованиям [5]:

- каждая строка матрицы факторной структуры должна содержать хотя бы один нулевой элемент;
- в каждом столбце матрицы факторной структуры должно быть не менее  $m$  нулей (так как  $m$  — число обобщенных факторов, то каждый фактор в своей гиперплоскости координат должен определиться не менее чем  $m$  признаками);
- для каждой пары столбцов матрицы факторной структуры найдется несколько признаков, соответствующие элементы которых в матрице равны нулю в одном столбце и не равны нулю в другом столбце;
- если число факторов равно четырем или больше, то достаточно велика доля признаков, имеющих в любой паре столбцов одновременно нулевые элементы;
- для любой пары столбцов найдется мало параметров, соответствующие элементы которых в обоих столбцах отличны от нуля.

Существует три подхода к проблеме вращения. *Первый подход* — графический [9]. Вращение заключается в проведении новых осей, которые соответствуют некоторому критерию простой и легко интерпретируемой структуры. Простая структура получается в том случае, если в факторном пространстве есть явные скопления переменных (результатов наблюдений). Они легко различимы, т. е. легко отделяемы друг от друга, и через них можно провести координатные оси (рис. 6.7). Если такое разделение не обеспечивается или число факторов велико, то графический метод практически не применим.

*Второй подход* [9] связан с аналитическими методами. В этом случае выбирается некоторый объективный критерий, которым надо руководствоваться при выполнении вращения. В рамках этого подхода различают два вида вращения: ортогональное и косоугольное (они, в свою очередь, имеют многочисленные вариации).

*Третий подход* [9] заключается в задании априорной целевой матрицы. Например, при вращении по часовой стрелке пользуются матрицей  $T$ :

$$T = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}, \quad (6.25)$$

где  $\alpha$  — угол поворота, °.

На практике обычно производится одновременный поворот всех осей координат по стандартным программам при помощи матрицы  $T$  [5].

Как показывает опыт, при решении задач технической эксплуатации автомобильного транспорта обычно обходятся одной или дву-

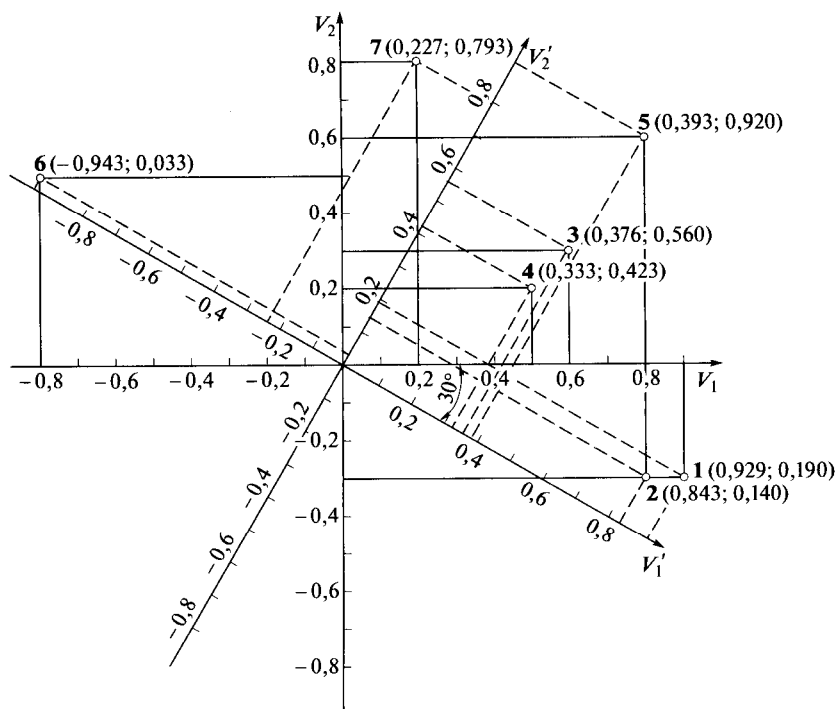


Рис. 6.7. Графическая интерпретация преобразования главных компонент в обобщенные факторы:

$V_1$  — первая главная компонента;  $V_2$  — вторая главная компонента;  $V_1'$  — первый обобщенный фактор;  $V_2'$  — второй обобщенный фактор

мя главными компонентами. Поэтому рассмотрим пример поворота главных компонент в двухмерном пространстве.

Пусть дана матрица  $A$  первых двух главных компонент:

$$A = \begin{pmatrix} 0,9 & -0,3 \\ 0,8 & -0,3 \\ 0,6 & +0,3 \\ 0,5 & +0,2 \\ 0,8 & +0,6 \\ -0,8 & +0,5 \\ 0,2 & +0,8 \end{pmatrix}$$

Необходимо преобразовать данную матрицу в матрицу, состоящую из обобщенных факторов, путем поворота ее на  $30^\circ$  по часовой стрелке. Результаты представить аналитически.

Решение сводится к следующему. Определяется матрица  $T$ :

$$T = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} = \begin{pmatrix} 0,866 & 0,50 \\ -0,50 & 0,866 \end{pmatrix}$$

Затем находится матрица  $A'$ :

$$A' = AT = \begin{pmatrix} 0,9 & -0,3 \\ 0,8 & -0,3 \\ 0,6 & +0,3 \\ 0,5 & +0,2 \\ 0,8 & +0,6 \\ -0,8 & +0,5 \\ 0,2 & +0,8 \end{pmatrix} \begin{pmatrix} 0,866 & -0,5 \\ -0,5 & 0,866 \end{pmatrix} = \begin{pmatrix} 0,929 & +0,190 \\ 0,843 & +0,140 \\ 0,367 & +0,560 \\ 0,333 & +0,423 \\ 0,393 & +0,920 \\ -0,943 & +0,033 \\ 0,227 & +0,793 \end{pmatrix}$$

Исходные признаки представим в координатах первых двух главных компонент ( $V_1$  и  $V_2$ ) и двух обобщенных факторах ( $V_1'$  и  $V_2'$ ) (см. рис. 6.7).

При решении задач технической эксплуатации автомобилей получили распространение методы ортогонального и косоугольного вращения, а также методы, основанные на введении вторичных осей. Данные методы легко реализуются (двумя-тремя щелчками «мышь») соответствующими опциями модуля «Анализ фактора» программы STATISTICA.

**Методы ортогонального вращения.** Поскольку каждая переменная имеет нагрузку только на один фактор, интерпретация переменных не представляет особого труда. Но для численного использования эта характеристика степени сложности не совсем удобна. Одной из возможных мер сложности модели является вариация квадрата факторной нагрузки для каждой строки (для каждой переменной). Принимается в рассмотрение квадрат нагрузок только для того, чтобы избежать осложнений, связанных с учетом знака.

Известно, что дисперсия определяется как математическое ожидание квадрата отклонений от среднего, поэтому при фиксированном числе факторов и заданных общностях дисперсия максимальна, если одно из значений квадратов нагрузок равно общности, а все остальные элементы в строке нулевые. Иначе говоря, дисперсия квадратов факторных нагрузок переменной есть мера факторной сложности этой переменной [9]:

$$\left( \begin{array}{l} \text{Факторная} \\ \text{сложность} \end{array} \right) = \frac{1}{r} \sum_{j=1}^r (b_{ij}^2 - \bar{b}_{ij}^2)^2, \quad (6.26)$$

где  $r$  — число столбцов факторной матрицы;  $b_{ij}$  — факторная нагрузка  $j$ -го фактора на  $i$ -ю переменную;  $\bar{b}_{ij}$  — среднее значение квадратов факторных нагрузок в  $i$ -й строке.

Это соотношение можно представить в другом виде:

$$q_i = \frac{\sum_{j=1}^r (b_{ij}^4) - \left( \sum_{j=1}^r (b_{ij}^2) \right)^2}{r^2}. \quad (6.27)$$

Общей мерой сложности может служить сумма  $q_i$  всех переменных:

$$q = \sum_{i=1}^n q_i = \sum_{i=1}^n \frac{\sum_{j=1}^r (b_{ij}^4) - \left( \sum_{j=1}^r (b_{ij}^2) \right)^2}{r^2}. \quad (6.28)$$

Использование критерия *квартимакс* основано на вращении осей таким образом, чтобы результирующие факторные нагрузки максимизировали  $q$ .

Применяя этот критерий, можно обеспечить простоту интерпретации переменных за счет простоты интерпретации факторов. В частности, описание переменной упрощается при уменьшении числа общих факторов, связанных с ней.

В то же время описание фактора становится проще, если небольшое число переменных имеют существенные нагрузки на этот фактор, а остальные переменные — нулевые нагрузки. Таким образом, метод *квартимакс* имеет тенденцию к выделению генерального фактора [9].

Метод *варимакс* использует немного другой критерий, в котором добиваются упрощения описания столбцов факторной матрицы. Вместо дисперсии квадратов нагрузок переменной рассматривается дисперсия квадратов нагрузок фактора. Общая мера простоты задается критерием *варимакс*:

$$V = \sum_{j=1}^r v_j = \frac{\sum_{j=1}^r n \sum_{i=1}^n b_{ij}^4 - \sum_{j=1}^r \left( \sum_{i=1}^n b_{ij}^2 \right)^2}{n^2}. \quad (6.29)$$

Обычно нормированные факторные нагрузки применяют, чтобы избавиться от нежелательного влияния на результат вращения переменных с большой общностью. Хотя алгоритмически метод *квартимакс* проще, чем метод *варимакс*, последний обычно дает лучшее разделение факторов [9].

Так как критерий *квартимакс* основан на упрощении описания строк, а критерий *варимакс* — на упрощении описания столбцов, то существует некоторый совместный критерий, имеющий следующий вид:

$$\alpha Q + \beta V = \max. \quad (6.30)$$

Учитывая выражения (6.28) и (6.29), можно записать:

$$\sum_{j=1}^r \sum_{i=1}^n b_{ij}^4 - \gamma \sum_{j=1}^r \left( \sum_{i=1}^n b_{ij}^2 \right)^2 / n = \max, \quad (6.31)$$

где  $\gamma = \beta / (\alpha + \beta)$ .

Таким образом, если  $\gamma = 0$ , то образуется критерий *квартимакс*, а если  $\gamma = 1$ , то критерий *варимакс*.

При  $\gamma = r/2$  и  $\gamma = 0,5$  получаются особые критерии, названные критериями *эквимакс* и *биквартимакс* соответственно [9].

**Методы косоугольного вращения.** Косоугольное вращение является более общим, чем ортогональное, так как здесь нет ограничений, связанных с некоррелированностью факторов. Преимущество косоугольного вращения состоит в следующем: когда в результате его выполнения получаются ортогональные факторы, можно быть уверенным, что эта ортогональность действительно им свойственна, а не привнесена методом вращения. Поскольку косоугольные вращения производятся с учетом корреляций между факторами, существуют многочисленные методы интерпретации результатов такого анализа.

Для объяснения корреляции между факторами в ряде случаев вводят факторы второго и более высокого порядков. Кроме того, существуют два подхода к косоугольному вращению: использование вторичных осей и первичной матрицы факторного отображения [9].

Критерий *облимакс* основан на упрощении факторной структуры по принципу увеличения числа значительных и пренебрежимо малых нагрузок за счет остальных коэффициентов структуры. Этот критерий эквивалентен критерию *квартимакс* в случае ортогонального вращения, но приводит к решению, отличному от метода *квартимин* при использовании его без ограничения, связанного с ортогональностью [9].

Следует отметить еще два метода вращения: метод *ортоблик* и метод *максплейн* [9].

**Методы, основанные на введении вторичных осей.** Эти методы основаны на том, что если существуют разделимые скопления точек, определяемые первичными факторами, то они будут иметь почти нулевые проекции на все вторичные оси, за исключением одной. Таким образом, можно определить критерий, называемый *квартимин*, который аналогичен критерию *квартимакс*:

$$N = \sum_{i=1}^n \sum_{j=1}^r a_{ij} a_{ik}. \quad (6.32)$$

В данном случае цель вращения — нахождение таких факторных нагрузок, которые минимизируют величину критерия  $N$ .



По аналогии с ортогональным критерием *варимакс* вводится критерий *коваримин*. В этом случае минимизируется ковариация квадратов проекций на вторичные оси [9]:

$$C = \sum_{j < k=1}^r \left( n \sum_{i=1}^n a_{ij}^2 a_{ik}^2 - \sum_{i=1}^n a_{ij}^2 \sum_{i=1}^n a_{ik}^2 \right). \quad (6.33)$$

Применительно к одним и тем же данным критерий *коваримин*, как правило, дает меньше косоугольных факторов, чем *квартимин*. Объединение этих двух критериев приводит к обобщенному критерию *облимину* [9]:

$$B = \alpha N + \beta C / n = \min. \quad (6.34)$$

Подставив в это выражение формулы (6.32), (6.33), получим:

$$B = \sum_{j < k=1}^r \left( n \sum_{i=1}^n a_{ij}^2 a_{ik}^2 - \gamma \sum_{i=1}^n a_{ij}^2 \sum_{i=1}^n a_{ik}^2 \right), \quad (6.35)$$

где  $\gamma = \beta / (\alpha + \beta)$ .

Этот критерий при  $\gamma = 0$  переходит в *квартимин*; при  $\gamma = 0,5$  — в *биквартимин*; при  $\gamma = 1$  — в *коваримин*.

### 6.3. Примеры использования компонентного анализа при решении задач управления процессами технической эксплуатации автомобилей

**Пример 1.** Исследованиями установлено, что условия эксплуатации и тип маршрута существенно влияют на целый ряд технико-эксплуатационных показателей работы городских автобусов (в частности, на расход топлива, ресурс шин, потери линейного времени, напряженность труда водителей, выбросы вредных веществ в отработавших газах и т.д.). Следовательно, могут объективно формироваться параметры сложности маршрута движения.

Среди них наибольшее влияние приходится на условия движения (50, 33, 50, 33 и 33 % соответственно), транспортные условия (33, 17, 33, 50 и 50 % соответственно) и дорожные условия (17, 50, 17, 17 и 17 % соответственно).

Среди факторов условий движения наиболее значимыми являются частота плановых и внеплановых остановок, длина перегона технологического цикла и скорость движения автобуса на перегоне. Среди факторов транспортных условий наиболее значимыми являются наполняемость салона автобуса по перегонам, интенсивность движения транспортного потока и вид перекрестка. Среди факторов дорожных условий наиболее значимыми являются состояние и тип дорожного покрытия.

Корреляционный анализ указанных факторов и физическая сущность процесса позволили выявить наиболее значимые из них, но они оказались взаимосвязанными (табл. 6.1).

Таблица 6.1

Коэффициенты парной корреляции факторов сложности маршрута движения городского автобуса

Факторы	$\gamma$	$l_n$	$v_s$	$\rho$
Коэффициент использования пассажироместимости $\gamma$	1,0	-0,79	-0,75	0,77
Средняя длина перегона $l_n$	-0,79	1,00	0,96	-0,74
Средняя эксплуатационная скорость $v_s$	-0,75	0,96	1,00	-0,72
Плотность транспортного потока $\rho$	0,77	-0,74	-0,72	1,00

Поэтому при помощи модуля «Анализ фактора» программы STATISTICA был выполнен компонентный анализ указанных факторов и получена первая главная компонента:

$$V = 0,867 Z_\gamma - 0,937 Z_{l_n} - 0,946 Z_{v_s} + 0,868 Z_\rho, \quad (6.36)$$

где  $Z_\gamma, Z_{l_n}, Z_{v_s}, Z_\rho$  — стандартизованные значения факторов сложности маршрута движения городских автобусов: коэффициента использования пассажироместимости, средней длины перегона, средней эксплуатационной скорости автобуса на маршруте и средней плотности транспортного потока на маршруте соответственно (переход к новым значениям выполнен на основе стандартного преобразования, приведенного в подразд. 6.1).

Также было определено собственное число наиболее значимой первой главной компоненты ( $\lambda_1 = 2,67$ ).

Далее в соответствии с процедурой компонентного анализа был осуществлен переход от модели на стандартизованных значениях факторов сложности маршрута движения городских автобусов к математической модели обобщенного параметра сложности маршрута движения городских автобусов на исходных признаках (путем подстановки соответствующих выражений для стандартных преобразований исходных признаков в зависимость (6.36)):

$$P_c = 9,12 + 11,1\gamma - 9,1l_n - 0,64v_s + 1,06\rho, \quad (6.37)$$

где  $P_c$  — обобщенный параметр сложности маршрута движения городских автобусов на исходных признаках.

Анализ математической модели обобщенного параметра сложности маршрута движения городских автобусов показывает, что наибольшее влияние на его формирование оказывает средняя эксплуатаци-

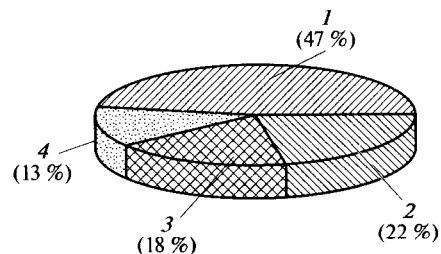


Рис. 6.8. Веса факторов в математической модели обобщенного параметра сложности маршрута движения городских автобусов:

1 — средняя эксплуатационная скорость; 2 — средняя длина перегона; 3 — коэффициент использования пассажироместности; 4 — средняя плотность транспортного потока

онная скорость на маршруте (47%), далее идут средняя длина перегона (22%), коэффициент использования пассажироместности (18%) и средняя плотность транспортного потока (13%).

Графическая интерпретация вклада факторов сложности в формирование обобщенного параметра сложности маршрута движения городских автобусов представлена на рис. 6.8. Полученный показатель сложности автобусного маршрута является комплексным. Следовательно, данный подход позволяет более точно классифицировать автобусные маршруты по сложности. Таким образом, был реализован четвертый вариант принципиальной схемы взаимодействия влияющих факторов «дерева систем» ТЭА при компонентном анализе (см. подразд. 6.1).

Пятый вариант принципиальной схемы взаимодействия влияющих факторов «дерева систем» ТЭА при компонентном анализе (см. подразд. 6.1) был реализован при построении обновленного параметра сложности маршрута движения городских автобусов. При этом на первом этапе были собраны статистические данные по значимым факторам сложности маршрута движения городских автобусов.

На втором этапе была проведена их обработка с использованием математического аппарата корреляционного и компонентного анализа. Обработка проводилась с использованием современных программных комплексов STATISTICA 6.0 и SPSS 11.0 for Windows. Особенность обработки заключалась в том, что были построены математические модели обобщенного параметра сложности маршрута движения городских автобусов на главных компонентах без вращения и с вращением компонент. При вращении использовались наиболее распространенные методы [5, 9]: *варимакс*, *эквиماكс*, *квартимакс*, *облимин* и *промакс*. Корреляционный анализ (табл. 6.2) подтвердил взаимосвязь выбранных факторов сложности маршрута движения городского автобуса.

На третьем этапе были проанализированы полученные математические модели обобщенного параметра сложности маршрута движения городских автобусов на главных компонентах без вращения и с вращением компонент следующего вида:

- без вращения главных компонент:

Таблица 6.2

Коэффициенты парной корреляции факторов сложности маршрута движения городского автобуса

Факторы	$v_s$	$l_n$	$n_n$	$\gamma$	$\rho$
Средняя эксплуатационная скорость $v_s$	1,00	0,86	-0,29	-0,75	-0,52
Средняя длина перегона $l_n$	0,86	1,00	-0,33	-0,77	-0,64
Удельное число поворотов $n_n$	-0,29	-0,33	1,00	0,18	0,13
Коэффициент использования пассажироместности $\gamma$	-0,75	-0,77	0,18	1,00	0,7
Плотность транспортного потока $\rho$	-0,52	-0,64	0,13	0,7	1,00

$$V = -0,285Z_{v_s} - 0,207Z_{l_n} + 0,241Z_{n_n} + 0,23Z_{\gamma} + 0,236Z_{\rho}, \quad (6.38)$$

где  $Z_{v_s}$ ,  $Z_{l_n}$ ,  $Z_{n_n}$ ,  $Z_{\gamma}$ ,  $Z_{\rho}$  — стандартизованные значения факторов сложности маршрута движения городских автобусов: средней эксплуатационной скорости автобуса на маршруте, средней длины перегона, удельного числа поворотов на маршруте, коэффициента использования пассажироместности и средней плотности транспортного потока на маршруте соответственно;

- при вращении главных компонент по методу *варимакс*:

$$V = -0,031Z_{v_s} + 0,211Z_{l_n} + 1,241Z_{n_n} - 0,29Z_{\gamma} - 0,167Z_{\rho}; \quad (6.39)$$

- при вращении главных компонент по методу *эквиماكс*:

$$V = -0,57Z_{v_s} + 1,307Z_{l_n} - 0,035Z_{n_n} - 0,041Z_{\gamma} + 0,016Z_{\rho}; \quad (6.40)$$

- при вращении главных компонент по методу *квартимакс*:

$$V = -0,505Z_{v_s} + 0,158Z_{l_n} + 0,688Z_{n_n} - 0,036Z_{\gamma} - 0,055Z_{\rho}; \quad (6.41)$$

- при вращении главных компонент по методу *облимин*:

$$V = -0,964Z_{v_s} - 0,011Z_{l_n} + 0,13Z_{n_n} + 0,11Z_{\gamma} + 0,12Z_{\rho}; \quad (6.42)$$

- при вращении главных компонент по методу *промакс*:

$$V = 1,0Z_{n_n}. \quad (6.43)$$

При вращении главных компонент по методу *промакс* была получена простейшая структура для любого влияющего фактора при достаточной значимости главных компонент ( $\lambda_1 = 3,43$ ;  $\sigma_1^2 = 68,687$ ), но это сводит формирование обобщенного параметра сложности маршрута движения городских автобусов к воздействию одного ка-

кого-то фактора, что, в свою очередь, огрубляет исследуемый процесс и искажает его эмергентные свойства.

При вращении главных компонент по методу *эквимакс* также была получена простейшая структура (формула (6.40)), отражающая фактически значимое влияние двух факторов (средней эксплуатационной скорости и средней длины перегона) при приемлемых характеристиках компонент ( $\lambda_1 = 3,43$ ;  $\sigma_1^2 = 68,687$ ), но искажающая влияние фактора средней длины перегона. При увеличении средней длины перегона значения обобщенного параметра сложности маршрута движения городских автобусов на главных компонентах будут увеличиваться, т.е. чем больше перегон, тем сложнее маршрут. Это противоречит общепринятым критериям.

Аналогичным образом была выведена из рассмотрения фактически трехфакторная модель (формула (6.41)), полученная по методу *квартимакс*.

Пятифакторная модель (формула (6.39)), полученная по методу *варимакс*, была выведена из рассмотрения, поскольку искажает влияние на сложность маршрута движения таких факторов, как средняя длина перегона, коэффициент использования пассажироместимости и плотность транспортного потока.

В ходе анализа математических моделей (формулы (6.38) и (6.42)) было установлено, что наиболее представительной является математическая модель обобщенного параметра сложности маршрута движения городских автобусов на главных компонентах без вращения (формула (6.38)). Она в меньшей степени искажает эмергентные свойства исследуемого процесса, поскольку отражает влияние пяти факторов (как и в реальной действительности), на двух главных компонентах она отображает 82 % общей дисперсии (69 % общей дисперсии описывает первая главная компонента и 13 % общей дисперсии описывает вторая главная компонента).

На четвертом этапе с использованием формулы (6.8) был осуществлен переход от стандартизованных значений факторов сложности маршрута движения городских автобусов к их исходным значениям. В результате была получена математическая модель обобщенного параметра сложности маршрута движения городских автобусов на исходных факторах:

$$P_c = -0,314 - 0,059v_s - 0,506l_n + 0,713n_n + 0,764\gamma + 0,246\rho, \quad (6.44)$$

где  $P_c$  — обобщенный параметр сложности маршрута движения городских автобусов на исходных факторах;  $v_s$  — средняя эксплуатационная скорость, км/ч;  $l_n$  — средняя длина перегона, км;  $n_n$  — удельное число поворотов, шт./км;  $\gamma$  — коэффициент использования пассажироместимости;  $\rho$  — плотность транспортного потока, количество автомобилей на 100 м.

На пятом этапе был проведен анализ математической модели обобщенного параметра сложности маршрута движения городских

автобусов на исходных факторах. В ходе анализа были определены минимальное и максимальное значения обобщенного параметра сложности, его чувствительность и вклад каждого из факторов сложности.

Минимальное и максимальное значения обобщенного параметра сложности маршрута движения городских автобусов определялись путем подстановки экстремальных значений исходных факторов сложности в математическую модель (формула (6.44)). В результате расчетов было установлено, что минимальное значение обобщенного параметра сложности маршрута движения городских автобусов составило  $-3,2$  ( $P_c^{\min} = -3,2$ ), а максимальное его значение составило  $2,85$  ( $P_c^{\max} = 2,85$ ).

Чувствительность обобщенного параметра сложности маршрута движения городских автобусов проверялась по коэффициенту эластичности. Было установлено, что при изменении средней эксплуатационной скорости на 1 % обобщенный параметр сложности маршрута движения городских автобусов изменяется на 0,31 %. При изменении средней длины перегона на 1 % обобщенный параметр сложности маршрута движения городских автобусов изменяется на 0,12 %. При изменении удельного числа поворотов на 1 % обобщенный параметр сложности маршрута движения городских автобусов изменяется на 0,4 %. При изменении коэффициента использования пассажироместимости на 1 % обобщенный параметр сложности маршрута движения городских автобусов изменяется на 0,3 %. При изменении плотности транспортного потока на 1 % обобщенный параметр сложности маршрута движения городских автобусов изменяется на 0,33 %. Таким образом, было установлено, что наибольшей чувствительностью обладает фактор удельного числа поворотов на маршруте движения городского автобуса.

Анализ вклада влияющих факторов сложности маршрута движения показал, что наибольшее влияние на формирование обобщенного параметра сложности маршрута движения городских автобусов оказывает средняя эксплуатационная скорость на маршруте (33 %), далее идут средняя плотность транспортного потока (19 %), удельное число поворотов (16 %), коэффициент использования пассажироместимости (15 %) и средняя длина перегона (8 %).

Графическая интерпретация вклада факторов сложности в формирование обобщенного параметра сложности маршрута движения городских автобусов представлена на рис. 6.9.

Полученный измеритель сложности автобусного маршрута является комплексным параметром, количественно характеризует совокупное влияние условий эксплуатации (дорожных и транспортных условий и условий движения) на эффективность работы городских автобусов на линии.

На его основе была уточнена общепринятая классификация условий эксплуатации подвижного состава автомобильного транспор-

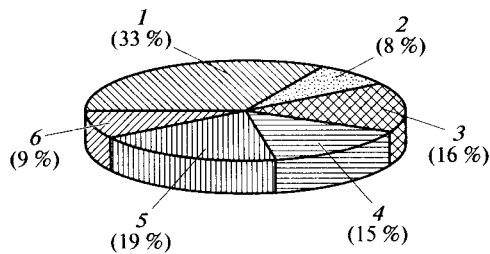


Рис. 6.9. Веса факторов в математической модели обобщенного параметра сложности маршрута движения городских автобусов:

1 — средняя эксплуатационная скорость; 2 — средняя длина перегона; 3 — удельное число поворотов; 4 — коэффициент использования пассажироместности; 5 — средняя плотность транспортного потока; 6 — прочие факторы

та, регламентированная «Положением о ТО и ремонте ПС АТ», и разработаны коэффициенты корректирования нормативов по ТО и ТР применительно к городским автобусам (табл. 6.3).

Оперативное корректирование нормативов технической эксплуатации городских автобусов производится путем изменения количественного значения основных нормативов по ТО и ТР. При этом используются следующие формулы:

$$l_{\text{ТО}} = l_{\text{ТО}}^{\text{н}} K_{1c} K_3; \quad (6.45)$$

$$L_{\text{к.р}} = L_{\text{к.р}}^{\text{н}} K_{1c} K_2 K_3; \quad (6.46)$$

$$t_{\text{т.р}} = t_{\text{т.р}}^{\text{н}} K_{1c} K_2 K_3 K_4 K_5; \quad (6.47)$$

$$H_{3,ч} = H_{3,ч}^{\text{н}} K_{1c} K_2 K_3, \quad (6.48)$$

где  $l_{\text{ТО}}$  — периодичность технического обслуживания, тыс. км;  $l_{\text{ТО}}^{\text{н}}$  — норматив периодичности технического обслуживания, тыс. км;  $L_{\text{к.р}}$  — пробег до капитального ремонта (ресурс), тыс. км;  $L_{\text{к.р}}^{\text{н}}$  — норматив пробега до капитального ремонта, тыс. км;  $t_{\text{т.р}}$  — удельная трудоемкость текущего ремонта, чел.-ч на 1 000 км;  $t_{\text{т.р}}^{\text{н}}$  — норматив удельной трудоемкости текущего ремонта, чел.-ч на 1 000 км;  $H_{3,ч}$  — расход запасных частей;  $H_{3,ч}^{\text{н}}$  — норматив расхода запасных частей;  $K_{1c}$  — коэффициент, учитывающий условия эксплуатации и категорию сложности маршрута движения;  $K_2$  — коэффициент, учитывающий модификацию подвижного состава и организацию его работы;  $K_3$  — коэффициент, учитывающий природно-климатические условия;  $K_4$  — коэффициент, учитывающий возраст подвижного состава (пробег с начала эксплуатации);  $K_5$  — коэффициент, учитывающий число обслуживаемых и ремонтируемых автомобилей на АТП и число технологически совместимых групп подвижного состава.

Нормативы периодичности ТО, ресурса, удельной трудоемкости ТР и коэффициенты  $K_2, K_3, K_4, K_5$  берутся из «Положения о ТО и ремонте подвижного состава автомобильного транспорта» и инструкций по эксплуатации той или иной модели автобусов. Значения норматива расхода запасных частей для той или иной модели автобусов определяются в соответствии с рекомендациями, изложенными в [8].

Значения коэффициента  $K_{1c}$  определяются по табл. 6.3.

При определении значений коэффициента корректирования  $K_{1c}$  необходимо знать категорию сложности маршрута движения. Для его определения необходимо в формулу (6.44) подставить численные значения факторов и получить численное значение обобщенного параметра сложности маршрута движения, по которому, используя табл. 6.4, определить категорию сложности.

Таблица 6.3

Коэффициенты корректирования нормативов  $K_{1c}$  в зависимости от категории условий эксплуатации и сложности маршрута

Категория условий эксплуатации	Категория сложности маршрута	Коэффициент корректирования $K_{1c}$			
		периодичности ТО	удельной трудоемкости ТР	ресурса	расхода запасных частей
1	—	1,00	1,00	1,00	1,00
2	—	0,90	1,10	0,90	1,10
3	1	0,80	1,20	0,80	1,25
	2	0,78	1,24	0,78	1,28
	3	0,76	1,28	0,76	1,31
	4	0,74	1,32	0,74	1,34
	5	0,72	1,36	0,72	1,37
4	1	0,70	1,40	0,70	1,40
	2	0,68	1,42	0,68	1,45
	3	0,66	1,44	0,66	1,50
	4	0,64	1,46	0,64	1,55
	5	0,62	1,48	0,62	1,60
5	—	0,60	1,50	0,60	1,65

Таблица 6.4

**Интервальные оценки разбиения по категориям сложности маршрутной сети ГУП «Мосгортранс» с учетом значений среднего квадратичного отклонения**

Категория сложности маршрута	Начальное значение интервала	Длина интервала	Конечное значение интервала
1	$-\infty$	1,536	-0,583
2	-0,583	0,862	0,278
3	0,278	0,862	1,140
4	1,140	0,862	2,002
5	2,002	2,528	$\infty$

С учетом изложенного были определены пробеги до списания современных автобусов типа МАЗ-103, ЛиАЗ-5256.25, ЛиАЗ-6212 (и др.) и разработаны рекомендации по списанию указанного подвижного состава в условиях автобусных парков. При этом было рассмотрено два варианта.

**Первый вариант.** В соответствии с рекомендациями заводов-изготовителей нормативный пробег до капитального ремонта у указанных автобусов составляет 500 тыс. км ( $L_{к.р}^н = 500\ 000$  км). Тогда в соответствии с Положением о ТО и ремонте подвижного состава автомобильного транспорта нормативный пробег до списания  $L_{сп}^н$  будет равен:

$$L_{сп}^н = 1,8L_{к.р}^н = 1,8 \cdot 500\ 000 = 900\ 000 \text{ км.}$$

Скорректируем этот нормативный пробег (фактический) до списания для условий Москвы для маршрутов различной категории сложности. Для 1-й категории сложности автобусного маршрута

$$L_{сп}^ф = 1,8L_{к.р}^н K_{1c} K_2 K_3 = 1,8 \cdot 500\ 000 \cdot 0,8 \cdot 1,0 \cdot 1,0 = 720\ 000 \text{ км.}$$

Для 2-й категории сложности автобусного маршрута

$$L_{сп}^ф = 1,8L_{к.р}^н K_{1c} K_2 K_3 = 1,8 \cdot 500\ 000 \cdot 0,78 \cdot 1,0 \cdot 1,0 = 702\ 000 \text{ км.}$$

Для 3-й категории сложности автобусного маршрута

$$L_{сп}^ф = 1,8L_{к.р}^н K_{1c} K_2 K_3 = 1,8 \cdot 500\ 000 \cdot 0,76 \cdot 1,0 \cdot 1,0 = 684\ 000 \text{ км.}$$

Для 4-й категории сложности автобусного маршрута

$$L_{сп}^ф = 1,8L_{к.р}^н K_{1c} K_2 K_3 = 1,8 \cdot 500\ 000 \cdot 0,74 \cdot 1,0 \cdot 1,0 = 666\ 000 \text{ км.}$$

Для 5-й категории сложности автобусного маршрута

$$L_{сп}^ф = 1,8L_{к.р}^н K_{1c} K_2 K_3 = 1,8 \cdot 500\ 000 \cdot 0,72 \cdot 1,0 \cdot 1,0 = 648\ 000 \text{ км.}$$

**Второй вариант.** На практике нормативный пробег до капитального ремонта составляет 400 тыс. км ( $L_{к.р}^н = 400\ 000$  км). Тогда в соответствии с Положением о ТО и ремонте подвижного состава автомобильного транспорта нормативный пробег до списания  $L_{сп}^н$  будет равен:

$$L_{сп}^н = 1,8L_{к.р}^н = 1,8 \cdot 400\ 000 = 720\ 000 \text{ км.}$$

Скорректируем этот нормативный пробег до списания для условий Москвы для маршрутов различной категории сложности. Для 1-й категории сложности автобусного маршрута

$$L_{сп}^ф = 1,8L_{к.р}^н K_{1c} K_2 K_3 = 1,8 \cdot 400\ 000 \cdot 0,8 \cdot 1,0 \cdot 1,0 = 576\ 000 \text{ км.}$$

Для 2-й категории сложности автобусного маршрута

$$L_{сп}^ф = 1,8L_{к.р}^н K_{1c} K_2 K_3 = 1,8 \cdot 400\ 000 \cdot 0,78 \cdot 1,0 \cdot 1,0 = 561\ 600 \text{ км.}$$

Для 3-й категории сложности автобусного маршрута

$$L_{сп}^ф = 1,8L_{к.р}^н K_{1c} K_2 K_3 = 1,8 \cdot 400\ 000 \cdot 0,76 \cdot 1,0 \cdot 1,0 = 547\ 200 \text{ км.}$$

Для 4-й категории сложности автобусного маршрута

$$L_{сп}^ф = 1,8L_{к.р}^н K_{1c} K_2 K_3 = 1,8 \cdot 400\ 000 \cdot 0,74 \cdot 1,0 \cdot 1,0 = 532\ 800 \text{ км.}$$

Для 5-й категории сложности автобусного маршрута

$$L_{сп}^ф = 1,8L_{к.р}^н K_{1c} K_2 K_3 = 1,8 \cdot 400\ 000 \cdot 0,72 \cdot 1,0 \cdot 1,0 = 518\ 400 \text{ км.}$$

В ГУП «Мосгортранс» сроки до списания и ресурс автобусов принято рассчитывать по второму варианту.

Далее был проведен расчет условного числа автобусов, подлежащих списанию:

$$N_{сп}^м = \frac{L_{г}^м}{L_{сп}^м},$$

где  $N_{сп}^м$  — условное число автобусов, подлежащих списанию на конкретном маршруте, ед.;  $L_{г}^м$  — годовой пробег на конкретном маршруте, км;  $L_{сп}^м$  — фактический пробег до списания автобусов с учетом категории сложности маршрута, км.

Годовые пробеги на конкретном маршруте задаются экономическим отделом автобусного предприятия. Фактические пробеги до списания автобусов с учетом категории сложности маршрута в данном случае берутся по двум рассмотренным ранее вариантам.

Расчет условного числа автобусов, подлежащих списанию по ГУП «Мосгортранс» за 2004 г. (2-й вариант), представлен в табл. 6.5. Таким

Таблица 6.5  
**Расчет планового условного числа автобусов, подлежащих списанию по ГУП «Мосгортранс» за 2004 г. (фрагмент)**

Номер маршрута	Номер автобусного парка	Обобщенный параметр сложности	Категория сложности	Пробег до списания, км	Годовой пробег автобусов		Условное число автобусов большого класса, подлежащих списанию, ед.	Условное число автобусов особо большого класса, подлежащих списанию, ед.
					большого класса, км	особо большого класса, км		
22	11	0,032	3	547 200	566 458	80 923	1,04	0,15
63	11	-0,097	3	547 200	226 161	376 934	0,41	0,69
65	11	-0,540	3	547 200	0	817 937	0,00	1,49
70	11	0,644	4	532 800	65 089	585 805	0,12	1,10
72	11	-0,381	3	547 200	73 780	811 575	0,13	1,48
87	11	-0,007	3	547 200	279 990	0	0,51	0,00
90	11	0,271	3	547 200	63 638	700 021	0,12	1,28
114	11	0,280	3	547 200	625 860	0	1,14	0,00
123	11	0,091	3	547 200	560 785	0	1,02	0,00
138	11	-0,760	2	561 600	156 465	0	0,28	0,00
149	11	0,161	3	547 200	76 527	765 273	0,14	1,40
170	11	-0,030	3	547 200	360 803	0	0,66	0,00
188	11	-0,910	2	561 600	85 177	255 532	0,15	0,46

191	11	-0,166	3	547 200	627 178	209 059	1,15	0,58	
194	11	-0,377	3	547 200	272 304	680 760	0,50	1,24	
199	11	-0,492	3	547 200	574 962	287 481	1,05	0,53	
200	11	-0,354	3	547 200	212 413	566 435	0,39	1,04	
233	11	0,110	3	547 200	115 259	288 147	0,21	0,53	
270	11	-0,404	3	547 200	229 555	0	0,42	0,00	
505	11	-3,960	1	576 000	265 496	0	0,46	0,00	
656	11	-0,125	3	547 200	325 447	0	0,59	0,00	
665	11	-0,049	3	547 200	129 125	0	0,24	0,00	
672	11	-0,116	3	547 200	227 438	758 127	0,42	1,39	
673	11	-0,289	3	547 200	304 585	304 585	0,56	0,56	
698	11	-0,368	3	547 200	121 219	0	0,22	0,00	
739	11	-0,896	2	561 600	155 770	0	0,28	0,00	
745	11	-0,967	2	561 600	100 589	201 178	0,18	0,36	
748	11	-0,141	3	547 200	241 560	0	0,44	0,00	
851	11	-1,650	2	561 600	0	2 112 113	0,00	3,76	
Всего:							12,83	17,84	
Число автобусов, подлежащих списанию по филиалу «11 автобусный парк» ГУП «Мосгортранс», ед.								13	18

Номер маршрута	Номер автобусного парка	Обобщенный параметр сложности	Категория сложности	Пробег до списания, км	Годовой пробег автобусов		Условное число автобусов большого класса, подлежащих списанию, ед.	Условное число автобусов особо большого класса, подлежащих списанию, ед.
					большого класса, км	особо большого класса, км		
5	14	0,135	3	547 200	151 304	0	0,28	0,00
32	14	-0,515	3	547 200	823 829	457 683	1,51	0,84
66	14	1,020	4	532 800	286 944	143 472	0,54	0,27
166	14	0,729	4	532 800	179 340	0	0,34	0,00
272	14	-2,550	1	576 000	192 826	128 551	0,33	0,22
611	14	-1,098	2	561 600	0	1 532 936	0,00	2,73
630	14	0,021	3	547 200	724 948	362 474	1,32	0,66
686	14	0,021	3	547 200	270 767	0	0,49	0,00
689	14	0,660	4	532 800	146 034	0	0,27	0,00
695	14	0,471	3	547 200	202 691	0	0,37	0,00
707	14	0,543	3	547 200	59 981	449 857	0,11	0,82
718	14	0,651	4	532 800	219 206	350 729	0,41	0,66
720	14	0,715	4	532 800	328 475	1 231 783	0,62	2,31

729	14	0,092	3	547 200	185 452	0	0,34	0,00
734	14	-0,091	3	547 200	83 009	83 009	0,15	0,15
746	14	-0,234	3	547 200	694 485	0	1,27	0,00
750	14	0,247	3	547 200	236 582	0	0,43	0,00
752	14	0,145	3	547 200	0	1 049 139	0,00	1,92
767	14	-0,613	2	561 600	322 926	1 227 120	0,58	2,19
785	14	-0,414	3	547 200	392 052	0	0,72	0,00
793	14	0,800	4	532 800	150 206	400 550	0,28	0,75
809	14	-0,496	3	547 200	280 210	70 052	0,51	0,13
810	14	0,757	4	532 800	248 554	894 794	0,47	1,68
812	14	-0,192	3	547 200	254 663	0	0,47	0,00
814	14	0,605	4	532 800	137 338	68 669	0,26	0,13
830	14	-1,194	2	561 600	609 527	870 753	1,09	1,55
Всего:				13,16			17,01	
Число автобусов, подлежащих списанию по филиалу «14 автобусный парк» ГУП «Мосгортранс», ед.				14	14	18		

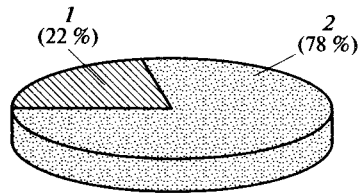


Рис. 6.10. Влияние «возраста» в годах 1 и пробега с начала эксплуатации в километрах 2 на срок службы городского автобуса

образом, было определено общее минимально необходимое число подлежащих списанию автобусов по филиалам «11 автобусный парк» и «14 автобусный парк ГУП «Мосгортранс» за 2004 г. (2-й вариант).

Оно должно составить 31 и 32 автобуса соответственно. Из них 27 автобусов большого класса и 36 автобусов — особо большого класса.

Для определения очередности списания того или иного автобуса необходимо использовать количественный критерий, называемый обобщенным параметром возраста и определяемый по формуле

$$V_{\text{гар.номер}} = 0,413B + 0,0000252L - 19,76,$$

где  $B$  — возраст списываемого автобуса, лет;  $L$  — пробег списываемого автобуса с начала эксплуатации, км.

Указанная математическая модель обобщенного параметра возраста была построена с учетом рекомендаций, изложенных в подразд. 6.1. Графическая интерпретация совместного влияния возраста списываемого автобуса в годах и пробега его с начала эксплуатации в километрах на срок службы представлена на рис. 6.10.

Первоочередному списанию подлежат автобусы, прошедшие нормативные пробеги и имеющие наибольшие значения обобщенного параметра возраста.

**Пример 1.** В филиале «14-й автобусный парк» ГУП «Мосгортранс» имеются семь автобусов «Икарус-260», подлежащих списанию. Их характеристика представлена в табл. 6.6.

Необходимо определить очередность списания указанного подвижного состава.

Находим численные значения обобщенного параметра возраста для каждого автобуса. Например, для автобуса гаражный номер 14 537:

$$V_{14537} = 0,413 \cdot 14,3 + 0,0000252 \cdot 726\,493 - 19,76 = 4,45.$$

Аналогично находим значения обобщенного параметра возраста для других автобусов. Данные заносим в табл. 6.6. Первоочередному списанию подлежат автобусы, имеющие наибольшие значения обобщенного параметра возраста. Таким образом, очередность списания автобусов «Икарус-260» в филиале «14-й автобусный парк» ГУП «Мосгортранс» должна быть следующей:

Таблица 6.6

Расчет первоочередного списания автобусов «Икарус-260» в филиале «14-й автобусный парк» ГУП «Мосгортранс»

Порядковый номер	Марка автобуса	Государственный номер	Гаражный номер	Дата ввода в эксплуатацию	Пробег с начала эксплуатации, км	Возраст автобуса, лет	Обобщенный параметр возраста
1	И-260	С 855 ВУ 77	14537	Ноябрь 1988	726 493	14,3	4,45
2	И-260	С 601 ВУ 77	14223	Август 1988	636 728	15,4	2,65
3	И-260.37	С 831 ВУ 77	14513	Сентябрь 1988	735 709	13,9	4,52
4	И-260.37	С 835 ВУ 77	14517	Май 1989	671 512	13,8	2,86
5	И-260.37	С 687 ВУ 77	14327	Декабрь 1987	611 039	15,3	1,96
6	И-260.37	С 706 ВУ 77	14347	Сентябрь 1989	604 612	15,1	1,71
7	И-260.37	С 779 ВУ 77	14444	Май 1989	647 602	13,8	2,26



- 1) «Икарус-260.37» (гаражный номер 14513);
- 2) «Икарус-260» (гаражный номер 14537);
- 3) «Икарус-260.37» (гаражный номер 14517);
- 4) «Икарус-260» (гаражный номер 14223);
- 5) «Икарус-260.37» (гаражный номер 14444);
- 6) «Икарус-260.37» (гаражный номер 14327);
- 7) «Икарус-260.37» (гаражный номер 14347).

**Пример 2.** Одной из основных задач технической эксплуатации автомобилей является обеспечение работоспособности подвижного состава на линии. Решение этой задачи обычно начинается с анализа факторов, влияющих на работоспособность. Применительно к городским автобусам корреляционный анализ показал, что на работоспособность оказывают влияние следующие факторы:

- $X_1$  — среднеквартальная температура воздуха;
- $X_2$  — средний пробег автобусов с начала эксплуатации;
- $X_3$  — удельный вес в парке автобусов с пробегом до  $0,25L_{к.р}$ ;
- $X_4$  — удельный вес в парке автобусов с пробегом  $0,25L_{к.р} - 0,5L_{к.р}$ ;
- $X_5$  — удельный вес в парке автобусов с пробегом  $0,5L_{к.р} - 0,75L_{к.р}$ ;
- $X_6$  — удельный вес в парке автобусов с пробегом  $0,75L_{к.р} - 1,0L_{к.р}$ ;
- $X_7$  — удельный вес в парке автобусов с пробегом свыше  $1,0L_{к.р}$ ;
- $X_8$  — средний возраст автобусов, лет;
- $X_9$  — средняя эксплуатационная скорость автобусов на маршруте;
- $X_{10}$  — среднесуточный пробег;
- $X_{11}$  — удельный пассажиропоток на маршруте;
- $X_{12}$  — длина маршрута движения;
- $X_{13}$  — удельное число остановок на маршруте;
- $X_{14}$  — удельное число перекрестков на маршруте;
- $X_{15}$  — удельное число светофоров на маршруте;
- $X_{18}$  — обеспеченность маршрута движения водителями;
- $X_{19}$  — удельный вес водителей первого класса;
- $X_{20}$  — коэффициент комплектации экипажей;
- $X_{21}$  — коэффициент числа замен водителя на маршруте;
- $X_{23}$  — общий стаж работы водителей;
- $X_{24}$  — среднемесячная заработная плата водителей;
- $X_{26}$  — периодичность ТО-2;
- $X_{27}$  — время в наряде;
- $X_{30}$  — удельное число автомобилемест для ТО и ТР в АТП;
- $X_{31}$  — коэффициент качества ТО и ТР автобусов;
- $X_{33}$  — средняя месячная заработная плата рабочих;
- $X_{34}$  — удельный расход запасных частей.

При этом было установлено, что влияющие на потери линейного времени по техническим причинам факторы связаны между собой как по физическому смыслу, так и математически (Приложение 6).



Рис. 6.11. Общая схема интерпретации главных компонент, определяющих работоспособность городских автобусов на линии

С учетом данного обстоятельства и закономерностей системного подхода была сформирована иерархическая структура факторов, определяющих работоспособность городских автобусов на линии (рис. 6.11).

Затем для каждого взаимосвязанного набора влияющих факторов был проведен компонентный анализ. В результате были получены первые главные компоненты для каждого взаимосвязанного набора влияющих факторов:

- по возрастной структуре:

$$V_1 = 0,55Z_2 - 0,5Z_3 + 0,05Z_4 + 0,01Z_5 + 0,42Z_6 + 0,27Z_7 + 0,44Z_8; \quad (6.49)$$

- по маршруту движения:

$$V_2 = -0,68Z_{12} + 0,58Z_{13} + 0,34Z_{14} + 0,29Z_{15}; \quad (6.50)$$

- по водительским кадрам:

$$V_3 = 0,09Z_{18} - 0,52Z_{19} - 0,48Z_{20} + 0,5Z_{21} - 0,32Z_{23} - 0,38Z_{24}; \quad (6.51)$$

- по интенсивности эксплуатации:

$$V_4 = 0,14Z_1 + 0,14Z_9 + 0,56Z_{10} + 0,56Z_{11} + 0,58Z_{27}; \quad (6.52)$$

- по системе и организации ТО и ремонта:

$$V_5 = 0,5Z_{26} - 0,57Z_{30} - 0,65Z_{34}; \quad (6.53)$$

- по уровню организации труда ремонтных рабочих:

$$V_6 = 0,71Z_{31} + 0,71Z_{33}, \quad (6.54)$$

где  $V_j$  ( $j = 1, 2, \dots, 6$ ) — главные компоненты, описывающие соответствующие группы факторов;  $Z_j$  — стандартизованные значения соответствующих влияющих факторов, определяемые по общепринятой методике (см. подразд. 6.1).

Затем на найденных первых главных компонентах была построена регрессионная модель параметра работоспособности городских автобусов на линии:

$$t_{\text{пр}}^T = 1,684 + 0,0755V_1 + 0,078V_2 + 0,2035V_3 + 0,0923V_4 + 0,037V_5 - 0,0533V_6, \quad (6.55)$$

где  $t_{\text{пр}}^T$  — потери линейного времени по техническим причинам на 1000 км, ч.

Результаты анализа полученной математической модели (табл. 6.7) показали, что на величину потерь линейного времени по техническим причинам при совместном действии исходных факторов в наибольшей степени оказывают влияние: водительские кадры  $V_3$ , интенсивность эксплуатации  $V_4$ , возрастная структура автобусов на маршруте  $V_1$ . Далее идут маршрут движения  $V_2$ , уровень организации труда ремонтных рабочих  $V_6$  и система и организация ТО и ремонта  $V_5$ .

Изменение значений компоненты, характеризующей возрастную структуру, на величину своего стандартного отклонения приводит к изменению потерь линейного времени по техническим причинам на 8,5%. Данное соотношение для маршрута движения составляет 6,1%; для водительских кадров — 21,8%; для интенсивности эксплуатации — 8,6%; для системы и организации ТО и ремонта — 3%; для уровня организации труда ремонтных рабочих — 4%.

В целях детального анализа работоспособности автобусов на линии была определена степень влияния вошедших в модель факторов на величину потерь линейного времени по техническим причинам. При этом был осуществлен переход от регрессии на главных компонентах к регрессии на исходных признаках:

Таблица 6.7

Влияние групп факторов на величину потерь линейного времени по техническим причинам

Показатель	Компоненты $V_j$					
	1	2	3	4	5	6
Вес компоненты $g_j$	0,162	0,078	0,514	0,193	0,020	0,033
Стандартизованный коэффициент регрессии $\beta_j$	0,204	0,166	0,589	0,233	0,082	-0,106

$$t_{\text{пр}}^T = 1,684 + 0,042Z_2 - 0,387Z_3 + 0,0038Z_4 + 0,0007Z_5 + 0,032Z_6 + 0,0204Z_7 + 0,038Z_8 - 0,053Z_{12} + 0,045Z_{13} + 0,026Z_{14} + 0,0226Z_{15} + 0,018Z_{18} - 0,1058Z_{19} - 0,0977Z_{20} + 0,1018Z_{21} - 0,065Z_{23} - 0,037Z_{24} + 0,0129Z_9 + 0,0517Z_{10} + 0,0517Z_{11} + 0,0749Z_{27} + 0,0185Z_{26} - 0,021Z_{30} - 0,024Z_{34} - 0,0378Z_{31} - 0,0378Z_{33} + 0,0129Z_1. \quad (6.56)$$

Оценка степени влияния проводилась по значению коэффициента эластичности  $\mathcal{E}_j$ . Расчеты показали (табл. 6.8), что при раздельном действии факторов наибольшее влияние на величину потерь линейного времени по техническим причинам оказывает коэффициент качества ТО и ТР автобусов в АТП.

Таблица 6.8

Влияние основных факторов на величину потерь линейного времени по техническим причинам на городских автобусных маршрутах

Факторы	Условное обозначение	Значение коэффициента эластичности $\mathcal{E}_i$
Средний пробег автобуса с начала эксплуатации, тыс. км	$X_2$	0,127
Средний возраст автобусов на маршруте движения, лет	$X_8$	0,100
Средняя эксплуатационная скорость автобусов на маршруте движения, км/ч	$X_9$	0,111

Окончание табл. 6.8

Факторы	Условное обозначение	Значение коэффициента эластичности Э,
Среднесуточный пробег автобусов на маршруте движения, км	$X_{10}$	0,193
Удельный пассажиропоток на маршруте движения, тыс. чел./авт.	$X_{11}$	0,139
Протяженность маршрута движения, км	$X_{12}$	-0,144
Обеспеченность водительскими кадрами маршрута движения, чел./авт.	$X_{18}$	0,149
Удельный вес водителей 1-го класса	$X_{19}$	-0,207
Коэффициент комплектации экипажей	$X_{20}$	-0,300
Коэффициент, учитывающий число замен водителей на маршруте, шт./чел.	$X_{21}$	0,356
Общий стаж работы водителей, лет	$X_{23}$	-0,189
Средняя месячная заработная плата водителей, руб.	$X_{24}$	-0,640
Периодичность ТО-2, тыс. км	$X_{26}$	0,220
Удельное число автомобилемест для ТО и ТР	$X_{30}$	-0,174
Время в наряде, ч	$X_{27}$	0,266
Коэффициент качества ТО и ТР автобусов в парке	$X_{31}$	-0,704
Средняя месячная заработная плата ремонтных рабочих	$X_{33}$	-0,651

Изменение его значения на 1 % приводит к изменению показателя работоспособности городских автобусов на линии на 0,7 %. Повышение коэффициента качества ТО и ТР на величину его стандартного отклонения позволит сократить потери линейного времени по техническим причинам на 2,3 %.

Существенное влияние на работоспособность автобусов городских маршрутов оказывают факторы организационного характера. Среди них особенно выделяется материальное стимулирование труда водителей и ремонтных рабочих ( $X_{24}$  и  $X_{33}$ ). Повышение значений указанных факторов на величину их стандартного отклонения позволит сократить потери линейного времени по техническим причинам на 4,6 и 2,3 % соответственно.

К факторам организационного характера относятся также повышение квалификации водителей, закрепление за автобусом водителей одной квалификации и стабильность работы водителя на автобусе (факторы  $X_{19}$ ,  $X_{20}$  и  $X_{21}$ ). При увеличении значений первых двух из них на величину стандартного отклонения потери времени по техническим причинам на линии уменьшаются на 6,3 и 5,8 % соответственно. Если снизить показатель  $X_{21}$  на величину его стандартного отклонения, то потери линейного времени по техническим причинам сократятся на 6 %.

К факторам организационного характера следует относить и учет возраста при распределении автобусов по маршрутам движения. При изменении значений показателей  $X_2$  и  $X_8$  на величину их стандартного отклонения потери линейного времени по техническим причинам изменяется на 2,5 и 2,3 % соответственно.

Совершенствование режимов технического обслуживания также оказывает влияние на работоспособность автобусов на линии. Снижение периодичности ТО-2 автобусов ЛиАЗ-677 до 12 тыс. км позволит сократить потери времени по техническим причинам на линии на 3,2 %.

Существенность влияния на работоспособность городских автобусов на линии производственно-технической базы АТП подтверждает фактор  $X_{30}$ : условий и интенсивности эксплуатации — факторы  $X_{12}$ ,  $X_9$ ,  $X_{10}$ ,  $X_{11}$ ,  $X_{27}$ . Данные показатели являются практически управляемыми на уровне предприятия. Тем не менее они должны быть учтены при анализе и планировании работоспособности автобусов на линии. Далее было проведено сравнение влияния на работоспособность городских автобусов как отдельных факторов, так и их совокупностей с учетом степени их управляемости на уровне автобусного автотранспортного предприятия.

Результаты анализа показали, что наибольшее влияние на величину потерь линейного времени по техническим причинам среди управляемых на уровне автобусного парка факторов оказывает возрастная структура парка (рис. 6.12). Установлено, что подвижной состав, имеющий пробег с начала эксплуатации 400 тыс. км и более,

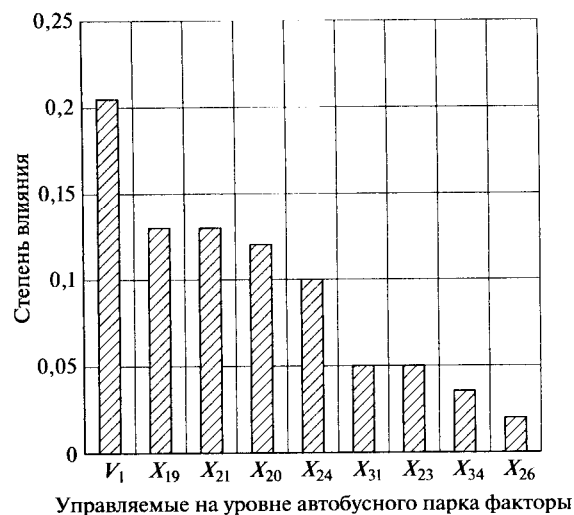


Рис. 6.12. Степень влияния управляемых на уровне автобусного парка факторов на потери линейного времени по техническим причинам

простаивает на линии по технической неисправности в 2 раза больше, чем автобусы с пробегом с начала эксплуатации до 95 тыс. км.

Существенно изменяются значения показателя работоспособности городских автобусов на линии в зависимости от сложности маршрута движения и интенсивности эксплуатации подвижного состава. При работе новых автобусов (пробег с начала эксплуатации не превышает 50 тыс. км) на маршрутах 10 и 90%-ной степеней сложности потери линейного времени по техническим причинам составляют соответственно 0,98 и 1,44 ч на 1 000 км пробега (или возрастают на 47 %).

Аналогичная зависимость наблюдается для линейной работоспособности и интенсивности эксплуатации. При повышении интенсивности эксплуатации нового подвижного состава на 80 % потери линейного времени по техническим причинам возрастают на 60 %. Следовательно:

- с помощью компонентного анализа было оценено влияние на работоспособность автобусов на линии не только отдельных факторов, но и их совокупностей. Это более объективно воспроизводит исследуемый физический процесс и дает более достоверные результаты при прочих равных условиях;

- важнейшим фактором повышения работоспособности подвижного состава на линии на уровне автотранспортного предприятия является управление возрастной структурой парка;

- приведение в действие выявленных резервов позволит предприятию сократить потери линейного времени по техническим причинам на 25... 40 %.

**Пример 3.** Для обоснованного выбора моторного масла в эксплуатации после проведения предварительного анализа смазочных материалов на основании существующих классификаций (рис. 6.13) необходимо руководствоваться технико-экономическим критерием, представленным в виде:

$$C_{м.уд} = \frac{C_m}{LV_m} K \rightarrow \min, \quad (6.57)$$

где  $C_{м.уд}$  — удельная стоимость моторного масла, руб./л на 1 000 км;  $C_m$  — стоимость моторного масла, руб./л;  $L$  — периодичность замены моторного масла, тыс. км;  $V_m$  — функция, характеризующая качество моторного масла;  $K$  — коэффициент, показывающий реализацию технико-эксплуатационных свойств моторных масел в эксплуатации.

Функция, характеризующая качество моторного масла, является обобщенным параметром технико-эксплуатационных свойств моторного масла. Параметры технико-эксплуатационных свойств моторного масла регламентируются ГОСТ Р 51634—2000 «Масла моторные автотракторные. Общие технические требования». Данный

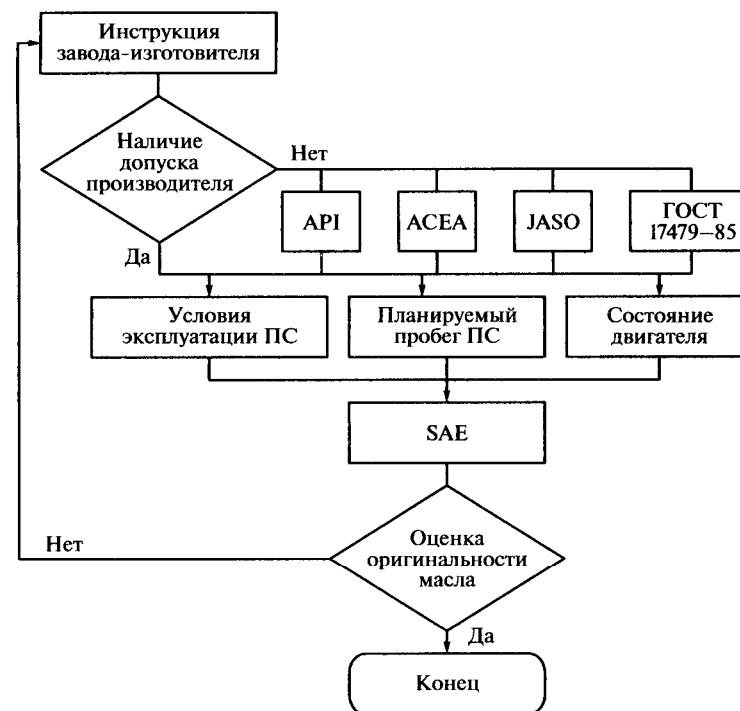


Рис. 6.13. Алгоритм предварительного выбора моторного масла

стандарт распространяется на группу однородной продукции — автотракторные моторные масла, устанавливает показатели качества, характеризующие безопасность продукции и подлежащие обязательному включению во все виды документации, по которой изготавливают моторные масла. К данным параметрам относятся:

- кинематическая вязкость;
- температура вспышки, определяемая в открытом тигле;
- массовая доля фосфора;
- сульфатная зольность;
- щелочное число.

Эти параметры характеризуют уровень эксплуатационных свойств и вязкостно-температурные свойства моторных масел и являются взаимосвязанными между собой. Корреляционный анализ показателей моторных масел, применяемых для городских автобусов в ГУП «Мосгортранс», подтвердил данное обстоятельство (табл. 6.9).

Таблица 6.9

**Коэффициенты парной корреляции параметров технико-эксплуатационных свойств моторных масел городских автобусов**

Факторы	$K_{40}$	$K_{100}$	$T_B$	$M_{\Phi}$	$C_3$	ЩЧ
$K_{40}$ — кинематическая вязкость при 40 °С, мм <sup>2</sup> /с	1,00	0,504	0,668	-0,072	0,028	0,53
$K_{100}$ — кинематическая вязкость при 100 °С, мм <sup>2</sup> /с	0,504	1,00	0,277	0,072	-0,338	0,064
$T_B$ — температура вспышки, определяемая в открытом тигле, °С	0,668	0,277	1,00	0,484	-0,286	0,315
$M_{\Phi}$ — массовая доля фосфора, %	-0,072	0,072	0,484	1,00	-0,415	0,123
$C_3$ — сульфатная зольность, %	0,028	-0,348	-0,286	-0,415	1,00	-0,106
ЩЧ — щелочное число, мг КОН на 1 г масла	0,53	0,064	0,315	0,123	-0,106	1,00

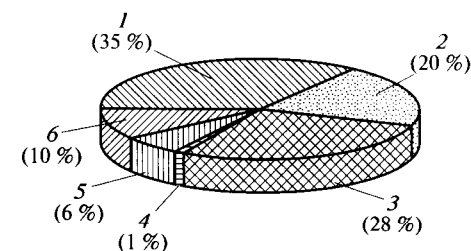


Рис. 6.14. Веса факторов в математической модели обобщенного параметра технико-эксплуатационных свойств моторного масла городских автобусов:

1 — кинематическая вязкость при 40 °С; 2 — кинематическая вязкость при 100 °С; 3 — температура вспышки, определяемая в открытом тигле; 4 — массовая доля фосфора; 5 — сульфатная зольность; 6 — щелочное число

Поэтому для нахождения функции  $V_m$ , характеризующей качество моторного масла, был проведен компонентный анализ, на основании которого была определена первая главная компонента:

$$V_m = 0,08K_{40} + 0,33K_{100} + 0,03T_B + 0,53M_{\Phi} - C_3 + 0,18ЩЧ - 15. \quad (6.58)$$

Указанная компонента может быть использована в качестве обобщенного параметра технико-эксплуатационных свойств моторного масла, так как включает в себя все показатели, регламентированные ГОСТ Р 51634—2000 «Масла моторные автотракторные. Общие технические требования».

Анализ математической модели обобщенного параметра технико-эксплуатационных свойств моторного масла показывает, что наибольшее влияние на его формирование оказывает кинематическая вязкость при 40 °С (35 %); далее идут температура вспышки, определяемая в открытом тигле (28 %), кинематическая вязкость при 100 °С (20 %), щелочное число (10 %), сульфатная зольность (6 %) и массовая доля фосфора (1 %).

Графическая интерпретация вклада индивидуальных параметров технико-эксплуатационных свойств моторного масла в формирование обобщенного параметра технико-эксплуатационных свойств моторного масла городских автобусов представлена на рис. 6.14.

Вращение полученных в результате компонентного анализа главных компонент не привело ни к повышению их собственных чисел, ни к увеличению доли описываемой ими дисперсии. Таким образом, был реализован 5-й вариант принципиальной схемы взаимодействия влияющих факторов «дерева систем» ТЭА при компонентном анализе (см. подразд. 6.1).

Коэффициент  $K$ , показывающий реализацию технико-эксплуатационных свойств моторных масел в эксплуатации при выборе свежего масла, принимается равным единице.

Качество моторного масла, находящегося в эксплуатации, определяется по формуле

$$K_j = \frac{V_j^H}{V_j^Ф}, \quad (6.59)$$

где  $V_j^H$  — нормативное значение обобщенного параметра технико-эксплуатационных свойств  $j$ -го моторного масла;  $V_j^Ф$  — текущее (фактическое) значение обобщенного параметра технико-эксплуатационных свойств  $j$ -го моторного масла.

Текущее значение обобщенного параметра технико-эксплуатационных свойств  $j$ -го моторного масла определяется по формуле (6.58) путем подстановки в них текущих значений каждого из параметров технико-эксплуатационных свойств отработавших моторных масел.

Нормативное значение обобщенного параметра технико-эксплуатационных свойств  $j$ -го моторного масла определяется также по формуле (6.58), только путем подстановки в нее нормативных значений каждого из параметров технико-эксплуатационных свойств свежих моторных масел. Нормативные значения кинематической вязкости (индекса вязкости), щелочного числа, массовой доли фосфора, сульфатной зольности и температуры вспышки приводятся в паспорте моторного масла.

**Пример 4.** При оценке влияния условий эксплуатации на расход топлива городского автобуса «Икарус-280», оборудованного нейтрализатором отработавших газов (НОГ) блочного типа, были определены наиболее значимые факторы:

- $\gamma_j$  — коэффициент использования пассажироместности оборотного рейса  $j$ -го маршрута, ед./км;
- $l_{nj}$  — средневзвешенная длина перегона  $j$ -го маршрута, км;
- $v_{эj}$  — средняя эксплуатационная скорость автобуса на  $j$ -м маршруте, км/ч;
- $\rho_{срj}$  — средневзвешенная плотность транспортного потока на  $j$ -м маршруте, число автомобилей на 100 м;
- $t$  — температура окружающей среды при  $j$ -м выходе автобуса на линию из парка, °С.

В то же время было установлено, что первые четыре фактора являются взаимосвязанными между собой (табл. 6.10).

При построении простой многофакторной регрессии расхода топлива на этих значимых факторах часть из них необходимо отбросить, так как модель должна строиться на независимых переменных.

В результате же компонентного анализа первые четыре фактора были объединены в главную компоненту:

$$V = -0,875\gamma + 0,928l_n + 0,937v_э - 0,887\rho. \quad (6.60)$$

Таблица 6.10

Коэффициенты парной корреляции исследуемых факторов с показателем маршрутного расхода топлива автобусов «Икарус-280», оборудованных НОГ модели НД59-14А-00000

Факторы	$Q$	$\gamma$	$l_n$	$v_э$	$\rho$	$t$
Расход топлива $Q$	1,00	0,64	-0,87	-0,88	0,62	-0,49
Коэффициент использования пассажироместности $\gamma$	0,64	1,00	-0,69	-0,71	0,79	-0,07
Средняя длина перегона $l_n$	-0,87	-0,69	1,00	0,95	-0,72	0,28
Средняя эксплуатационная скорость $v_э$	-0,88	-0,71	0,95	1,00	-0,72	0,29
Плотность транспортного потока $\rho$	0,62	0,79	-0,72	-0,72	1,00	0,074
Температура окружающей среды $t$	-0,49	-0,07	0,28	0,29	0,07	1,00

Было определено собственное число наиболее значимой главной компоненты ( $\lambda_1 = 3,292$ ) и на ее основе построена регрессионная модель на новых независимых переменных с суммарной дисперсией 84 %:

$$Q = 56,19 - 0,542V - 0,063t. \quad (6.61)$$

Затем был осуществлен переход от главной компоненты к исходным факторам:

$$Q = 56,19 + 0,474\gamma - 0,503l_n - 0,508v_э + 0,481\rho - 0,063t. \quad (6.62)$$

Проверка полученной модели на значимость и адекватность показала хорошую сходимость с экспериментальными данными. Основные статистические характеристики модели (формула (6.62)) представлены в табл. 6.11.

Таблица 6.11

Статистические характеристики адекватности модели

Коэффициент корреляции	Коэффициент детерминации	Критерий Фишера		Стандартная ошибка аппроксимации, %
		табличный	расчетный	
0,92	0,84	2,18	47,56	1,0

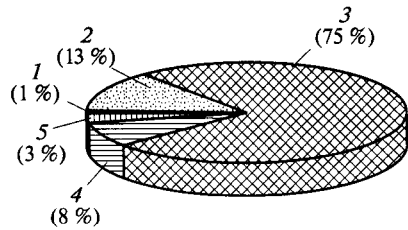


Рис. 6.15. Веса факторов в математической модели маршрутного расхода топлива автобуса «Икарус-280», оборудованного нейтрализатором НД59-14А-00000:

1 — коэффициент использования пассажироместности; 2 — средневзвешенная длина перегона; 3 — средняя эксплуатационная скорость; 4 — средневзвешенная плотность транспортного потока; 5 — температура окружающей среды

Таким образом, был реализован 2-й вариант принципиальной схемы взаимодействия влияющих на расход топлива факторов «дерева систем» ТЭА при компонентном анализе (см. подразд. 6.1).

Анализ полученной модели с использованием коэффициента эластичности показал, что первое место по эластичности занимает фактор «средневзвешенная длина перегона». Изменение его на 5% приводит к изменению функции отклика на 3,5%. Второе место по эластичности по отношению к маршрутному расходу топлива занимает фактор «средневзвешенная удельная плотность транспортного потока». Изменение данного фактора на 5% приводит к изменению функции отклика на 2,3%. Третье место по эластичности занимает фактор «средняя эксплуатационная скорость». Изменение значения данного фактора на 5% приводит к изменению показателя расхода топлива на 1,2%.

Значительно меньшей эластичностью обладают такие факторы, как «температура окружающей среды» и «коэффициент использования пассажироместности». Увеличение температуры окружающей среды на величину стандартного отклонения приводит к уменьшению расхода топлива на 1,5%. Возрастание коэффициента использования пассажироместности на значение стандартного отклонения приводит к повышению расхода топлива на 0,2%.

Графическая интерпретация математической модели представлена на рис. 6.15.

Приведенные данные свидетельствуют о том, что исключение из модели ряда взаимосвязанных факторов, как предполагает многошаговый регрессионный анализ, привело бы к определенным ошибкам при нормировании расхода топлива городского автобуса «Икарус-280», оборудованного нейтрализатором отработавших газов блочного типа в эксплуатации.

**Пример 5.** При оценке влияния условий эксплуатации на расход топлива городского автобуса ЛиАЗ-677 с нейтрализатором обрабо-

тавших газов типа 17.1206.010 были определены наиболее значимые факторы:

- $v_s$  — средняя эксплуатационная скорость на маршруте, км/ч;
- $\gamma$  — коэффициент использования пассажироместности (средневзвешенный уровень загрузки автобуса по перегону);
- $\rho$  — средняя плотность транспортного потока, число автомобилей на 100 м;
- $l_n$  — средняя длина перегона, км;
- $N_{ост}$  — удельное количество остановок (помехонасыщенность) на маршруте, ед./км.

При этом было установлено, что все они являются взаимосвязанными (табл. 6.12) и при многошаговом регрессионном анализе часть из них необходимо исключать, так как переменные должны быть не зависимыми величинами.

Для оценки комплексного влияния рассмотренных факторов на расход топлива городского автобуса ЛиАЗ-677 с НОГ типа 17.1206.010 с использованием математического аппарата компонентного анализа были определены главные компоненты.

В результате первого этапа компонентного анализа, при ортогональном преобразовании корреляционной матрицы, получены собственные значения дисперсии и факторные нагрузки двух главных компонент (табл. 6.13).

Таблица 6.12

**Коэффициенты парной корреляции исследуемых факторов с показателем маршрутного расхода топлива автобуса ЛиАЗ-677 с НОГ типа 17.1206.010**

Факторы	$Q$	$v_s$	$\gamma$	$\rho$	$l_n$	$N_{ост}$
Расход топлива $Q$	1,00	-0,45	0,58	0,21	-0,48	0,44
Средняя эксплуатационная скорость $v_s$	-0,45	1,00	-0,13	0,49	0,73	-0,75
Коэффициент использования пассажироместности $\gamma$	0,58	-0,13	1,00	0,41	-0,49	0,24
Плотность транспортного потока $\rho$	0,21	0,49	0,41	1,00	0,21	-0,24
Средняя длина перегона $l_n$	-0,48	0,73	-0,49	0,21	1,00	-0,75
Помехонасыщенность на маршруте $N_{ост}$	0,44	-0,75	0,24	-0,24	-0,75	1,00

Таблица 6.13

## Главные компоненты и их статистические характеристики

Метод вращения главных осей	Собственное значение		Дисперсия	
	Компоненты 1 ( $V_1$ )	Компоненты 2 ( $V_2$ )	Компоненты 1 ( $V_1$ )	Компоненты 2 ( $V_2$ )
Без вращения	$V_1 = 0,926v_3 + 0,739\gamma + 0,848\rho + 0,782I_n - 0,599N_{\text{ост}}$ (6.63)			
	$V_2 = 0,07v_3 + 0,446\gamma + 0,464\rho - 0,476I_n + 0,692N_{\text{ост}}$ (6.64)			
	3,09	1,12	0,62	0,22
Варимакс	$V_1 = 0,776v_3 + 0,858\gamma + 0,955\rho + 0,329I_n - 0,053N_{\text{ост}}$ (6.65)			
	$V_2 = 0,510v_3 + 0,098\gamma + 0,149\rho + 0,855I_n - 0,914N_{\text{ост}}$ (6.66)			
	2,36	1,86	0,47	0,37
Биквартимакс	$V_1 = 0,789v_3 + 0,860\gamma + 0,958\rho + 0,351I_n - 0,076N_{\text{ост}}$ (6.67)			
	$V_2 = 0,490v_3 + 0,076\gamma + 0,125\rho + 0,846I_n - 0,913N_{\text{ост}}$ (6.68)			
	2,41	1,81	0,48	0,36
Квартимакс	$V_1 = 0,806v_3 + 0,862\gamma + 0,962\rho + 0,381I_n - 0,109N_{\text{ост}}$ (6.69)			
	$V_2 = 0,461v_3 + 0,045\gamma + 0,091\rho + 0,833I_n - 0,909N_{\text{ост}}$ (6.70)			
	2,48	1,74	0,50	0,35

На втором этапе компонентного анализа в целях поиска оптимального положения главных компонент, представленных в нормализованном виде в многофакторном пространстве, выполнено вращение осей различными методами (в частности, методами *варимакс*, *биквартимакс*, *квартимакс*).

Затем, также используя стандартные программы (Mathcad 6.0 PLUS, Statistica 5.0), с помощью ПЭВМ на базе экспериментального материала были построены корреляционно-регрессионные математические модели маршрутного расхода топлива на двух главных компонентах с последующим переходом к исходным факторам. При этом с незначительным преимуществом предпочтение было отдано математическим моделям, полученным при вращении осей методом *варимакс* (табл. 6.14).

В результате анализа полученных математических моделей расхода топлива автобуса ЛиАЗ-677 с НОГ типа 17.1206.010 приходим к выводу, что доминирующее значение в формировании расхода топлива играет помехонасыщенность маршрута  $N_{\text{ост}} = 58\%$ . Меньший весовой вклад имеют средняя эксплуатационная скорость на маршруте  $v_3 = 22\%$ , средняя плотность транспортного потока  $\rho = 12\%$ , средняя длина перегона  $I_n = 6\%$  и коэффициент использования пассажироместности  $\gamma = 2\%$  (рис. 6.16).

Таблица 6.14

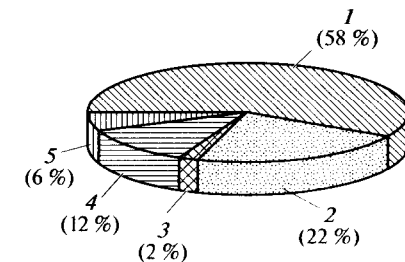
## Многофакторные математические модели маршрутного расхода топлива автобуса ЛиАЗ-677 с нейтрализатором отработавших газов типа 17.1206.010

Метод вращения главных осей	Коэффициент корреляции	Коэффициент детерминации	Критерий Фишера	Табличное значение критерия Фишера	Стандартная ошибка аппроксимации, %
Без вращения	$Q = 43,48 - 0,61v_3 + 2,96\gamma + 2,98\rho - 5,25I_n + 5,33N_{\text{ост}}$ (6.71)				
	0,677	0,458	9,301	2,14	2,631
Варимакс	$Q = 43,52 - 0,61v_3 + 2,96\gamma + 2,98\rho - 5,25I_n + 6,91N_{\text{ост}}$ (6.72)				
	0,677	0,459	9,316	2,14	2,629
Биквартимакс	$Q = 43,48 - 0,61v_3 + 2,96\gamma + 2,98\rho - 5,24I_n + 6,91N_{\text{ост}}$ (6.73)				
	0,677	0,458	9,300	2,14	2,631
Квартимакс	$Q = 43,48 - 0,61v_3 + 2,96\gamma + 2,98\rho - 5,25I_n + 6,91N_{\text{ост}}$ (6.74)				
	0,677	0,458	9,302	2,14	2,631

Полученные данные свидетельствуют о том, что удаление из модели любого из взаимосвязанных факторов ведет к существенным ошибкам при нормировании расхода топлива автобуса ЛиАЗ-677 с НОГ типа 17.1206.010 в эксплуатации. Таким образом, был реализован 3-й вариант принципиальной схемы взаимодействия влияющих на расход топлива факторов «дерева систем» ТЭА при компонентном анализе (см. подразд. 6.1).

**Пример 6.** Для нормирования маршрутного ресурса шин городских автобусов (Р) было необходимо выявить влияющие факторы и построить на них математическую модель. Для этого сначала

Рис. 6.16. Веса факторов в математической модели маршрутного расхода топлива автобуса ЛиАЗ-677, оборудованного нейтрализатором типа 17.1206.010:



1 — помехонасыщенность маршрута; 2 — средняя эксплуатационная скорость; 3 — коэффициент использования пассажироместности; 4 — средняя плотность транспортного потока; 5 — средняя длина перегона



был проведен корреляционный анализ предварительно отобранных факторов. Результаты корреляционного анализа представлены в табл. 6.15. Анализ табл. 6.15 показал, что они практически все являются взаимосвязанными. Например, средняя скорость сообщения  $v_c$  связана со средней эксплуатационной скоростью  $v_3$ , средней длиной перегона  $l_n$ , коэффициентом использования пассажироместности  $\gamma$  и плотностью транспортного потока  $\rho$ , а средняя длина перегона связана с удельным числом поворотов  $n_n$  и т. д.

Таблица 6.15

**Матрица коэффициентов парной корреляции факторов, влияющих на маршрутный ресурс шин городских автобусов ЛиАЗ-5256.25**

	P	$v_3$	$l_n$	$n_n$	$\gamma$	$\rho$	$v_c$
P	1,00	0,583	0,543	-0,155	-0,618	-0,670	0,781
$v_3$	0,583	1,00	0,826	-0,328	-0,215	-0,286	0,896
$l_n$	0,543	0,826	1,00	-0,239	-0,133	-0,281	0,717
$n_n$	-0,155	-0,328	-0,239	1,00	0,383	-0,025	-0,127
$\gamma$	-0,618	-0,215	-0,133	0,383	1,00	0,677	-0,343
$\rho$	-0,670	-0,286	-0,281	-0,025	0,677	1,00	-0,333
$v_c$	0,781	0,896	0,717	-0,127	-0,343	-0,333	1,00

Кроме того, было установлено, что скорость сообщения оказывает более существенное влияние на ресурс шины городских автобусов, чем средняя эксплуатационная скорость (коэффициент парной корреляции равен 0,781 вместо 0,583), поэтому последняя была выведена из дальнейшего рассмотрения (в частности, при проведении компонентного анализа и построении многофакторной математической модели маршрутного ресурса шины на главных компонентах).

Компонентный анализ проходил в два этапа.

При проведении первого этапа компонентного анализа были получены собственные значения, дисперсии и факторные нагрузки компонент для ресурса шин городского автобуса ЛиАЗ-5256.25. На втором этапе были определены три наиболее значимые главные компоненты (собственные значения главных компонент больше единицы, и на графике «каменистой осыпи» имеется характерный перегиб у третьей компоненты).

Для проведения этих операций использовалась компьютерная программа SPSS 11.0. При этом применение компонентного анализа при нормировании ресурса шин городских автобусов следует рассматривать как инновационное решение, направленное на повышение эффективности эксплуатации подвижного состава. В связи с тем,

что регрессионная модель ресурса шины городского автобуса ЛиАЗ-5256.25 строится на исходных факторах, имеющих разную размерность, необходимо определить стандартизованные значения переменных компоненты (см. формулу (6.8)).

Средние значения и среднеквадратические отклонения для каждой переменной приведены в табл. 6.16.

Таблица 6.16

**Среднее значение и среднеквадратическое отклонение переменной**

Параметр	P	$v_3$	$l_n$	$n_n$	$\gamma$	$\rho$	$v_c$
$\bar{X}$	68 933	20,72	1,03	1,044	0,466	1,878	22,16
$\sigma$	6 056	3,868	0,95	0,622	0,096	1,177	3,6

Для определения числа компонент, принимаемых в рассмотрение для построения математической модели, необходимо вычислить собственные значения (табл. 6.17), по которым можно также определить процент воспроизведения общей дисперсии. Для этого применялись критерий Кайзера, который предлагает принимать в рассмотрение компоненты с собственным значением больше единицы (результаты представлены в табл. 6.17) и критерий «каменистой осыпи» (результаты представлены на рис. 6.17).

Таблица 6.17

**Собственные значения компонент**

Компонента	$\lambda$	Процент воспроизведения общей дисперсии
1	2,344	46,881
2	1,165	23,294
3	1,028	20,552
4	0,343	6,853
5	0,121	2,42

По критерию Кайзера исходя из вычисленных собственных значений главных компонент рекомендуется принять в рассмотрение три компоненты в связи с тем, что они имеют собственные числа больше единицы и воспроизводят 91 % общей дисперсии.

По критерию «каменистой осыпи» с учетом данных графика, приведенного на рис. 6.17, можно сказать, что рекомендуется брать в рассмотрение также три компоненты. Затем были определены математические формулы по каждой значимой главной компоненте. В стан-

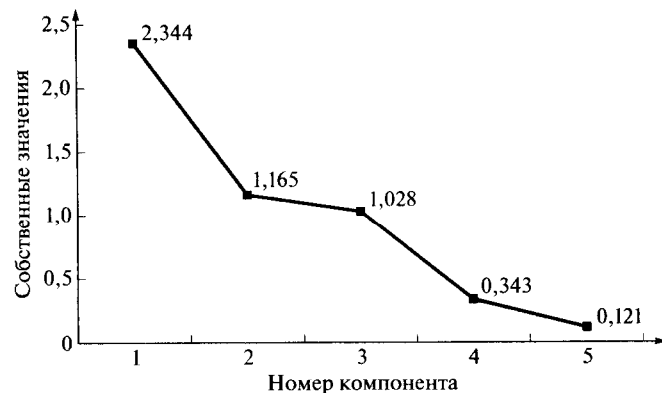


Рис. 6.17. График «каменистой осыпи» маршрутного ресурса шин городских автобусов

дартизованном виде полученные компоненты имеют следующий вид:

$$V_1 = -0,335Z_{v_c} - 0,304Z_{l_n} + 0,318Z_{\gamma} + 0,171Z_{n_n} + 0,303Z_{\rho}; \quad (6.75)$$

$$V_2 = 0,392Z_{v_c} + 0,531Z_{l_n} + 0,512Z_{\gamma} + 0,059Z_{n_n} + 0,396Z_{\rho}; \quad (6.76)$$

$$V_3 = 0,148Z_{v_c} + 0,013Z_{l_n} + 0,094Z_{\gamma} + 0,877Z_{n_n} - 0,415Z_{\rho}. \quad (6.77)$$

Модель маршрутного ресурса шин на главных компонентах имеет следующий вид:

$$P = 68\,933 - 5\,162V_1 + 25V_2 + 1\,256V_3. \quad (6.78)$$

Подставляя значения главных компонент в формулу (6.78), получаем модель маршрутного ресурса шины при стандартизованных значениях влияющих факторов:

$$P = 68\,933 + 1\,925Z_{v_c} + 1\,599Z_{l_n} - 1\,511Z_{\gamma} + 220,6Z_{n_n} - 2\,075,6Z_{\rho}. \quad (6.79)$$

С учетом формулы (6.8) и данных табл. 6.16 преобразуем модель на стандартизованных значениях влияющих факторов в математическую модель на исходных признаках:

$$P = 68\,917 + 534,8v_c + 1\,683,8l_n - 15\,636,7\gamma + 354,9n_n - 1\,764\rho. \quad (6.80)$$

Анализ статистических характеристик полученной математической модели (табл. 6.18) свидетельствует о том, что она является адекватной исследуемому процессу и исследуемым исходным данным.

Статистические характеристики математической модели маршрутного ресурса шин городского автобуса ЛиАЗ-5256.25 на главных компонентах

Коэффициент множественной корреляции	Коэффициент детерминации	Критерий Фишера	Табличное значение критерия Фишера	Средняя ошибка аппроксимации, %
0,877	0,77	7,798	2,66	5,03

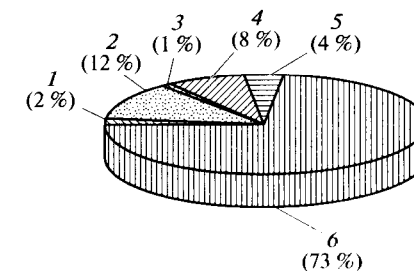
Анализ многофакторной регрессионной модели ресурса шин 11/70R-22,5 автобусов ЛиАЗ-5256.25 на главных компонентах показал (рис. 6.18), что доминирующее влияние на ресурс шины из факторов условий эксплуатации оказывает скорость сообщения — 12 %, далее идет коэффициент использования пассажироместности — 8 %, средняя плотность транспортного потока — 4 %, средняя длина перегона — 2 %, удельное число поворотов — 1 %.

Суммарное влияние факторов условий эксплуатации составляет 27 %. Остальное влияние (73 %) на ресурс шины 11/70R-22,5 городского автобуса ЛиАЗ-5256.25 оказывают конструкционные, технологические, производственные и природно-климатические факторы (рис. 6.19).

Исследование полученной регрессионной модели маршрутного ресурса шин городского автобуса ЛиАЗ-5256.25 на главных компонентах с помощью коэффициента эластичности показало, что среди факторов условий эксплуатации по отношению к маршрутному ресурсу шины до списания наибольшей эластичностью обладает коэффициент использования пассажироместности. При изменении его на 1 % маршрутный ресурс шины городского автобуса ЛиАЗ-5256.25 изменяется на 0,12 %. Далее по эластичности идут средняя плотность транспортного потока на маршруте — 0,054 %, средняя длина перегона — 0,028 %, скорость сообщения — 0,02 %, удельное число поворотов на маршруте — 0,006 %.

Рис. 6.18. Влияние факторов на ресурс шины 11/70R-22,5 автобуса ЛиАЗ-5256.25:

1 — средняя длина перегона; 2 — скорость сообщения; 3 — удельное число поворотов; 4 — коэффициент использования пассажироместности; 5 — средняя плотность транспортного потока; 6 — конструкционные, технологические, производственные и природно-климатические факторы



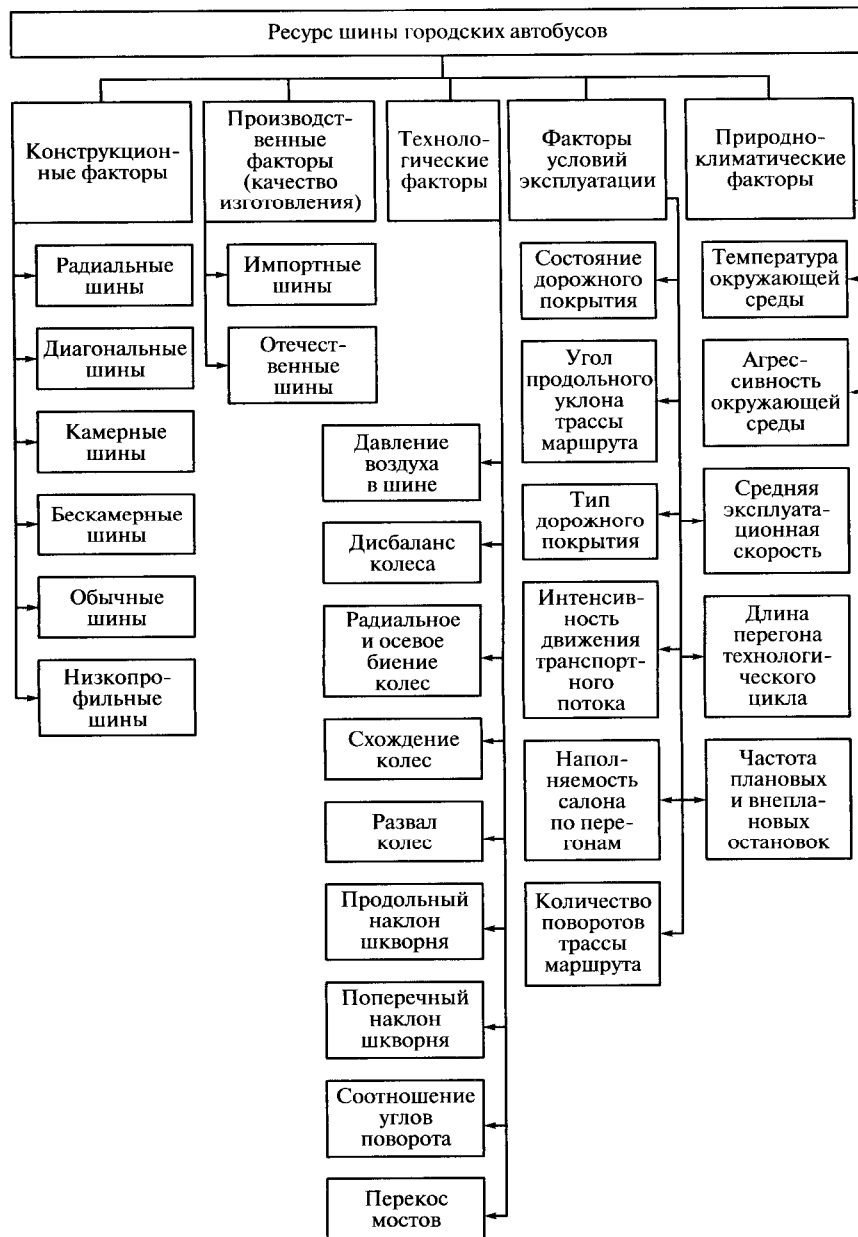


Рис. 6.19. Классификация факторов, влияющих на ресурс шин городских автобусов в эксплуатации

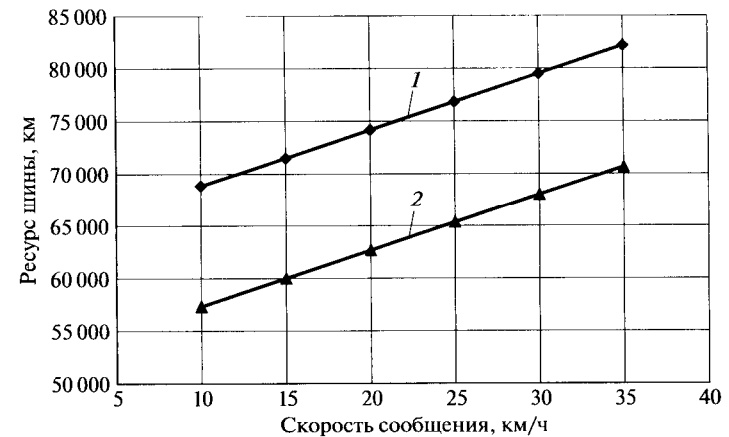


Рис. 6.20. Изолинии маршрутного ресурса шины 11/70R-22,5 автобуса ЛиАЗ-5256.25 в зависимости от скорости сообщения при минимальных (1) и максимальных (2) значениях значимых факторов условий эксплуатации

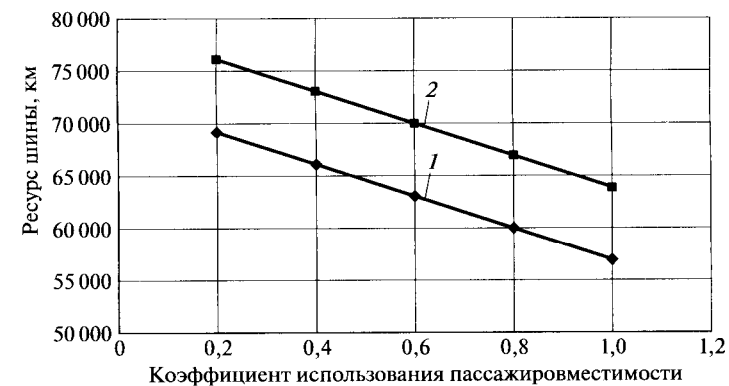


Рис. 6.21. Изолинии маршрутного ресурса шины 11/70R-22,5 автобуса ЛиАЗ-5256.25 в зависимости от коэффициента использования пассажироместности при минимальных (1) и максимальных (2) значениях значимых факторов условий эксплуатации

Зависимость маршрутного ресурса шины 11/70R-22,5 автобуса ЛиАЗ-5256.25 при экстремальных значениях значимых факторов условий эксплуатации от скорости сообщения и коэффициента использования пассажироместности представлена на рис. 6.20 и 6.21 соответственно.

Таким образом, был реализован 3-й вариант принципиальной схемы взаимодействия влияющих на ресурс шин городских автобусов

факторов «дерева систем» ТЭА при компонентном анализе (см. подразд. 6.1).

### КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Назовите основные принципиальные подходы при оценке влияющих факторов с использованием многошагового регрессионного и компонентного анализов.
2. Что такое главные компоненты?
3. Назовите основные числовые характеристики главных компонент.
4. По каким критериям отбираются главные компоненты?
5. Назовите основные методы ортогонального вращения главных компонент (факторов).
6. Назовите основные методы косоугольного вращения главных компонент (факторов).
7. Перечислите основные методы вращения главных компонент (факторов), основанные на введении вторичных осей.
8. Перечислите основные факторы, определяющие сложность маршрута движения городских автобусов.
9. В чем состоит оперативное корректирование основных нормативов ТЭА?
10. Какие факторы определяют обобщенный параметр возраста городских автобусов?
11. Перечислите основные факторы, влияющие на работоспособность городских автобусов.
12. Назовите основной параметр работоспособности городских автобусов на линии.
13. Перечислите пять основных факторов, влияющих на работоспособность городских автобусов на линии.
14. Каким критерием необходимо руководствоваться при выборе моторных масел в эксплуатации?
15. Какие факторы влияют на выбор моторных масел в эксплуатации?
16. Какие факторы условий эксплуатации влияют на расход топлива городских автобусов?
17. Какие факторы условий эксплуатации влияют на ресурс шин городских автобусов?
18. Какие факторы влияют на маршрутный ресурс шин городских автобусов?
19. По каким критериям оценивается адекватность математических моделей?
20. Какой физический смысл имеет коэффициент детерминации?
21. По какому параметру отбираются значимые факторы при корреляционном анализе?

## ГЛАВА 7

### Имитационное моделирование как метод получения количественных оценок перспективных организационных и технологических систем поддержания работоспособности автомобилей

#### 7.1. Возможности имитационного моделирования в исследовании вариантов применения внешнего и встроенного диагностирования на автомобильном транспорте

Полученные на основе системного подхода и системного анализа рекомендации по целесообразной структуре внешнего и встроенного диагностирования [3], массовое применение которых на автомобилях следует считать основой совершенствования процессов технической эксплуатации автомобилей в настоящее время, дают возможность целенаправленно осуществить следующий более сложный этап решения указанной проблемы при помощи имитационного моделирования, без которого переход к последующей широкой практике может оказаться малореальным.

*Имитационное моделирование* — это процесс построения и анализа имитационной модели в целях выявления внутрисистемных закономерностей и связей.

Имитационная модель — это объективно существующая зависимость (математическая, логическая, экономическая и др.) между выходом и входом рассматриваемой системы, полученная, как правило, с использованием ЭВМ и воспроизводящая реальный процесс или конструкцию технического устройства с определенной точностью и заданными ограничениями.

В области ТЭА имитационное моделирование применяется при определении периодичности ТО, в деловых играх, при описании и анализе работы производственных подразделений и технических механизмов, анализе стратегий развития систем поддержания работоспособности автомобилей в эксплуатации и др.

В данной главе рассмотрим возможности имитационного моделирования на автомобильном транспорте.

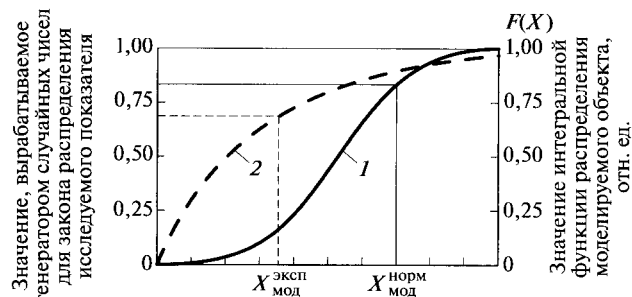
При этом задачей моделирования будет являться получение всесторонней оценки прогнозируемых изменений в системе поддержания работоспособности автомобилей, базирующихся на конкретных зна-

чениях технико-экономических показателей. Однако решение этой задачи для автомобильного транспорта в целом, включающего в себя автомобили индивидуального пользования, акционерных коллективов сельского хозяйства, небольших автотранспортных организаций, в настоящее время в полной мере не может быть осуществлено из-за отсутствия необходимых исходных данных, отражающих специфику их технической эксплуатации.

Что касается автотранспорта общего пользования, базирующегося на комплексных автотранспортных предприятиях, то здесь решение данной задачи является острой необходимостью в связи с тенденцией существенной перестройки структуры системы поддержания работоспособности, и для этого имеются все необходимые предпосылки. За 70-е годы XX в. был накоплен значительный опыт моделирования многих процессов АТП. Однако при этом в основном рассматривались задачи, связанные с внутрипроизводственной оптимизацией работы зон ТО и ремонта в рамках принятой двухступенчатой профилактической системы.

Моделирование использовалось также для оптимизации загрузки участков и зон первых вариантов стационарного диагностирования (Д-1 и Д-2), которые при этом рассматривались как элементы системы массового обслуживания. Все это существенно упрощало задачу моделирования новых процессов работоспособности автомобилей в АТП с более широкой оценкой последствий от их внедрения.

Как уже отмечалось ранее, важнейшими показателями такой оценки должны являться: изменение расхода топлива и загрязнения окружающей среды в зависимости от уровня поддерживаемого технического состояния; влияние этого уровня на безопасность движения; эффек-



Смоделированная величина наработки агрегата или узла, тыс. км., трудоемкости (чел.-мин) или времени (мин) его ремонта и т.д.

Рис. 7.1. Графическое представление процессов имитационного моделирования возникновения отказов и неисправностей исследуемых элементов автомобиля и восстановления их исправного состояния на основе использования интегральных оценок закономерностей рассеивания показателя:

1 — пример для нормального закона; 2 — пример для экспоненциального закона

тивность использования трудовых ресурсов ремонтно-обслуживающего персонала и водителей при восстановлении работоспособности автомобилей. Поэтому основная трудность решения поставленной задачи получения количественных оценок заключается в разработке подмоделей, учитывающих изменение указанного ранее комплекса показателей, имеющих социальное значение, при одновременной их увязке с общей технико-экономической оценкой. Разработка таких подмоделей в целом не встречает принципиальных затруднений.

На рис. 7.1 дано графическое представление моделирования процессов появления отказов или неисправностей для отдельных элементов автомобиля (узлов, агрегатов, систем и деталей), основанное на предварительно изученных закономерностях рассеивания их ресурса в эксплуатации (см. подразд. 2.7), а также процессов их восстановления (по времени) на основе ремонтных воздействий или обслуживания. Здесь за основу берутся интегральные оценки закономерностей рассеивания показателя (рис. 7.2, выражение (2.18)), вытекающие из дифференциальных функций, по которым на основе разыгрывания генератором случайных чисел для соответствующего закона распределения величины в диапазоне от 0 до 1 может быть получено соответствующее значение смоделированного показателя  $X_{\text{мод}}$ . Полученная величина запоминается и используется для дальнейших расчетов.

Указанная процедура многократно повторяется для всех исследуемых элементов одного автомобиля и для всего подвижного состава, находящегося в автотранспортной организации, а получаемые при этом выходные показатели по отдельным процессам обобщаются на основе суммирования и более сложных видов обработки. Таким образом, может быть смоделирована работа АТП за год, два и другие периоды, на которых могут быть изучены закономерности изменения представляющих интерес выходных показателей.

Естественно, что исследование на модели процесса поддержания работоспособности автомобилей в условиях АТП будет являться частным и далеко не характерным случаем для автомобильного транспорта в целом, поэтому могут возникнуть сомнения в ценности полученных на основании этого исследования выводов, в частности, касающихся путей развития диагностирования. Усилению этих сомнений способствует также перспектива применения на автотранспорте общего пользования централизованного обслуживания и ремонта, предусматривающая отход от комплексных автотранспортных предприятий. Однако подобная точка зрения в целом является неправомерной и может служить только аргументом в пользу более детального моделирования специфики технической эксплуатации автомобилей в различных организационных условиях, для решения текущих задач управления. Что касается стратегических задач, то общеметодическое значение результатов моделирования перспективных организационно-технических вариантов, необходимость разра-



Рис. 7.2. Структурная схема проведения имитационного моделирования для АТП общего пользования

ботки которых следует из этапа системных исследований, будет значительно шире рассматриваемых рамок АТП.

Основатель динамического программирования Р. Беллман считает, что для многих процессов более важными являются общие представления, чем конкретные значения констант. На основе моделирования возможных вариантов применения стационарного и встроенного диагностирования для условий работы комплексного автотранспортного предприятия и сопоставления выявленных при этом относительных изменений как раз и могут быть получены общие представления

о процессах использования диагностирования, которые могут быть положены в разработку новых концепций в области технической эксплуатации автомобилей.

Это связано с тем, что как в комплексных АТП, так и в акционерных предприятиях сельского хозяйства и небольших ведомственных транспортных организациях эксплуатируются одни и те же модели автомобилей. Совокупность таких автомобилей различных организаций, прикрепленных в рыночных условиях к одному централизованному предприятию технического обслуживания и ремонта, с позиций технической эксплуатации можно рассматривать как аналог одного комплексного АТП с присущими ему значениями констант. Поскольку одной из основ моделирования, отмеченных ранее, является сравнительное исследование вариантов, можно считать, что качественные относительные изменения, выявленные при моделировании определенного типа АТП, будут характерны и для АТП других типов (с другими значениями констант).

Необходимо также отметить, что полученные при моделировании значения констант всегда будут отличаться от значений для реального АТП, поскольку на модели может быть исследована некоторая идеализированная и достаточно жестко ограниченная ситуация. Приведенные обстоятельства, как уже отмечалось ранее, существенно повышают общеметодическую ценность результатов моделирования процессов использования диагностирования на примере АТП. Можно предполагать, что полученные при этом некоторые качественно-количественные зависимости будут правомерны и для автотранспорта индивидуального пользования.

Таким образом, наилучшая изученность основных технологических процессов АТП, с одной стороны, и наибольшая сложность их организации, с другой, повышают возможность реализации модельного исследования закономерностей использования диагностирования и значимость полученных при этом результатов. Моделирование должно проводиться на примере хорошо изученной как в отношении эксплуатационной надежности, так и в отношении процессов диагностирования модели автомобиля. Такой моделью в период проведения исследований являлся бензиновый автомобиль средней грузоподъемности.

Поскольку процессы диагностирования систем питания и зажигания являются практически неизменными для всех автомобилей с бензиновыми двигателями, в том числе и легковых, результаты моделирования по влиянию диагностирования на расход топлива и загрязнение окружающей среды имеют широкое обобщение.

Возможности диагностирования в своевременном выявлении потребности в текущем ремонте агрегатов, систем и узлов автомобиля, отказы которых происходят в результате постепенного изменения технического состояния, требуют специфического моделирования процессов как возникновения, так и выявления, предотвращения

и устранения этих отказов производственной зоной АТП. Аварийные отказы в подобных случаях более чем вдвое увеличивают трудовые и материальные затраты на ремонт по сравнению с их профилактическим предупреждением на стадии неполного использования ресурса. За счет диагностирования появляется возможность выявлять предотказное состояние, когда ресурс выработан практически полностью, но аварийные отказы еще не произошли, и на этой стадии производить профилактические ремонты.

Общие эксплуатационные затраты в такой ситуации будут зависеть от наличия соответствующих методов и средств диагностирования и их прогнозирующих возможностей, с одной стороны, и стратегий (вариантов) использования этих средств, определяющих возможности своевременного выявления предотказного состояния и пропуска аварийных отказов, с другой. Правильный учет этих обстоятельств приводит к значительному усложнению процесса моделирования параметра потока заявок на ремонт по сравнению с ранее рассматриваемыми задачами по оптимизации системы массового обслуживания машин, когда исходный поток требований мог описываться более упрощенными законами, в частности стационарным и нестационарным пуассоновским.

Основная трудность при этом будет заключаться в необходимости моделирования общего потока отказов, потребность проверки технического состояния которых стимулировало развитие методов и средств диагностирования. Одновременно по этим и другим агрегатам и узлам автомобиля должны моделироваться случайные отказы, которые не могут быть предупреждены методами диагностики. Согласно исследованиям доля таких отказов в общем потоке заявок, формирующих объем текущего ремонта, значительна (30... 50 %), и их неправильный учет может привести к переоценке роли диагностирования. Отмеченные ранее обстоятельства требуют более детального описания закономерностей появления отказов и неисправностей многих элементов автомобиля, чем это делалось ранее.

На рис. 7.2 приведена наиболее целесообразная структура имитационной модели процессов поддержания работоспособности грузовых автомобилей с карбюраторными двигателями (типа малотоннажных грузовиков Горьковского автозавода) в условиях АТП, удовлетворяющая перечисленным ранее требованиям и позволяющая проводить исследования в широком диапазоне вариантов использования диагностирования. Ее принципиальной особенностью является разделение на две взаимосвязанные, но в то же время достаточно независимые части: моделирование процессов изменения технического состояния и процессов его восстановления в зонах ремонта и обслуживания АТП.

Особенностью первой части является помимо потока заявок на ремонт и обслуживание с учетом постепенных и случайных отказов моделирование неисправностей, не вызывающих прекращения

транспортного процесса, но влияющих на расход топлива, загрязнение окружающей среды, безопасность движения и экономические затраты. Такое разделение позволяет осуществить моделирование изменения технического состояния для автомобилей, находящихся в любых специфических условиях эксплуатации, в том числе и индивидуального пользования. Естественно, что выходные показатели модели первой части будут существенно зависеть от организации процессов восстановления, определяющих не только временные потери на ремонт и обслуживание, но и глубину воздействия на техническое состояние. Поскольку указанная организация неодинакова для автотранспорта общего пользования, автомобилей акционерных коллективов сельского хозяйства и индивидуального пользования, соответственно должны отличаться и модели второй части.

В то же время указанный подход дает возможность на первом этапе моделирования получать некоторые обобщенные, в определенной мере идеализированные показатели процесса восстановления, непосредственно отражающие технические характеристики автомобиля, без ограничений, связанных с пропускными и производственными возможностями АТП, станций технического обслуживания (СТО) централизованного типа и автосервиса, а также потерями времени на перегоны автомобилей и ожидания в очереди. При этом возможно более рельефно оценить изменение технических характеристик автомобиля при различных вариантах использования диагностирования, а именно эта информация, в первую очередь, необходима для заводов-изготовителей, выпускающих автомобили универсального назначения и ответственных за совершенствование их эксплуатационных показателей.

Соответственно на втором этапе моделирование необходимо проводить с учетом производственных и организационных ограничений АТП и СТО и более реально оценить возможности диагностирования в различных условиях технической эксплуатации. Одновременно также можно произвести сравнительную оценку различных систем организации ТО и ремонта автомобилей. Как уже отмечалось ранее, второй этап моделирования реально возможен и для АТП общего пользования в связи с хорошей изученностью внутрипроизводственных процессов.

## **7.2. Основные стратегии поддержания исправного технического состояния для отдельного элемента (детали, узла, агрегата) автомобиля**

Примерно половина всех отказов автомобилей в эксплуатации обусловлена процессами изнашивания или постепенным накоплением усталостных повреждений и неисправностей [8]. Если отказав-

ший элемент такого типа заменить на новый в процессе ремонта или обслуживания, то имеется определенная гарантия того, что на последующем периоде эксплуатации отказ будет исключен (если элемент не имел дефектов производственного характера) до момента, когда интегральная величина износа или уровня накопления усталостных повреждений и неисправностей не превысит так называемого предельного уровня.

Интенсивность приближения к этому уровню зависит не только от степени технологического совершенства элемента, но и от условий эксплуатации и должна постоянно изучаться и анализироваться в целях выработки практических рекомендаций о целесообразных технических воздействиях на автомобиль, в первую очередь по пробегу, в целях исключения излишних материальных затрат и нежелательных в социальном отношении последствий от отказов.

Как было отмечено в гл. 2, наиболее распространенным методом оценки интенсивности изменения технического состояния является получение статистических закономерностей рассеивания ресурса однотипных элементов в эксплуатации, по параметрам которых можно делать определенные количественные оценки. Практическим путем установлено [3, 8], что для неслучайных отказов (для элементов с постепенным изнашиванием или накоплением неисправностей) характерными являются нормальный закон рассеивания ресурса (с коэффициентами вариации  $v = 0,25 \dots 0,35$ ) и Вейбулла ( $v = 0,4 \dots 0,7$ ); чисто случайные отказы интерпретируются экспоненциальным законом (со статистическими оценками коэффициента вариации  $v = 0,75 \dots 1,2$ ).

Если отказы и неисправности, описываемые экспоненциальным распределением, практически являются неуправляемыми в эксплуатации (за некоторым исключением, о чем будет сказано далее), то для законов нормального и Вейбулла в силу закономерного характера постепенного перехода в состояние неисправности и отказа имеются определенные возможности управления на основе использования различных стратегий обслуживания и ремонта, реализуемых, как правило, в определенных сочетаниях в последующих организационных и технологических системах поддержания работоспособного состояния автомобилей в эксплуатации. Использование принципов системного подхода подтверждает целесообразность выделения четырех основных стратегий применительно к отдельному элементу автомобиля, графическая интерпретация которых дана далее на примере нормального закона.

На рис. 7.3 приведена вероятностная схема 1-й стратегии обслуживания и ремонта на основе аварийных отказов, которые в силу случайных, непредсказуемых без дополнительных воздействий специального характера (типа технической диагностики) моментов их возникновения для конкретного индивидуального объекта являются в данном случае единственным естественным способом выявления факта полного использования ресурса элемента, заложенного заво-

дом-изготовителем. Штриховка всей площади под кривой распределения условно показывает 100%-ную вероятность аварийных отказов ( $q = 1$ ); при этом полностью реализуются индивидуальные ресурсы всей статистической совокупности элементов, а среднее значение ресурса в эксплуатации составит  $L_{cp}$ .

Естественная простота этой стратегии, а также полная реализация ресурса элементов автомобиля, заложенного заводом-изготовителем, обеспечили ее широкое распространение на заре развития автомобильного транспорта, которое сохраняется и до настоящего времени, несмотря на достижения научно-технического прогресса (в рядовой эксплуатации, даже без учета отказов случайного характера, по этой стратегии поддерживается работоспособное состояние в 80...90 % случаях восстановления дефектов автомобилей, находящихся в средних и крупных автотранспортных предприятиях общего пользования, и практически в 100 % случаев для автомобилей небольших предприятий и индивидуальных владельцев, а также автолюбителей).

С системных позиций необходимо также учитывать то обстоятельство, что именно в рамках 1-й стратегии осуществлялось формирование специфики технологических процессов обслуживания и ремонта автомобилей и развития производственно-технической базы, являющейся основой технической эксплуатации до настоящего времени. Однако данная стратегия имеет существенные недостатки как в экономическом, так и в социальном отношении, что приводит к необходимости существенного ограничения ее использования несмотря на все сложности данного процесса. В экономическом отношении это связано с тем, что последствия устранения любого, даже незначительного, аварийного отказа всегда значительно выше (в 2—10 раз) по сравнению с профилактическим воздействием [8], под которым следует понимать заблаговременное предупредительное техническое воздействие (обслуживание или замену элемента с уже выработанным ресурсом), выполняемое в условиях предприятия до начала выполнения автомобилем транспортного процесса и, таким образом, исключающего возникновение отказа на дороге.

Если обобщенную стоимость устранения аварийного отказа обозначить через  $C$  [3, 8], а стоимость профилактического воздействия через  $d$ , то можно ввести удобное понятие коэффициента опасности отказа  $k$ , которое способствует лучшему практическому восприятию идеологии широкого применения работ предупредительного характера по поддержанию работоспособного состояния, рекомендуемого как заводами-изготовителями, так и эксплуатационниками в течение уже 100-летнего периода развития автомобильного транспорта:

$$k = \frac{C}{d}. \quad (7.1)$$

С учетом введенных понятий целесообразно перейти к экономическим критериям оценки стратегий, без которых затруднительно



получить системно взвешенное решение по их выбору. Для этого удобно использовать критерий удельных затрат на поддержание работоспособного состояния [3, 8], который для 1-й стратегии согласно схеме, представленной на рис. 7.3, будет иметь вид

$$c_{уд}^I = \frac{C}{L_{ср}} = \frac{kd}{L_{ср}}. \quad (7.2)$$

Из выражения (7.2) видно, что, несмотря на полную реализацию индивидуального ресурса, затраты в сфере технической эксплуатации являются довольно большими ввиду значительных величин коэффициентов опасности аварийных отказов; это является недостатком данной стратегии в экономическом отношении. Другим существенным недостатком 1-й стратегии следует считать ее неприемлемость для поддержания работоспособного состояния элементов автомобиля, отказы которых помимо экономических имеют отрицательные социальные последствия (снижают уровни безопасности движения и экологической безопасности, несмотря на резервирование данных процессов, предусмотренных конструкцией автомобиля).

Отмеченные ранее недостатки в определенной степени устраняет 2-я, так называемая плано-предупредительная, стратегия, графическая интерпретация которой приведена на рис. 7.4. Здесь в целях исключения аварийных отказов для основной части элементов производится принудительное ограничение их используемого в эксплуатации ресурса величиной  $L_p$ , значительно меньшей величины среднего ресурса  $L_{ср}$ . При этом будет иметь место индивидуально

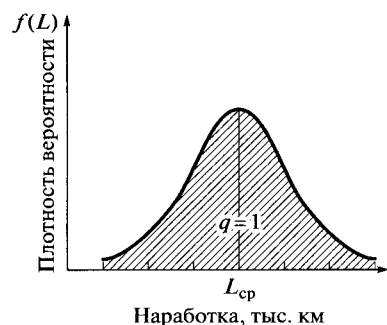


Рис. 7.3. Графическая интерпретация 1-й стратегии обслуживания и ремонта — по потребности по мере наступления аварийного отказа — для отдельного элемента автомобиля, изменение состояния которого связано с изнашиванием или постепенным накоплением неисправностей

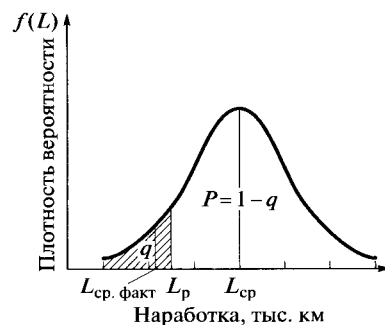


Рис. 7.4. Графическая интерпретация 2-й, плано-предупредительной, стратегии обслуживания и ремонта для отдельного элемента автомобиля, изменение состояния которого связано с изнашиванием или постепенным накоплением неисправностей

непредсказуемая вероятность аварийного отказа  $q$  (заштрихованная область под кривой рассеивания ресурса), величина которой может регулироваться задаваемым значением регламентного, или так называемого гамма-процентного, ресурса  $L_p$  [3, 8] (индекс «р» отражает так называемую вероятность безотказной работы  $p = 1 - q$ , отражаемую незаштрихованной частью площади под кривой рассеивания ресурса).

Из схемы, представленной на рис. 7.4, можно видеть, что снижение вероятности  $q$  достигается за счет значительного недоиспользования индивидуального ресурса элементов, заложенного заводом-изготовителем. Данное обстоятельство является основным недостатком плано-предупредительной стратегии, повсеместно вызывающей перекосы при ее использовании. Однако, несмотря на ограничение ресурса экономические показатели данной стратегии (при условии правильного подбора величины  $L_p$ ), оказываются значительно выше по сравнению с 1-й стратегией (удельные затраты на поддержание работоспособного состояния заметно снижаются за счет резкого сокращения доли аварийных отказов  $q$ ).

Теоретически величина удельных затрат может быть подсчитана по достаточно сложному выражению, учитывающего вероятностные процессы и характер кривой рассеивания ресурса [3, 8]:

$$c_{уд}^{II} = \frac{Cq(L_p) + d(1 - q(L_p))}{L_{ср.факт}} = \frac{kdq(L_p) + d(1 - q(L_p))}{L_{ср.факт}} \rightarrow \min. \quad (7.3)$$

Из схемы, представленной на рис. 7.4, можно видеть, что величина среднего фактического ресурса  $L_{ср.факт}$ , реализуемого 2-й стратегией, будет меньше ограничивающего ресурса  $L_p$  за счет неизбежных более ранних аварийных отказов для части элементов (для практических расчетов ресурсы можно принимать одинаковыми), а удельные затраты по выражению (7.3) могут быть минимизированы подбором вероятности  $q(L_p)$ , определяемой величиной  $L_p$ .

Сложный процесс минимизации выражения (7.3) для диапазона значений коэффициента опасности отказов  $k = 2 \dots 10$  и характерных для эксплуатации коэффициентов вариации рассеивания ресурса  $v = 0,3 \dots 0,7$  показывает, что при этом будет иметь место достаточно высокий целесообразный уровень вероятности аварийного отказа  $q = 0,10 \dots 0,25$ , который недостаточен для узлов и систем автомобиля, обеспечивающих безопасность движения, при уровне недоиспользования среднего ресурса в 2—2,5 раза. Указанный уровень недоиспользования интуитивно воспринимается как эксплуатационный резерв повышения эффективности работы автомобилей за счет определенного превышения рекомендуемых заводом-изготовителем

гамма-процентных ресурсов  $L_p$  по основным агрегатам, узлам, механизмам и деталям.

Следует отметить, что если раньше установление рациональных режимов  $L_p$  требовало сложного комплекса привязанных к базовым условиям эксплуатации исследований и непростую систему корректировки этих показателей для конкретных условий эксплуатации (поскольку заводы-изготовители были заинтересованы в условиях централизованного управления показывать завышенные «рекламные» результаты), то в рыночной экономике плано-предупредительные режимы должны быть отражены в так называемой сервисной книжке автомобиля, фактически являющейся составной частью инструкции по эксплуатации.

Для социально опасных узлов и систем автомобиля величины  $L_p$  при этом определяются исходя из граничного уровня вероятности отказа  $q = 0,05$  (а не по условию (7.3) [3, 8]), поскольку его дальнейшее ужесточение приводит уже к неприемлемым для эксплуатации автомобилей уровням недоиспользования ресурсов, заложенных заводами-изготовителями.

Как показывает практика, реальные удельные затраты на поддержание работоспособного состояния автомобиля при строгом соблюдении рекомендаций плано-предупредительного подхода, что обычно имеет место для автомобилей, выполняющих ответственные социально-значимые перевозки (например, автобусные) в условиях областных и краевых центров и крупных городов, значительно выше по сравнению с более простыми периферийными условиями, где допускаются менее строгие подходы к технической эксплуатации на основе 1-й стратегии. Несомненным преимуществом 2-й стратегии являются лучшие возможности для крупного предприятия предупредительного планирования работ по поддержанию работоспособности автомобилей, хотя для условий автомобиля индивидуального владельца подобное планирование, без наличия определенного взаимно-обязательного договора с сервисным центром, является уже несущественным.

Плановые возможности 2-й стратегии в целом оказываются довольно ограниченными, поскольку в связи с большим количеством критических в отношении надежности узлов и агрегатов автомобиля, имеющих свои сильно различающиеся режимы  $L_p$ , в идеале необходимо иметь большое количество согласованных по наработке (пробегу) «ступеней» технологических воздействий [3], что практически неосуществимо. Реальное количество таких ступеней в передовом мировом опыте не превышало четырех-пяти, а для отечественных условий равнялось двум (без учета ежедневного обслуживания).

Плано-предупредительная стратегия оказала определенное воздействие на формирование производственной базы крупных автотранспортных предприятий СССР, в частности, обусловила появление так называемых поточных линий ТО-1 и ТО-2, состоящих из

трех-четырех проездных постов, которые на практике зачастую использовались в другом режиме (линия ТО-1 в режиме одного, иногда двух технологических постов, линия ТО-2 в режиме стационарных постов, хотя последнее приводило к трудностям в случае значительной неравномерности объемов работ по автомобилям, установленным на посты линии).

Указанные недостатки 2-й стратегии, в том числе касающиеся ограниченных возможностей процессов планирования, в значительной степени устраняются введением (для крупных автотранспортных предприятий) 3-й стратегии, графическая интерпретация которой приведена на рис. 7.5.

Здесь весь диапазон рассеивания ресурса как бы покрывается по наработке «сеткой» регламентного стационарного (т.е. выполняемого в условиях предприятия) диагностирования, выполняемого для всех без исключения автомобилей с периодичностью  $L_d$ , в процессе которого по факту превышения предельно допустимого значения диагностического параметра (см. подразд. 3.4) индивидуально определяется так называемое предотказное состояние контролируемого объекта и выполняется его предупредительный ремонт или обслуживание стоимостью  $d$ . Согласно исследованиям наработка диагностируемого объекта в предотказном состоянии составляет в среднем около 10 % от его индивидуального ресурса, поэтому позволяет при владении системы периодического контроля определенным образом исключать из эксплуатации объекты с предельным состоянием до момента их аварийного отказа.

Таким образом, после каждой плановой диагностической проверки все автомобили разделяются на два класса: неисправные и исправные. Исправные автомобили могут эксплуатироваться с незначительной вероятностью отказов  $q_{di}$  на  $i$ -м межконтрольном пробеге  $L_{di}$ , уровень которой может регулироваться (уменьшаться или увеличиваться) соответствующим изменением  $L_d$ . Вероятность отказов  $q_{di}$  на схеме рис. 7.5 графически отражена заштрихованной областью на некотором интервале, предшествующем моменту планового диагностирования, поскольку после его проведения такие отказы практи-

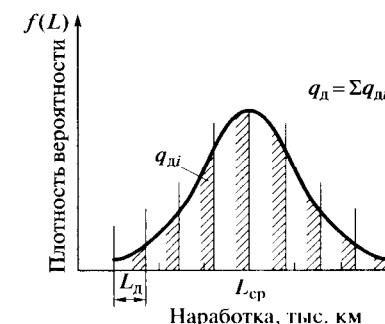


Рис. 7.5. Графическая интерпретация 3-й стратегии обслуживания и ремонта — на основе стационарного регламентного (периодического) диагностирования — для отдельного элемента автомобиля, изменение состояния которого связано с изнашиванием или постепенным накоплением неисправностей

чески невозможны в связи с выявлением и исключением из эксплуатации объектов, находящихся в предотказном состоянии (теоретическая возможность ошибочного диагноза здесь исключена).

Общая вероятность отказов  $q_d$  при такой системе будет определяться суммированием  $q_{di}$  по всем необходимым для данного диапазона рассеивания числом межконтрольных пробегов, как это показано на рис. 7.5, а снижение периодичности  $L_d$  (установление более мелкой сетки), приводящее к уменьшению вероятности отказов  $q_d$ , неизбежно сопровождается увеличением затрат на диагностирование.

Таким образом, основным достоинством 3-й стратегии является полная реализация среднего ресурса, заложенного заводом-изготовителем, при расширенных возможностях эксплуатационного регулирования уровней вероятности аварийных отказов. Это дает значительное снижение экономических и временных затрат на обслуживание и ремонт автомобилей, которые теоретически могут быть отражены следующей зависимостью для удельных затрат:

$$c_{уд}^{III} = \frac{k d q_d(L_d) + d(1 - q_d(L_d))}{L_{cp}} + \frac{c_d}{L_d} \rightarrow \min. \quad (7.4)$$

Данное выражение уже содержит составляющие, отражающие затраты на проведение диагностирования в виде стоимости единичной контрольной проверки  $c_d$  и периодичности ее выполнения  $L_d$ . При известном (фиксированном) значении первой величины удельные затраты могут быть оптимизированы (минимизированы) путем подбора соответствующего значения  $L_d$ . Очевидно, что повышение стоимости проверки (при использовании дорогого оборудования или сложной технологии), а также не технико-экономическое уменьшение периодичности в целях снижения вероятности пропуска аварийных отказов может привести к такой ситуации, когда в затратном отношении 3-я стратегия окажется невыгодной. Чтобы этого не произошло, необходимо на начальной стадии ее введения иметь показатели по  $c_d$  и  $L_d$ , гарантирующие ее эффективность.

Эти показатели могут быть рассчитаны из условия

$$c_{уд}^{II} \geq c_{уд}^{III}. \quad (7.5)$$

При расчетах удобно перейти от конкретных к универсальным относительным оценкам, в частности определять значения так называемого коэффициента максимально допустимой стоимости единичного диагностирования  $\epsilon_{max}$ , равного отношению  $c_{dmax}/d$  для конкретных значений коэффициентов опасности отказов  $k$  и вариации рассеивания ресурса  $v$ , и соответствующего ему из условия оптимизации (7.4) значения коэффициентов относительной максимальной периодичности проведения диагностирования  $\mu_{max}$ , равного отношению  $L_{dmax}/L_{cp}$ .

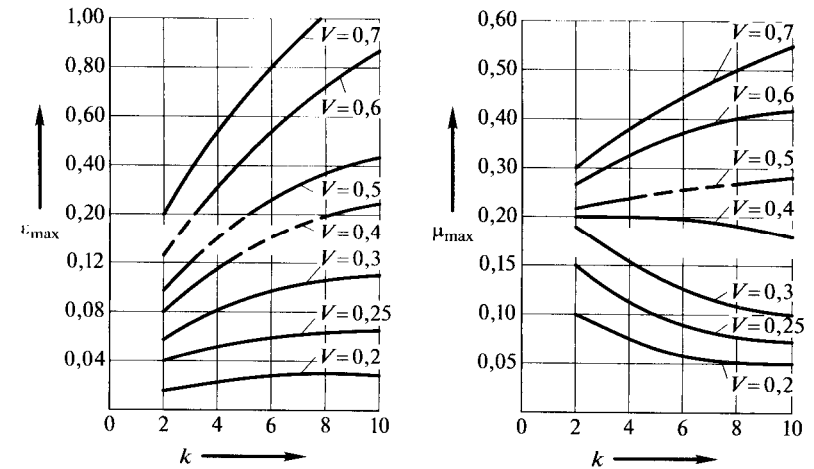


Рис. 7.6. Значения коэффициентов максимально допустимых величин стоимости единичного диагностирования  $\epsilon_{max}$  и периодичности его проведения  $\mu_{max}$  в зависимости от вариации рассеивания долговечности  $V$  контролируемых механизмов и коэффициента опасности их отказов  $k$

Удобные для практического использования номограммы, отражающие результаты такого расчета, приведены на рис. 7.6.

Из них можно видеть, что для наиболее характерных вариаций рассеивания ресурса  $v = 0,4 \dots 0,5$  в широком диапазоне значений коэффициента опасности отказов  $k$  величина  $\mu_{max}$  практически стабилизируется на уровне 0,2, т.е. периодичность диагностирования не должна превышать 1/5 части среднего ресурса. Данное значение удобно использовать при планировании работ по организации диагностики. Также величина  $\epsilon_{max}$  может применяться для подбора соответствующего диагностического оборудования как по его стоимости, так и по технологичности выполнения операций (дорогое и сложное оборудование, как правило, связано с незначительными затратами времени на проведение диагностики).

Дополнительным преимуществом 3-й стратегии является то обстоятельство, что при оптимизации ее режимов (периодичности контроля) суммарное количество аварийных отказов получается примерно в два раза ниже по сравнению со 2-й стратегией, что существенно для социально значимых элементов автомобиля. Также она позволяет находить рациональные режимы в случае экспоненциального распределения моментов появления отказов или неисправностей, если неисправное состояние сохраняется определенный период и могут быть оценены его экономические последствия (для практических расчетов в этом случае достаточно использовать расчетные значения номограмм рис. 7.6 при коэффициенте вариации  $v = 0,7$ ).

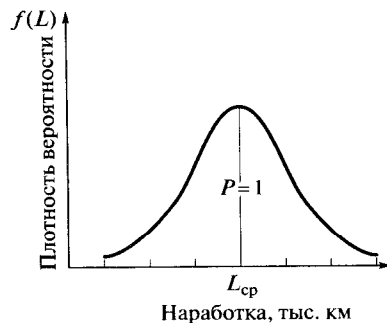


Рис. 7.7. Графическая интерпретация 4-й стратегии обслуживания и ремонта — по фактическому состоянию, выявляемого на основе встроенного (непрерывного) диагностирования — для отдельного элемента автомобиля, изменение состояния которого связано с изнашиванием или постепенным накоплением неисправностей

Однако наиболее важным является то обстоятельство, что только в рамках 3-й стратегии теоретически можно перейти к перспективной одноступенчатой системе предупредительного обслуживания и ремонта на основе принудительного диагностирования, добиваясь за счет незначительной стоимости диагностирования долговечных узлов и агрегатов снижения их рациональных режимов до значений, соответствующих наименее надежным и критическим для технического состояния автомобиля в целом и его элементов (тормозной системы или элементов, влияющих на экологическую безопасность).

Недостатками 3-й стратегии являются: потребность в дорогом и сложном диагностическом оборудовании, преимущественное использование в рамках крупных автотранспортных организаций или кооперативной группы более мелких предприятий с существенно более высоким уровнем организации управления (как элемент более качественной технологии обслуживания и ремонта стационарная диагностика не показала особых преимуществ и даже отвергалась вообще). Недостатком является также невозможность полного исключения аварийных отказов для социально значимых узлов и систем автомобиля, хотя на практике технологические и организационные недостатки 2-й стратегии приводят к более худшим показателям.

В связи с этим потенциальное значение будет иметь так называемая идеальная, или 4-я, стратегия, основанная на непрерывном контроле технического состояния, осуществляемого в процессе движения бортовыми или встроенными в автомобиль системами диагностирования. Графическая интерпретация данной стратегии приведена на рис. 7.7. Здесь стоимость диагностирования можно считать равной нулю, поскольку отпадает необходимость во внешнем стационарном комплексе, а за счет непрерывности контроля исключается

пропуск аварийных отказов и обеспечивается полная реализация индивидуально ресурса, заложенного заводом-изготовителем.

Удельные затраты на поддержание работоспособности при этом имеют самое минимальное значение и определяются соотношением

$$c_{уд}^{IV} = \frac{d}{L_{ср}}. \quad (7.6)$$

Данная стратегия является единственной стратегией, которая теоретически может решить социальную проблему исключения в эксплуатации аварийных отказов по системам, влияющим на безопасность движения и загрязнения окружающей среды. Кроме того, бортовая (встроенная) диагностика на перспективу может являться единственным доступным видом подобного контроля технического состояния для легковых автомобилей индивидуального пользования и грузовых автомобилей и автобусов, принадлежащих частным лицам.

Хотя интуитивно достоинства подобной стратегии были очевидны с самого начала появления автомобилей, ее реализация практически стала возможной только в эпоху научно-технического прогресса, благодаря достижениям электроники и электрических систем измерения.

При этом сразу же стал очевиден такой ее существенный недостаток, как необходимость большого количества разнообразных по физическим принципам работы встроенных систем диагностирования для всех критических в отношении эксплуатационной надежности элементов автомобиля (более 100), что могло сделать данную стратегию нерациональной для реализации даже на длительную перспективу. Высокие рекламные оценки возможности снижения затрат в сфере технической эксплуатации при переходе от 1-й к 4-й стратегии, который может быть потенциально осуществлен для всего автомобильного парка (в 7—10 раз для отдельных объектов, что очевидно при сравнении выражений (7.2) и (7.6)), обусловили на Западе с середины 70-х годов XX в. интенсивные работы по созданию систем встроенного диагностирования, которые реализованы в последних моделях автомобилей.

Однако анализ особенностей этих автомобилей показывает, что весь еще далеко до широкой реализации принципов 4-й стратегии, а их техническая эксплуатация, наоборот, стала более сложной и дорогой. Это свидетельствует о недостаточной перспективности ранее использованных простейших идей в области встроенного диагностирования и необходимости привлечения более сложных методологических подходов для решения данной проблемы, в частности, рассмотренных ранее системного подхода и системного анализа, а также комплекса оценок, полученных на основе имитационного моделирования.

### 7.3. Основные организационно-технологические варианты обслуживания и ремонта автомобилей на АТП общего пользования, подлежащие модельному исследованию

Сформулируем основные организационно-технологические варианты обслуживания и ремонта автомобилей в АТП общего пользования, которые необходимо промоделировать для выявления путей рационального использования диагностирования.

*Первый вариант* (рис. 7.8) — выполнение ремонтов автомобилей по потребности, без проведения плановых обслуживания и диагностирования. Этот вариант отражает наиболее простую, естественно складывающуюся стратегию поддержания работоспособности одиночного автомобиля индивидуального пользования и имеет место также для небольших групп автомобилей акционерных коллективов сельского хозяйства и ведомственных автотранспортных организаций. Данная стратегия приводит к наиболее плохим технико-экономическим показателям, однако несоблюдение плана и нарушение регламентной технологии обслуживания, обычно наблюдающиеся в той или иной степени даже на крупных АТП общего пользования, приводит к выходным эксплуатационным показателям, характерным для этого простейшего варианта.

Особенностями данного варианта являются отсутствие поточного метода обслуживания (при наличии поточных линий с осмотровыми канавами и последним постом для механизированного выполнения смазочно-очистительных работ, в соответствии со старыми типовыми проектами, они используются в качестве постов ТР тупикового типа, хотя при этом возможно блокирование автомобиля, стояще-

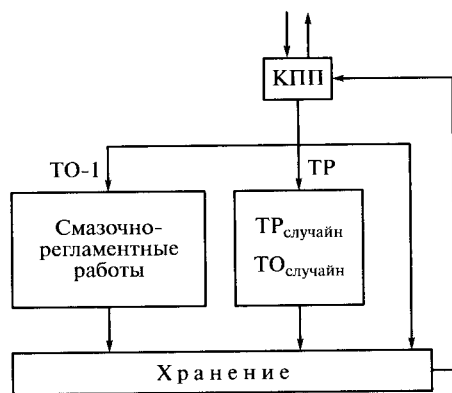


Рис. 7.8. Схема 1-го простейшего варианта обслуживания и ремонта автомобилей на АТП общего пользования

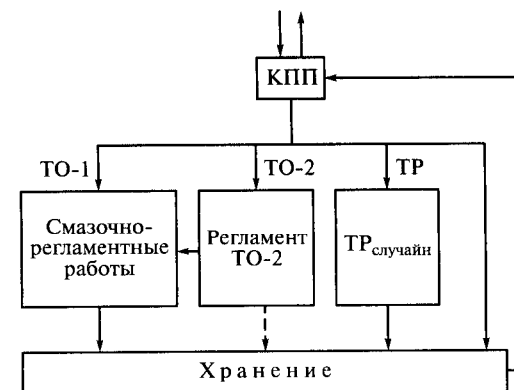


Рис. 7.9. Схема 2-го варианта организации обслуживания и ремонта автомобилей на АТП общего пользования, базирующегося на планово-предупредительном выполнении работ ТО-1 и ТО-2 [8]

го на посту в середине линии) и проведение всех видов работ на универсальных постах. Другой характерной особенностью как 1-го, так и остальных вариантов является выполнение при ТО-1 не всего комплекса работ, предусмотренного ранее рекомендованными типовыми документами, а только его части (около 15...20%), составляющих регламент смазочно-очистительных операций.

Моделирование 1-го варианта помимо теоретического изучения его возможностей по комплексу отмеченных ранее выходных показателей также необходимо для оценки последствий от использования средств встроенного диагностирования на автомобилях, эксплуатирующихся вне крупных автотранспортных предприятий общего или ведомственного пользования. Как уже отмечалось ранее, именно такие автомобили составляют основу автопарка страны.

*Второй вариант* (рис. 7.9) — двухступенчатое плановое обслуживание (ТО-1, ТО-2) без применения средств диагностирования и с проведением ремонта основных агрегатов, не влияющих на безопасность движения, по потребности, как в 1-м варианте [3, 8]. Второй вариант и в настоящее время повсеместно рекомендуется ведомственными нормативными документами к обязательному использованию для всех видов государственного автотранспорта. Для крупных АТП работы по ТО-1 рекомендуются выполнять на 3—4-постовой поточной линии в межсменное время. При сложившейся организации даже для крупных АТП основные работы ТО-2 выполняются на тупиковых постах, а завершающие смазочные работы осуществляются на специализированном посту поточной линии ТО-1.

В целях повышения эксплуатационной надежности автомобилей при каждом ТО-2 рекомендуется помимо тормозной системы и рулевого управления проводить также ревизию технического состояния

карбюратора, прерывателя-распределителя, свечей, генератора и реле-регулятора, которая может быть выполнена в цеховых условиях АТП при демонтаже указанных устройств с автомобиля. Однако в связи с тем, что безотказность этих устройств в последнее время существенно повысилась, а вероятность внесения дополнительных дефектов в процессе демонтажно-монтажных работ остается значительной, указанные рекомендации в настоящее время не выполняются и ремонт и обслуживание систем питания, зажигания и электрооборудования производится по потребности при обнаружении их явных разрегулировок или отказов. Специальных видов обслуживания, направленных на повышение экономичности и снижение загрязнения окружающей среды, профилактическая система ТО-1, ТО-2 не предусматривает.

Методическое значение моделирования 2-го варианта помимо общей сравнительной оценки эффективности от использования диагностирования заключается также в возможности проверки адекватности модели в отношении трудовых и временных затрат на обслуживание и текущий ремонт, которые достаточно хорошо нормированы для всех моделей автомобилей.

*Третий вариант* (рис. 7.10) — двухступенчатое плановое обслуживание (ТО-1, ТО-2), дополненное стационарным диагностированием Д-1 и Д-2 [6]. Подобная организация диагностирования ранее рекомендовалась нормативными документами для автотранспортных предприятий общего пользования, хотя при этом предусматривались различные технологические варианты — от централизованных постов

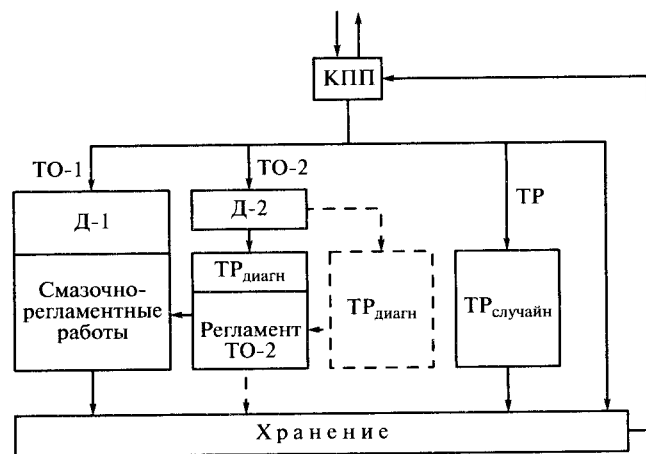


Рис. 7.10. Схема 3-го варианта организации обслуживания и ремонта автомобилей на АТП общего пользования, базирующегося на планово-предупредительном выполнении работ ТО-1 и ТО-2 с предшествующим контролем технического состояния стационарным диагностированием Д-1 и Д-2 [6]

диагностирования на основе комбинированных механических стендов до поточных линий диагностирования, выделенных или совмещенных с процессами ТО-1 и ТО-2.

Принципиальной особенностью данного варианта, как это можно видеть из схемы, представленной на рис. 7.10, является, в первую очередь, разделение объемов текущего ремонта, ранее выполняемого индивидуально непредсказуемым (случайным) образом, на две части:

- ТР<sub>диагн</sub>, планируемый для базовых агрегатов автомобиля по результатам диагностирования Д-2, предшествующего ТО-2 (для этого первое рекомендовалось проводить заблаговременно, чтобы подготовить производство для эффективного выполнения выявленных объемов ТР);
- работы случайного характера ТР<sub>случайн</sub>, не связанные с диагностированием.

Диагностирование Д-1 при этом должно было осуществлять выявление (с последующим устранением) возникающих до постановки в плановое ТО-2 неисправностей или скрытых отказов тормозов, рулевого управления, систем питания и зажигания, влияющих на загрязнение окружающей среды (теоретически они могли иметь место примерно у 20 % автомобилей). Работы ТР<sub>диагн</sub> как можно видеть из схемы, представленной на рис. 7.10, могли выполняться совместно с регламентными работами ТО-2, хотя для крупных АТП для этого рекомендовались специализированные посты в зоне ТР. Для завершения ТО-2 автомобиль должен был поступить на механизированный пост смазочно-очистительных работ линии ТО-1.

За основу при моделировании был принят самый сложный, но наиболее экономичный для крупных АТП вариант организации, разработанный в МАДИ и предусматривающий максимальное совмещение профилактических и диагностических операций Д-1 и Д-2 (при размещении постов диагностики на линии ТО-1), сопровождающийся существенными изменениями технологического процесса, при сохранении принципов поточного метода обслуживания на специализированных постах и участках.

Данный вариант размещения стационарного диагностического оборудования на поточной линии с обзорной канавой на примере легковых автомобилей с передним приводом показан на рис. 7.11 (в настоящее время его целесообразно реализовывать для заднеприводных автомобилей и грузовых автомобилей и микроавтобусов на их базе типа «Газель», «Соболь» для организационно-производственных структур типа централизованного специализированного производства (ЦСП); при этом расстояние между стендами для проверки тормозов и тягово-экономических показателей целесообразно увеличить для лучшей синхронизации работ постов поточной линии).

Пост для проверки тягово-экономических показателей должен быть оборудован системой отсоса выхлопных газов, обычно использующей гофрированный гибкий шланг, подключенный к подвешен-

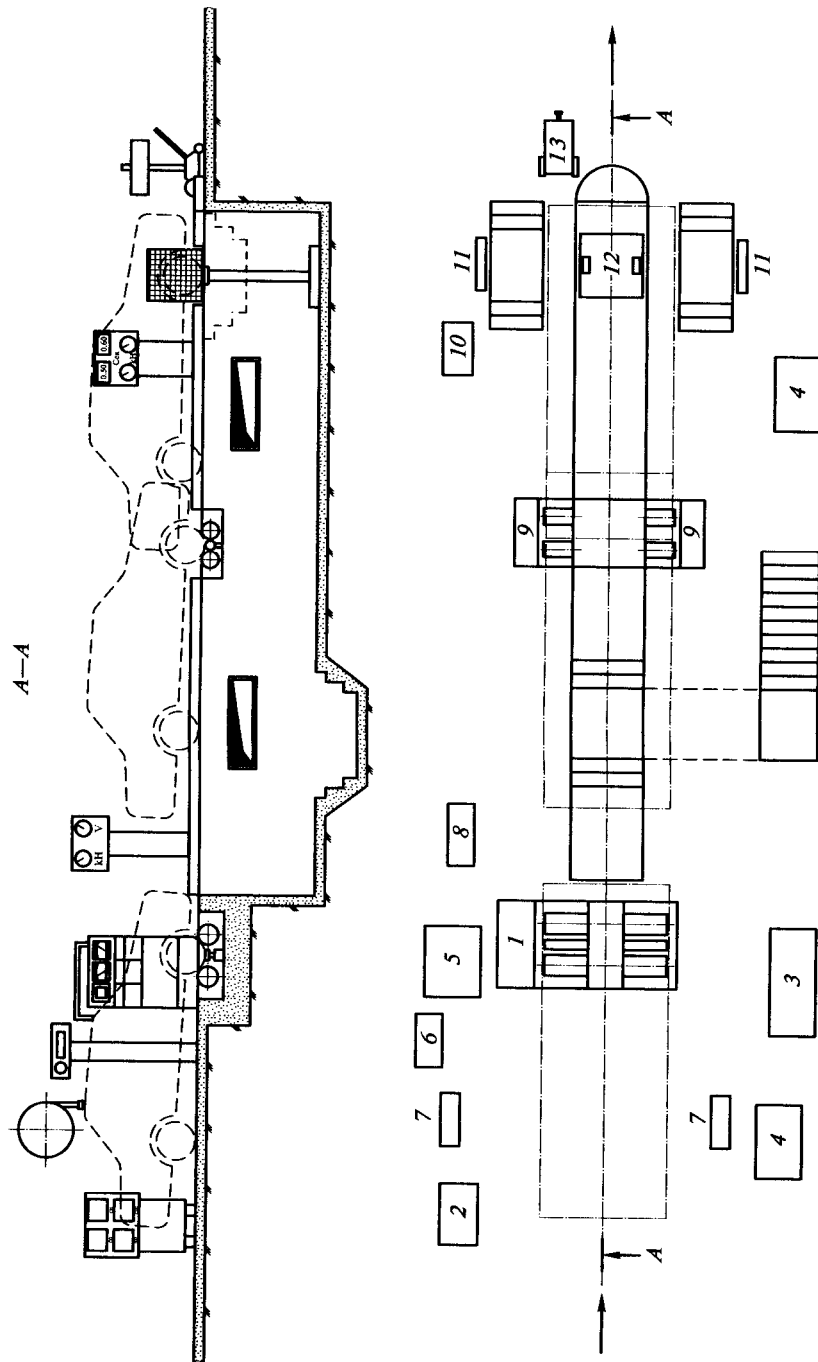


Рис. 7.11. Вариант размещения стационарного диагностического оборудования на поточной линии с осмотровой канавой на примере легковых автомобилей с передним приводом:

1 — динамометрический стенд для проверки тягово-экономических показателей; 2 — пятикомпонентный газоанализатор; 3 — верстак слесарный; 4 — шкаф-стеллаж для хранения переносного диагностического оборудования и запасных частей; 5 — мотор-тестер; 6 — автоматическая воздухоподаточная колонка; 7 — барабаны с самонаматывающимися шлангами для подкачки шин; 8 — пульт управления динамометрическим стендом; 9 — стенд для проверки тормозных сил колес одной оси и времени срабатывания тормозов, с отдельными каретками силовых роликов, удобный для монтажа на осмотровой канаве; 10 — пульт управления тормозным стендом; 11 — лазерно-оптический стенд для углов установки управляемых колес; 12 — двухстоечный передвижной электромеханический подъемник, не перекрывающий размещение операторов по осмотровой канаве; 13 — прибор для проверки фар

ному под потолком вытяжному вентилятору, или другие удобные конструкции (на рис. 7.11 не приведена).

Удобство данного варианта заключается в возможности совмещения диагностирования с операциями ТО, направленными на устранение выявляемых при диагностике неисправностей, а также сосредоточения диагностического оборудования в одном месте (стенд тягово-экономических показателей, составляющий основу поста Д-2, может быть расположен на отдельном посту тупикового типа).

Подобная поточная линия в начале 1980-х годов была распространена на станциях технического обслуживания «Сельхозтехники» и автотранспортных предприятиях общего и ведомственного пользования для одиночных грузовых автомобилей с бензиновыми двигателями типа ГАЗ и ЗиЛ (к настоящему времени указанное оборудование уже практически вышло из строя и не используется). Она имела те же специализированные посты, что и на рис. 7.11, но другие типоразмеры механических стендов. Кроме того, на 3-м посту ранее устанавливался стенд для проверки схождения колес по боковой силе (в настоящее время он может быть модернизирован за счет использования ненагруженных беговых барабанов, о чем будет сказано далее) и при помощи других приборов и устройств контролировалось давление в шинах, состояние рулевого механизма, рулевых тяг и шкворневых втулок переднего моста. Кроме того, типовой проект предусматривал расположение на 1-м посту стенда для проверки схождения, а на 3-м посту — стенд для проверки тягово-экономических показателей. Указанная последовательность размещения механических стендов была принята за основу при проведении моделирования.

Хотя 3-й вариант организации к моменту модельного исследования более 10 лет в той или иной степени апробировался на практике, его теоретические возможности до конца не были определены. Главная причина этого заключалась в отсутствии методов и средств диагностирования, полностью удовлетворяющих потребности технологиче-



ского процесса (согласно выполненным оценкам определение объемов текущего ремонта современными средствами диагностирования может быть осуществлено примерно наполовину от возможного). Поэтому полное выявление показателей 3-го варианта представляется возможным только на основе моделирования.

**Четвертый вариант** (рис. 7.12) — одноступенчатое плановое обслуживание на основе полнокомплектного диагностирования (ЕТОД). Его периодичность в первом приближении равнялась половине пробега до ТО-2. Данный вариант по результатам системного анализа был признан наиболее предпочтительным для автотранспортных предприятий общего пользования и получил определенное признание за рубежом. Так, намеченная в США на 1980-е годы программа совершенствования технического обслуживания грузовых автомобилей также основывалась на одной плановой ступени с предварительным диагностированием и обработкой информации.

Для моделирования была принята технологическая схема проведения ЕТОД на 4-постовой поточной линии, как и в случае 3-го варианта. Четвертый пост специализирован для выполнения очистительно-смазочных работ. Обслуживание и ремонт всех элементов и узлов автомобиля, подвергающихся диагностированию, при ЕТОД целесообразно проводить на постах тупикового типа зоны текущего ремонта по потребности (по состоянию).

Информационно-технологические возможности диагностирования при моделировании ЕТОД были приняты такими же, как и для 3-го варианта. Трудоемкость технологического процесса поточной

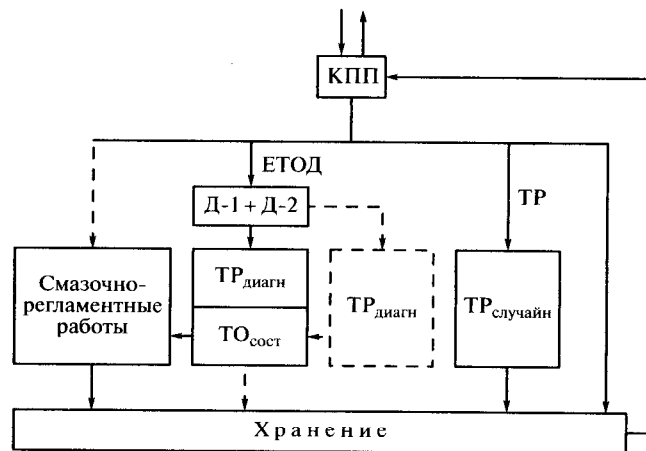


Рис. 7.12. Схема 4-го варианта организации обслуживания и ремонта автомобилей на АТП общего пользования: «Единое (одноступенчатое) техническое обслуживание на основе регламента комплексного стационарного диагностирования (Д-1 + Д-2)» — ЕТОД

линии ЕТОД с распределением работ по постам была рассчитана путем корректировки уже рекомендуемой технологии «Сельхозтехники» и уточнена экспериментально. Параллельно с моделированием могли быть осуществлены экспериментальные исследования возможностей ЕТОД, проведение сопоставление полученных теоретических и практических результатов.

При переходе к большегрузным дизельным автомобилям и автопоездам, а также сочлененным автобусам расположение механических стендов диагностирования на поточной линии уже встречало значительные затруднения (длина линии увеличивалась до 90... 100 м, а удаление выхлопных газов системой отсоса при перемещении автомобиля на такие расстояния требовало постоянной перестановки гибких шлангов). Исходя из принципов системного подхода данная проблема могла быть решена на основе создания однопостового участка комплексного диагностирования, перемещение по которому ограничивалось бы только технологическими операциями по последовательной проверке тормозов всех мостов автомобиля или автопоезда и не превышала бы 30... 36 м при независимом расположении и 48 м в сочетании с постом для выполнения смазочно-очистительных работ. Пример такого поста показан на рис. 7.13.

Принципиальным отличием данного участка от описанных ранее является отсутствие динамометрического стенда и связанного с ним комплекса нагрузочных испытаний автомобиля (для имитации дорожного движения здесь использованы простые по конструкции ненагруженные беговые барабаны, которые могут быть изготовлены силами даже небольшого предприятия). Это существенно упрощает технологический процесс, а теряемая при этом информативность может быть компенсирована применением других методов и средств (для бензинового автомобиля, как это было показано в подразд. 3.4, объем информации даже возрастает за счет применения пятикомпонентного газоанализатора; для дизеля такие испытания нетехнологичны и неинформативны).

Для облегчения скоростных испытаний, на которых можно проверить вибрацию и шумы агрегатов трансмиссии, не различаемые на дороге, а также легкость переключения передач и (в перспективе) экологические показатели, третий мост автомобилей типа КамАЗ может быть вывешен передвижным двухстоечным подъемником. Другой функцией данного подъемника является проверка шкворневых втулок переднего моста и механизма рулевого управления при вывешивании колес (до установки автомобиля на стенд). При установке передних колес на ненагруженные беговые барабаны затем можно проверить и отрегулировать их сходжение при помощи линейки, поворачивая ручным приводом колеса через барабаны вместе с линейкой без перекатывания автомобиля.

Однако в связи с неизбежными в эксплуатации деформациями балок переднего моста, приводящими к их изгибу и скручиванию,



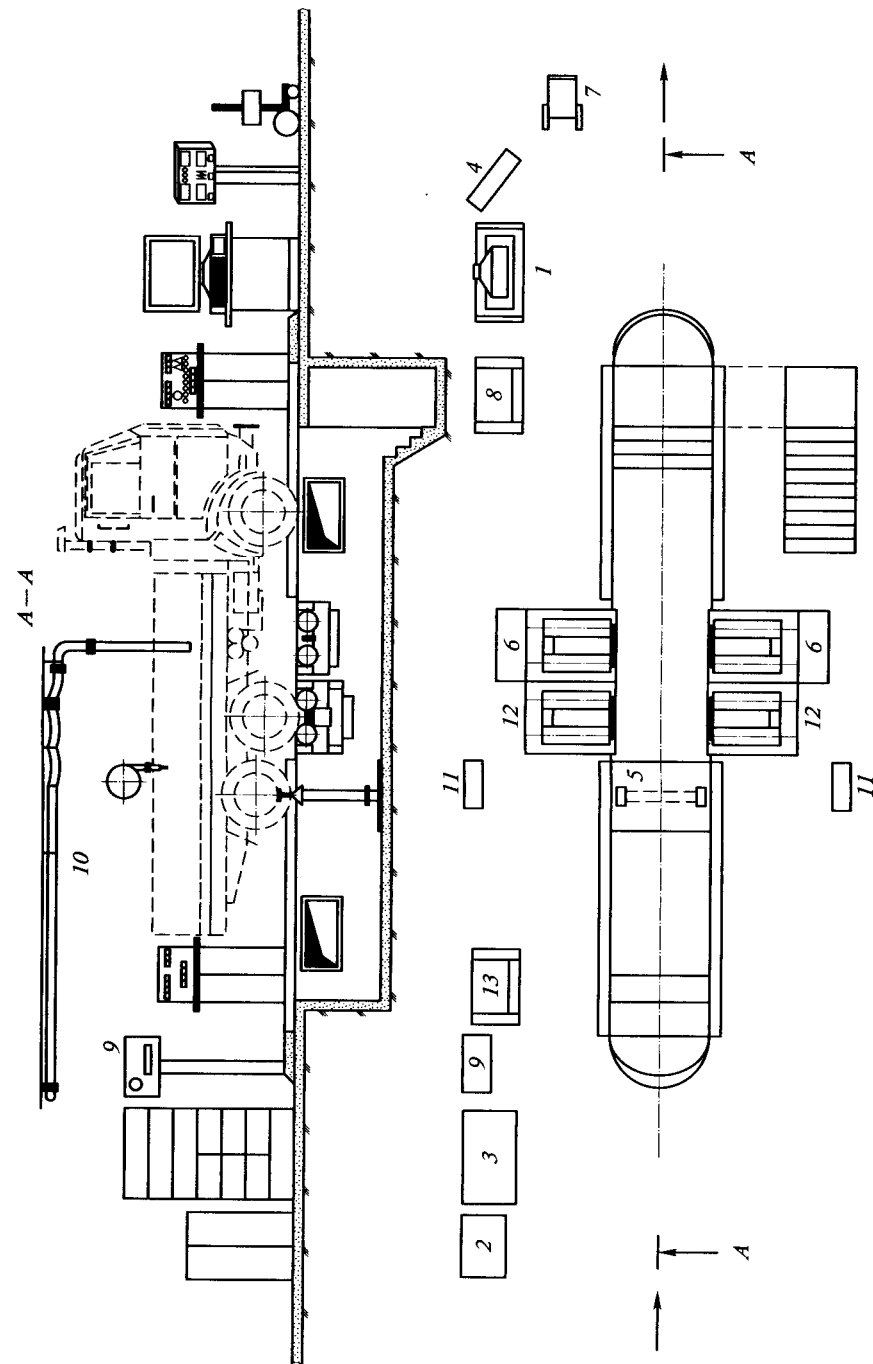


Рис. 7.13. Вариант организации однопостового участка комплексного диагностирования:

1 — автоматизированное рабочее место оператора-диагноста; 2 — шкаф для хранения инструмента и переносного диагностического оборудования; 3 — стеллаж для ремонтируемых или обслуживаемых деталей и небольших агрегатов, временно снимаемых с автомобиля; 4 — пульт управления стендом для проверки тормозов; 5 — передвижной канавный двухстоечный электромеханический подъемник, не перекрывающий перемещения операторов по осмотровой канаве; 6 — силовой роликовый стенд для проверки тормозных сил на колесах одной оси и времени срабатывания тормозов; 7 — прибор для проверки и регулировки фар; 8 — переносной дизель-тестер на передвижной тумбочке; 9 — автоматическая воздухоподкачивающая колонка; 10 — система отсоса выхлопных газов с гибким гофрированным шлангом, перемещающимся вместе с автомобилем; 11 — барабаны с самонаматывающимися шлангами для подкачки шин; 12 — ненагруженные беговые барабаны; 13 — переносной дымомер на передвижной тумбочке

предварительно необходимо проверить значения углов развала (и по возможности продольного наклона шкворней), используя навесные переносные приборы, в том числе и самодельного типа [3], для чего также необходимо повернуть колеса на барабанах на  $180^\circ$ . Если деформация балки находится в допустимых пределах, то необходимая для данного автомобиля величина схождения определяется по номограмме, представленной на рис. 7.14. Указанная технология заменяет метод регулировки схождения по боковой силе и является более эффективной, так как осуществляется контроль деформации балки.

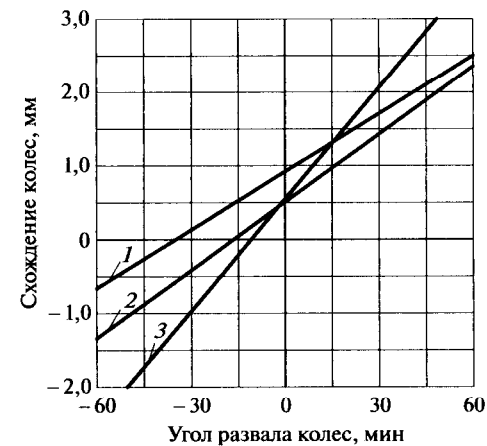


Рис. 7.14. Номограмма для определения схождения передних колес грузовых автомобилей в зависимости от средней величины угла развала:

1 — автомобили Маз, КраЗ; 2 — автомобили ЗиЛ; 3 — автомобили ГАЗ, ЛиАЗ, КамАЗ

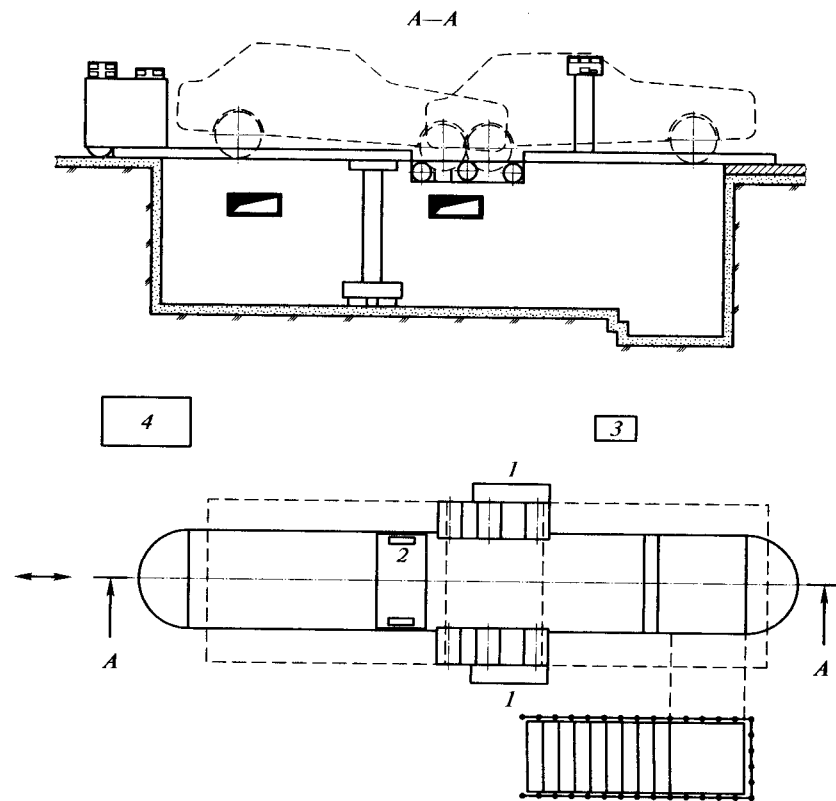


Рис. 7.15. Пост комплексного диагностирования легковых автомобилей простейшего типа:

1 — комбинированный стенд с ненагруженными беговыми барабанами и силовым роликом для проверки времени срабатывания тормозов простейшего типа; 2 — подъемник канавный двухстоечный передвижной; 3 — пульт управления комбинированного стенда; 4 — тумбочка передвижная с пятикомпонентным газоанализатором и переносным мотор-тестером

Другой вариант участка диагностирования на основе комбинированного стенда простейшей конструкции для легковых автомобилей приведен на рис. 7.15 с соответствующей номограммой (рис. 7.16). Подобные комбинированные стенды можно изготавливать на кооперативной основе не только для легковых, но и для грузовых автомобилей, используя для привода переднего силового ролика червячный редуктор; при этом стоимость стационарного оборудования резко снижается, что существенно для малых предприятий. При проверке тормозов здесь решающее значение будет иметь параметр времени их срабатывания (для неработающих тормозов он автоматически не определяется), нарушение нормативов которого вызывает заносы

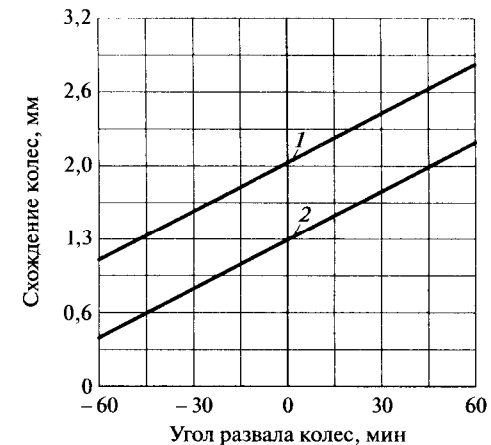


Рис. 7.16. Номограмма для определения схождения передних колес легковых автомобилей в зависимости от средней величины угла развала [3]:

1 — автомобили ГАЗ-24-31, «Газель», «Соболь»; 2 — заднеприводные автомобили ВАЗ, «Москвич», ИЖ

автомобиля при торможении на скользкой дороге и «складывание» автопоездов.

При необходимости для автомобилей небольшой массы («Газель», «Соболь») тормозная сила может быть проверена в статике, путем измерения усилия проворачивания заторможенного колеса от барабанов через закрепленный рычаг длиной 1,57 м (10-кратное значение радиуса барабана). При этом усилие на конце рычага будет не выше 200...300 Н и может быть реализовано вручную. В крайнем варианте для технологического процесса можно использовать просто ненагруженные беговые барабаны, достаточные для выявления не тормозящих колес и применения других методов диагностирования.

Более рациональным и системно вписывающимся в сложившуюся практику следует считать двухступенчатый вариант ЕТОД (рис. 7.17), при котором на первой ступени проводится обслуживание по состоянию на основе диагностирования Д-1, а на второй, примерно совпадающей со ступенью ТО-2, на которой обычно осуществлялась замена моторного масла и фильтров, — на основе полнокомплектного диагностирования (Д-1 + Д-2). Указанные виды диагностирования целесообразно выполнять на одном участке по разным технологиям, что не сложно в организационном отношении. Общая трудоемкость диагностирования при этом может быть существенно сокращена.

*Пятый вариант* (рис. 7.18) — одноступенчатое плановое обслуживание на основе стационарного диагностирования (ЕТОД) для автомобилей, имеющих оперативно встраиваемые при эксплуатации

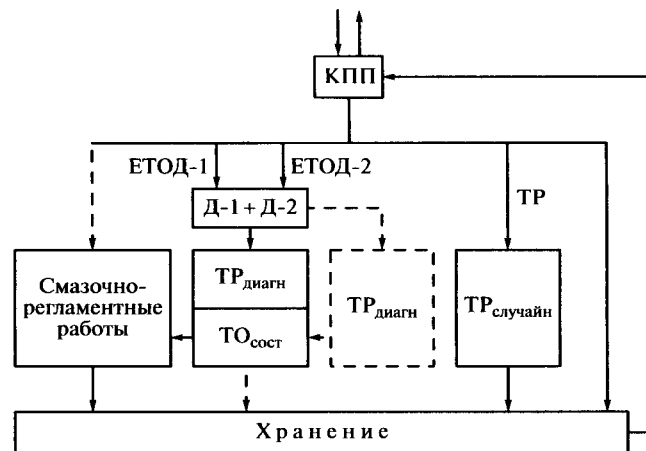


Рис. 7.17. Схема рационального 4-го варианта организации обслуживания и ремонта автомобилей на основе ЕТОД на АТП общего пользования, предусматривающая две степени регламентного диагностирования: Д-1 для ЕТОД-1 (с периодичностью, вдвое меньшей периодичности ТО-2); (Д-1 + Д-2) для ЕТОД-2 (с периодичностью ТО-2 при планировании работ по замене моторного масла)

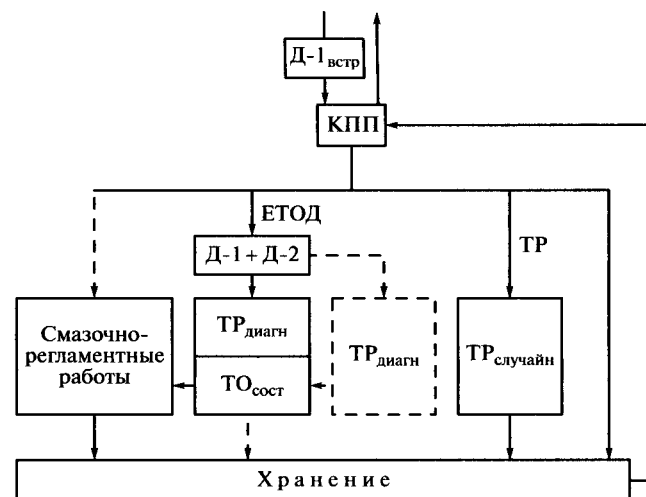


Рис. 7.18. Схема 5-го варианта организации обслуживания и ремонта автомобилей на основе ЕТОД на АТП общего пользования при дополнительном использовании возможностей оперативно встраиваемых в автомобиль средств диагностирования для непрерывного контроля тормозов и состояния систем, влияющих на загрязнение воздушного бассейна

устройства для диагностирования систем, влияющих на расход топлива, загрязнение окружающей среды и безопасность движения (в данном случае систем зажигания, питания и тормозов). При этом мелкие неисправности, выявляемые встроенным диагностированием, в основном зажигания и холостого хода, а также разрегулировки тормозов, могут устраняться силами водителей (в рамках ежедневного обслуживания) в целях сокращения количества внеплановых выездов в стационарный комплекс ЕТОД.

Поскольку информационные возможности стационарного диагностирования значительно выше по сравнению со встроенным, целесообразно сохранение плановых заездов и регламента операций ЕТОД, хотя возможна большая периодичность по сравнению с 4-м вариантом. При обнаружении встроенной системой неустранимых неисправностей тормозов автомобиль должен поступать в стационар ЕТОД ранее планируемого срока. По результатам системного анализа подобное сочетание стационарного и встроенного диагностирования признано наиболее перспективным для всего автомобильного транспорта. Значимость результатов моделирования 5-го варианта может быть повышена при конкретизации возможностей встроенного диагностирования на основе предварительной разработки методов и макетов этих средств.

*Шестой вариант* — одноступенчатое внеплановое обслуживание на основе стационарного диагностирования для автомобилей, имеющих встроенное диагностирование систем зажигания, питания и тормозов. В 6-м варианте, в отличие от 5-го, автомобиль будет проходить регламентное обслуживание на линии ЕТОД при выявлении встроенным диагностированием неустранимых неисправностей тормозов. Средняя периодичность обслуживания при этом будет увеличена, по сравнению с 5-м вариантом, однако можно предполагать увеличение затрат на ремонты основных агрегатов за счет их несвоевременного выявления средствами стационарного диагностирования.

Несмотря на возможные худшие показатели 6-го варианта (по сравнению с 5-м) последний может иметь преимущественное распространение для автомобилей акционерных коллективов сельского хозяйства, небольших ведомственных автотранспортных организаций и индивидуальных владельцев, сложные виды обслуживания и ремонта которых могут выполняться централизованными специализированными предприятиями.

*Седьмой вариант* (рис. 7.19) — выполнение ремонтов по потребности для автомобилей, имеющих встроенное диагностирование систем питания, зажигания и тормозов. Встроенное диагностирование делает излишним проведение профилактических плановых обслуживания вида ТО-1, ТО-2, направленных в основном на поддержание надежности тормозов и рулевого управления, однако систематическое применение более сложных видов обслуживания на основе стандо-

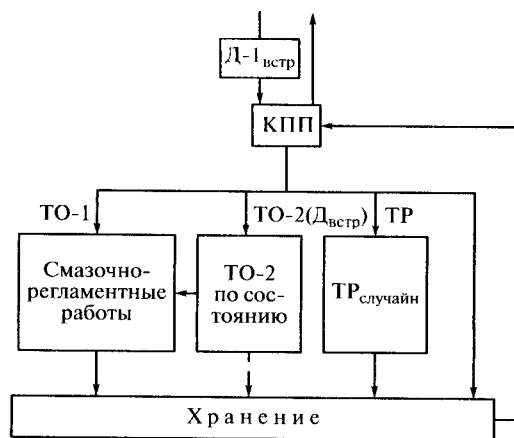


Рис. 7.19. Схема 7-го варианта организации обслуживания и ремонта автомобилей на АТП общего пользования, не имеющего стационарного диагностирования, при дополнительном использовании возможностей оперативного встраиваемых в автомобиль средств для непрерывного контроля тормозов и состояния систем, влияющих на загрязнение воздушного бассейна

вого диагностирования для многих автомобилей, работающих в отрыве от баз, будет при этом невозможным или затруднительным. Как уже отмечалось ранее, таких автомобилей при переходе к рыночной экономике станет подавляющее большинство и применение средств встроенного диагностирования является единственным способом совершенствования их технической эксплуатации.

Таким образом, сформулированные варианты охватывают широкий диапазон технической эксплуатации современных и перспективных автомобилей, а результаты моделирования на частном примере модели автомобиля и условий АТП общего пользования могут иметь более широкое теоретическое значение для понимания процессов поддержания работоспособности автомобилей в целом.

#### 7.4. Результаты моделирования основных вариантов организации ТО и ремонта на основе использования стационарного и встроенного диагностирования на автотранспортных предприятиях общего пользования

Для всех пяти базовых вариантов использования стационарного и встроенного диагностирования, сформулированных ранее, действует принцип планового предупреждения отказов, без которого

невозможно обеспечение необходимого уровня эксплуатационной надежности; варианты отличаются только полнотой, глубиной и режимами планового диагностирования агрегатов, узлов и систем автомобиля. Кроме того, в эксплуатации (в зависимости от ее конкретных условий) одновременно могут иметь место несколько базовых вариантов организации с различной степенью реализации. Поэтому оценка на основе моделирования «в чистом виде» каждого из них важна не только с позиций оптимизации, но и в целях выявления и учета возможных последствий от неполноты применения наиболее предпочтительного варианта.

Наиболее сложным и решающим этапом данных исследований является разработка программы компьютерного моделирования. К сожалению, данное направление еще не освоено несмотря на его актуальность. Пока здесь просматриваются два подхода. *Первый подход* обуславливает создание универсальной программы для исследований стратегического характера с дружественным интерфейсом, понятным рядовому потребителю, например начальнику производственно-технического отдела (ПТО) АТП или техническому руководителю небольшой автотранспортной организации. Предварительные опыты показали, что это возможно осуществить на языке программирования типа «С++» в среде WINDOWS.

*Второй подход* может быть связан с использованием специализированного языка программирования для имитационного моделирования вида GPSS, приспособленного для персональных компьютеров. Ранее данный язык широко применялся для ЭВМ серии ЕС с оперативной памятью 4 Мбайт и позволял моделировать совместно специфические особенности как коммерческой, так и технической эксплуатации всех автомобилей крупного АТП. К сожалению, первый вариант языка GPSS для ПЭВМ типа «Х-тишки» и выше обладал значительно более скромными возможностями, а главное, работал только в DOS, а оперативная память для моделирования не превышала 300 Кбайт. Хотя при этом возможно было решать достаточно сложные задачи оперативного управления при условии хорошего владения языком, однако применительно к задачам моделирования работы АТП получить качественные результаты оказалось невозможным и исследования в этом направлении на протяжении порядка 20 лет не проводились.

Появление на рынке современного варианта языка GPSS для среды WINDOWS с постоянно расширяющейся оперативной памятью может обеспечить решение более сложных задач моделирования работы автомобилей, связанных с рыночной экономикой, разукрупнением АТП и появлением независимого централизованного специализированного производства (ЦСП) для выполнения ТО и ремонта. Поэтому работы по созданию подобных моделей должны проводиться с возрастающей интенсивностью. Ценность этих работ можно видеть по представленным далее результатам моделирования работы

типового АТП 20-летней давности, сохраняющим свою общеметодическую значимость до настоящего времени.

Для проведения моделирования помимо знания принципиальных возможностей ранее известных и вновь разработанных методов и средств диагностирования необходимо иметь подробное описание закономерностей изменения технического состояния всех элементов автомобиля, поддающихся диагностированию, а также показатели, характеризующие технологические процессы обслуживания и ремонта. Принятые исходные данные (частично приведены в табл. 7.1 ... 7.3) были получены на основе анализа многочисленных исследований в области надежности и диагностики бензиновых автомобилей средней грузоподъемности за длительный период. Проверка показала, что они обеспечивали соответствие конечных результатов моделирования для двухступенчатой профилактической системы ее нормативным величинам, что также подтверждало и адекватность принятой модели в целом. Соответственно была проверена и подтверждена адекватность подмоделей для описания отдельных (уже изученных экспериментально) процессов диагностирования.

Наибольшую сложность представляло моделирование вариантов организации диагностирования, поскольку неизбежно приходилось пользоваться некоторыми допущениями и предположениями в отношении перспективных методов и средств, т.е. только в первом приближении изученным на практике. В связи с этим применялись наименее выгодные в отношении диагностирования значения показателей из полученных экспериментально. В первую очередь, подобные оценки имели место в отношении повышения расхода топлива при появлении неисправностей карбюратора и системы зажигания, а также падения давления в шинах. При моделировании процессов текущего ремонта считалось, что период предотказного состояния, который может быть зафиксирован средствами диагностики, составляет 10 % от значения индивидуального ресурса, а пропуск отказа увеличивает трудоемкость (и стоимость) текущего ремонта всего на 20 %.

Недостаточная детализация по операциям текущего ремонта также приводила при моделировании к занижению оценок эффективности от использования диагностики. Учитывалось, что в общем потоке заявок на ремонт, поступающих от коробки передач, 45 % отказов являются случайными, недиагностируемыми; для главной передачи доля случайных отказов составляла 20 %. Вероятность того, что средствами диагностирования будет обнаружена скрытая неисправность без разборки агрегата, принята равной 0,97. Коэффициент вариации фактической трудоемкости устранения диагностируемых отказов принят равным 0,3.

Возможности вычислительной техники (машина ЕС 1022) позволили осуществить процесс моделирования только для АТП мощностью 100 автомобилей на периоде их работы 200 тыс. км. Производствен-

Таблица 7.1

**Параметры законов распределения пробегов до отказа основных диагностируемых агрегатов бензинового автомобиля средней грузоподъемности и средние трудоемкости устранения этих отказов при текущем ремонте при нормальном законе распределения ( $\nu = 0,3$ ), примененные при имитационном моделировании**

Агрегат или узел	Вид отказа	Закон распределения	Параметры законов		Трудоемкость ТР, чел.-ч
			$L_{\text{ср}}$ , тыс. км	$\nu$	
Двигатель (ЦПГ + подшипник)	1-й	Нормальный	150	0,24	20,5
	2-й	Вейбулла $a = 72, b = 1,32$	64	0,76	20,5
Двигатель (ГРМ)	1-й	Вейбулла $a = 75, b = 1,7$	67	0,60	7,8
	2-й				
Сцепление (ведомый диск)	1-й	Вейбулла $a = 61, b = 1,3$	54	0,77	5,0
	2-й	Вейбулла $a = 43, b = 1,6$	38	0,64	5,0
Коробка передач (в сборе)	1-й	Вейбулла $a = 72, b = 1,32$	58	0,69	2,7
	2-й	Экспоненциальный	28	0,90	2,7
Главная передача (в сборе)	1-й	Вейбулла $a = 65, b = 1,92$	58	0,73	3,9
	2-й	Вейбулла $a = 45, b = 2,2$	40	0,60	3,9
Передний мост (шкворневое соединение)	1-й	Вейбулла $a = 117, b = 2,7$	104	0,40	6,3
	2-й	Вейбулла $a = 83, b = 1,6$	74	0,65	6,3
Тормоза передние (накладки)	1-й	Нормальный	48	0,27	1,7
	2-й	Вейбулла $a = 39, b = 3,0$	35	0,37	1,7

Окончание табл. 7.1

Агрегат или узел	Вид отказа	Закон распределения	Параметры законов		Трудоемкость ТР, чел.-ч
			$L_{\text{ср}}$ , тыс. км	$\nu$	
Тормоза задние (накладки)	1-й	Вейбулла $a = 41, b = 3,0$	36	0,33	4,1
	2-й	Вейбулла $a = 83, b = 1,6$	28	0,40	4,1
Наработка на отказ случайного характера		Экспоненциальный	1,25	1,0	—
Трудоемкость устранения случайного отказа		Вейбулла $a = 1,5, b = 1,32$	$T_{\text{ср}} = 1,3$ чел.-ч	0,76	—

Таблица 7.2

**Параметры законов распределения наработок на неисправность основных диагностируемых узлов и систем бензинового автомобиля средней грузоподъемности, влияющих на безопасность движения, расход топлива и загрязнение окружающей среды, примененные при имитационном моделировании**

Узел или система	Закон распределения	Параметры законов	
		$L_{\text{ср}}$ , тыс. км	$\nu$
Тормозной механизм передних колес	Вейбулла $a = 11,2, b = 2,0$	10	0,55
Тормозной механизм задних колес	Вейбулла $a = 13,5, b = 2,0$	12	0,55
Неисправность свечи зажигания	Вейбулла $a = 7,5, b = 1,7$	6,7	0,60
Отказ свечи зажигания	Вейбулла $a = 37, b = 1,7$	33	0,60
Неисправность поплавкового механизма карбюратора	Нормальный	38	0,28
Неисправность насоса ускорителя карбюратора	Нормальный	30	0,30

Окончание табл. 7.2

Узел или система	Закон распределения	Параметры законов	
		$L_{\text{ср}}$ , тыс. км	$\nu$
Неисправность системы холостого хода	Нормальный	8	0,30
Наработка на падение давления в шинах	Вейбулла $a = 4,8, b = 1,7$	4,3	0,60
Скорость изменения боковой силы управляемых колес, $H/1000$ км	Вейбулла $a = 5,5, b = 1,7$	4,9	0,60

Таблица 7.3

**Параметры нормальных законов распределения показателей технологических процессов обслуживания зонами диагностирования бензинового автомобиля средней грузоподъемности, примененные при имитационном моделировании**

Показатель	Единица измерения	Параметры распределения		
		$X_{\text{ср}}$	$\nu$	
Среднесуточный пробег	км	200	0,26	
Пробег между ТО-1	»	2 000	0,26	
Пробег между ТО-2	»	8 000	0,15	
Трудоемкость ТО-1 с Д-1 на постах:	чел.-ч	2,7	0,30	
		№ 1	0,82	0,25
		№ 2	0,80	0,25
		№ 3	0,66	0,26
№ 4	0,50	0,28		
Трудоемкость ТО-2, из них:	»	10,8	0,25	
		Д-2	1,3	0,25
		ТО-2 постовые	6,8	0,25
		ТО-2 с Д-2	2,7	0,30
Трудоемкость ЕТОД на постах:	»	5,84	0,30	
		№ 1	1,63	0,28
		№ 2	1,69	0,28
		№ 3	1,82	0,28
№ 4	0,70	0,30		

ные возможности автотранспортного предприятия (количество рабочих, тупиковых постов текущего ремонта и обслуживания) при моделировании всех вариантов организации принимались неизменными и соответствовали действующим нормативам. Также моделирование технологических процессов ТО-1, ТО-1 с Д-1 и ЕТОД проводилось при условии четырехпостовой поточной линии, являющейся по соответствующим нормам проектирования типовой для АТП средней мощности. Технологические характеристики постов ЕТОД определялись на основе предварительных расчетов. Вероятность выполнения плана ТО-1, Д-1 с ТО-1 и ЕТОД принималась равной 0,9.

В табл. 7.4 приведены основные (выходные) результаты сравнительного моделирования сформулированных ранее базовых вариантов организации обслуживания и ремонта автомобилей как с диагностикой, так и без ее использования, которые являются наиболее существенными с позиций рассматриваемой проблемы. Методическое значение моделирования первых двух вариантов (без использования диагностирования) заключалось как в проверке адекватности модели по трудовым и временным затратам на восстановление автомобилей, так и в более полном понимании сущности возможных изменений выходных показателей технической эксплуатации, вносимых диагностированием. В то же время значимость диагностирования невозможно оценить каким-либо одним или двумя универсальными обобщающими показателями, как это обычно делалось применительно к профилактической системе, где рассматривались только техническая готовность и затраты на ее обеспечение. Как показал системный анализ, в настоящее время главенствующую роль в оценке качества технической эксплуатации приобретает совокупность показателей, имеющих социальное значение, поэтому их необходимо рассматривать в первую очередь.

Рассмотрим подробнее все выбранные 11 показателей (см. табл. 7.4) с технико-экономических и социальных позиций, одновременно поясняя особенности их оценок в процессе моделирования. В частности, вместо общепринятого коэффициента технической готовности при моделировании определялся коэффициент технического использования, учитывающий все потери линейного времени в связи с простоями автомобиля при обслуживании и ремонте, в том числе ожидание в очереди из-за занятости постов ТО и ТР, и по организационным причинам, когда вышедший из ремонта автомобиль не может быть мгновенно включен в перевозочный процесс.

По результатам моделирования коэффициент технического использования (КТИ) изменялся незначительно (не более 5%), хотя потребность в ремонте и обслуживании по отдельным вариантам изменялась вдвое (показатели 5, 6 и 7). При условии повышения надежности автомобилей в 1,5 раза разница в КТИ уже не наблюдалась (для всех вариантов составила 0,92). Это свидетельствует о решающем значении организационных факторов управления про-

Таблица 7.4  
Сравнительные оценки вариантов организации ТО и ТР для бензиновых автомобилей средней грузоподъемности на АТП общего пользования, полученные по результатам имитационного моделирования

№ п/п	Показатель	Единица измерения	Вариант								
			1-й	2-й	3-й	4-й	5-й	6-й	7-й		
1	Коэффициент технического использования	Отн. ед.	0,83	0,845	0,86	0,87	0,86	0,87	0,86	0,84	0,83
2	Перерасход топлива	%	26	26	6	5	1	2	1	2	7
3	Пробег с неисправностями системы холостого хода	»	80	80	6	8	1	2	1	2	3
4	Пробег с технологическими неисправностями тормозов	»	48	28	5	20	1	2	1	2	3
5	Стоимость обслуживания	Усл. ед. на 1000 км	0,92	3,55	3,81	1,30	1,40	1,30	1,40	1,24	0,92
6	Стоимость заявочного ремонта (постовые работы)	»	7,22	4,57	1,06	0,80	1,10	0,80	1,10	1,97	7,04
7	Стоимость предупредительного ремонта (постовые работы)	»	—	—	1,63	1,75	1,60	1,60	1,60	1,20	—
8	Стоимость перерасходоанного топлива	»	15,65	15,57	3,06	2,85	0,40	0,40	0,40	0,57	5,87
9	Затраты на шины (передний мост)	»	5,49	5,48	4,03	4,50	4,60	4,50	4,60	4,62	5,19
10	Всего затрат и издержек (с учетом потерь от простоев)	»	39,66	38,70	19,83	17,90	16,30	17,90	16,30	17,39	28,02
11	Эквивалент изменения производительности труда при ТО и Р (с учетом цеховых работ)	%	99,9	100	112	140	137	140	137	132	101

изводством для снижения времени простоев в ТО и ремонте и объясняет отсутствие (не использование) показателя коэффициента технической готовности (КТГ) в зарубежной практике.

Полученная оценка относительного перерасхода топлива (показатель 2) наиболее полно отражает возможности повышения качества технической эксплуатации вносимым стационарным и встроенным диагностированием. Она имеет и социальное значение, так как тесно связана с повышением уровня загрязнения атмосферы (получение конкретных оценок является достаточно сложной научно-технической задачей). При отсутствии диагностических воздействий (1-й и 2-й варианты) перерасход топлива достигает значительной величины (26 %, причем на зажигание приходилось около 5 %), а введение планового стационарного диагностирования (3-й и 4-й варианты) обеспечивает его резкое снижение (до 5...6 %).

Дополнительное использование встроенного диагностирования в сочетании со стационарным (5-й и 6-й варианты) уже не очень существенно влияет на снижение перерасхода топлива. Однако при отсутствии стационарного диагностирования (7-й вариант), что и в перспективе будет характерно для подавляющей части автомобильного парка, именно встроенное диагностирование может обеспечить приемлемый уровень перерасхода топлива, что свидетельствует об его огромной экономической и социальной значимости. Полученная для 7-го варианта оценка перерасхода топлива в 7 % обусловлена главным образом отсутствием плановых воздействий на шины и передний мост, а также меньшей глубиной обслуживания карбюратора.

С перерасходом топлива мало связан показатель вероятности работы с повышенной токсичностью холостого хода двигателя (показатель 3), имеющий чисто социальное значение. Он также свидетельствует о решающей роли диагностирования в снижении загрязнения атмосферы токсичными компонентами, в первую очередь, встроенного диагностирования с возможностью осуществления оперативных регулировок системы холостого хода силами водителя (5—7-й варианты). Полученные для условий отсутствия диагностирования (1-й и 2-й варианты) высокие значения вероятности работы с повышенной токсичностью (80 %) хорошо согласуются с многочисленными экспериментальными данными. При внедрении стационарного диагностирования с периодичностью проведения от 2 до 6 тыс. км (3-й и 4-й варианты) вероятность неисправной работы может быть снижена до приемлемых величин (6...8 %).

Наиболее важным в социальном отношении следует считать вероятность работы автомобилей с неисправными тормозами (показатель 4). При моделировании указанное состояние тормозов принималось в случае, когда неисправность или предельный износ накладок фиксировались для двух или трех колес автомобиля, что соответствует реальной диагностической практике. При наличии неисправностей тормозных механизмов всех четырех колес считалось,

что падение тормозной эффективности достигало предельного уровня, фиксируемого водителем, и предусматривалось полное восстановление тормозной системы с принудительной заменой колодок независимо от степени изношенности накладок. Последнее накладывало более жесткие ограничения на изменение состояния тормозов по сравнению с реальной практикой. Не моделировалось изменение состояния самой пневматической системы (в связи с отсутствием надежных данных), однако это также способствовало ужесточению ограничений, а конечные результаты оказывалось возможным интерпретировать для тормозных систем с гидроприводом.

Хотя полученные таким образом результаты моделирования дают немного заниженную оценку эффективности диагностирования, их сравнительный анализ обеспечивает выявление существа вносимых изменений от его использования. Характерным является высокая вероятность (0,5) эксплуатации автомобилей с неисправными тормозами при отсутствии предупредительных воздействий (1-й вариант). Хотя эта вероятность немного ниже вероятности, полученной при массовой проверке автомобилей на одной из дорог (0,77), однако она совпадает с зарубежными оценками. Принудительная профилактика тормозов с заменой колодок с периодичностью ТО-2 почти вдвое снижает вероятность неисправного состояния (до 0,28, 2-й вариант).

Высокое значение оценки вероятности эксплуатации с неисправным состоянием тормозов для 2-го варианта позволяет сделать вывод о принципиальной невозможности полного решения этой социальной проблемы на основе чисто профилактических систем поддержания работоспособности автомобилей, поскольку теоретически любая принудительная профилактика целесообразна при определенном уровне накопления неисправностей.

Значительное снижение данной вероятности возможно только на основе диагностирования, причем именно встроенное диагностирование (а не стационарное и не инструментальный контроль, выполняемый с большой периодичностью при годовых технических осмотрах) обеспечивает практическое исключение эксплуатации с неисправными тормозами (5-й и 7-й варианты). Низкое значение вероятности имеет место и для 3-го варианта, когда диагностирование Д-1 дополняет профилактическое восстановление тормозов при ТО-2, однако диагностирование Д-1 на практике не получило широкого распространения из-за системной недоработки социологических, организационных и других причин.

Поскольку ЕТОД (4-й вариант) по результатам системного анализа являлся наиболее предпочтительным вариантом организации стационарного диагностирования в АТП, полученная для него достаточно высокая вероятность неисправного состояния тормозов (0,2) заслуживает дополнительного анализа. Надо иметь в виду, что



периодичность ЕТОД, определенная при моделировании по минимуму удельных издержек на эксплуатацию, составила 6 тыс. км, т.е. оказалась в три раза большей периодичности Д-1.

В то же время указанный минимум не имел ярко выраженного характера, а удельные затраты изменялись незначительно (в пределах 5 %) в диапазоне от 4 000 до 8 000 км пробега. Данный результат, полученный теоретическим путем, имеет важное практическое значение, поскольку указывает прежде всего на наименьшую опасность нарушения режимов профилактики, всегда имеющих место на практике в случае ЕТОД. Кроме того, он подтверждает рациональность отмеченного ранее простейшего подхода к организации ЕТОД — назначения его режимов вдвое меньшими по сравнению с режимами ТО-2 (в рассматриваемом случае — 4 000 км). Это приведет к увеличению затрат на обслуживание, зато (по принципу «кашу маслом не испортишь») системно обеспечит дополнительную подстраховку своевременности контроля технического состояния тормозов (при этом вероятность неисправного состояния может составить 13... 15 %) и экологических показателей повышенной значимости (например, описанного в подразд. 3.4 метода проверки бензинового автомобиля на частично нагруженных режимах).

Минимальная сложность проведения комплексного диагностирования Д-1 в режиме ЕТОД с максимально возможной периодичностью (в рассматриваемом случае — до 8 000 км) делает возможным (и целесообразным при отсутствии встроенного диагностирования) создание на этой основе государственной системы контроля за техническим состоянием грузовых автомобилей и автобусов индивидуальных владельцев и небольших предприятий при выдаче им лицензий на перевозку пассажиров и грузов повышенной опасности. Назначаемые при этом повышенные режимы прохождения обязательных проверок (календарные или по пробегу) могут оперативно корректироваться по результатам текущего контроля. Это системно позволит снять с организаций, выдающих лицензии (районные отделения Российской транспортной инспекции), долю юридической ответственности за возможные ДТП по техническому состоянию.

Значение вероятности эксплуатации с технологическими неисправностями тормозов в 20 % немного ниже по сравнению с профилактической системой (2-й вариант), однако рационально с технологических позиций предупредительного восстановления на основе диагностирования (см. подразд. 7.2). О социальной неопасности этой величины косвенно свидетельствует общее улучшение функционирования тормозов, зафиксированное в активных экспериментах при внедрении диагностирования в технологические процессы обслуживания. Кроме того, более важное значение имеет оперативный контроль и управление состоянием тормозов (на основе изменения диагностических нормативов и периодичности плановых воздействий), легко осуществляемое в процессе ЕТОД. Как показали сис-

темные исследования, решение проблемы повышения безопасности в основном связано с правильным выбором режима движения, хотя пониженное внимание к состоянию тормозов для автомобилей индивидуального пользования и находящихся в небольших автотранспортных организациях косвенно, а иногда и непосредственно, может явиться причиной ДТП.

Вторым традиционным критерием для оценки эффективности технической эксплуатации являются предложенные д-ром техн. наук, проф. Г.В.Крамаренко [8] суммарные удельные затраты на обслуживание и ремонт, по минимуму которых оптимизировались режимы технического обслуживания и обосновывалась стратегия системы в целом. Однако в случае использования диагностирования указанные показатели уже не отражают все стороны ее влияния и их необходимо расширить с учетом возможного изменения топливной экономичности, затрат на шины, а также потерь от вынужденного прекращения транспортного процесса. Отдельные показатели этого нового комплекса имеют социальное значение, тем не менее только их суммирование дает возможность получить единый обобщающий критерий и на его основе рационально учесть неизбежные противоречия и их текущую конъюнктуру. На фоне экономического обобщения возможен и более объективный учет социальных сторон.

Приведенный в табл. 7.4 комплекс экономических показателей (показатели 5... 10), полученный на основе соответствующих перерасчетов результатов моделирования, подтверждает сделанный ранее вывод. Наиболее сложными и принципиальными из них являются три показателя (показатели 5... 7), отражающие специфику изменения технологических процессов поддержания работоспособности автомобилей и выявляющие возможные резервы повышения производительности труда ремонтных рабочих. По этим показателям также можно внести предложения о возможном снижении расхода запасных частей, поскольку согласно исследованиям в общем балансе затрат на ремонт составляющая на запасные части (около 40 %) даже немного выше составляющей от используемого труда. Появление 7-го показателя связано с принципиальной возможностью диагностирования своевременного выявления потребности в некоторой части текущего ремонта и планирования его предупредительного проведения с периодичностью диагностирования.

Экспериментальные данные (см. табл. 7.1, 7.2) позволили при моделировании получить только изменение постовых работ текущего ремонта, трудоемкость которых для 2-го варианта организации считается равной трудоемкости цеховых работ, и указанный уровень цеховых работ был принят без изменения для всех остальных вариантов.

Тем не менее данные моделирования показывают значительные резервы повышения производительности труда при рациональном использовании диагностирования (показатель 11).

Для понимания сущности этих резервов необходимо более подробно остановиться на каждом варианте. Так, в 1-м варианте отсутствуют затраты на плановое обслуживание и имеет место только текущий ремонт по потребности в связи с аварийными остановками автомобиля. Как показывает практика, в подобных случаях с более или менее равномерной периодичностью выполняются смазочные работы, которые неправильно отождествляются с работами по ТО-1, также как работы по ремонту тормозов по потребности отождествляются с работами по ТО-2, что дает основание говорить об определенной степени выполнения профилактической планово-предупредительной системы (2-й вариант). Поэтому реально различия между 1-м и 2-м вариантами организации могут быть установлены только в теоретическом плане и базовой основой такой теории в настоящее время следует считать имитационное моделирование.

Его результаты показывают, что введение профилактической системы обеспечивает существенное снижение (примерно на 35 %) трудно прогнозируемых постовых работ текущего ремонта за счет восстановления тормозов (и рулевого управления) в процессе ТО-2, что помимо улучшения управляемости технической службой АТП и роста коэффициента использования приводит к повышению среднего уровня состояния систем, влияющих на безопасность движения. За указанные преимущества приходится расплачиваться общим повышением трудоемкости постовых работ (на 13 %), хотя в целом (с учетом цеховых работ) это не приводит к снижению производительности труда. Последнее в определенной мере объясняет интуитивное стремление работников АТП по возможности сокращать профилактические воздействия, в частности для новых автомобилей и при повышенном внимании к техническому состоянию со стороны водителей.

Дополнение двухступенчатой профилактической системы комплексом диагностических работ Д-1 и Д-2 (3-й вариант) может приводить к более чем четырехкратному сокращению постовых работ случайного ремонта, а основная часть текущего ремонта (60 %) будет планироваться по результатам текущего диагностирования, что в целом усиливает планово-предупредительный характер и управляемость системы. При этом возможно существенное снижение (на 40 %) объема постовых работ текущего ремонта, однако большая часть снижения может быть реализована только при условии использования современных методов и средств виброакустической диагностики агрегатов трансмиссии и газораспределительного механизма двигателя, что и на перспективу представляется проблематичным.

Тем не менее зафиксированное в реальной практике использование обычного комплекса Д-2 дает снижение трудоемкости текущего ремонта на 10 % (с учетом цеховых работ), что является достаточно весомым, и не случайно 3-й вариант считался базовым при внедрении диагностирования в АТП. Однако теоретически общее сниже-

ние производительности труда при данном варианте несущественно (на 12 % по сравнению с 1-м и 2-м вариантами) и не удовлетворяет перспективной потребности ее роста. Причиной является сохранение значительного объема профилактических предупредительных работ ТО-1 и ТО-2, противоречащих основной идее использования диагностирования.

Именно сокращение этих работ до уровня, соответствующего проведению обслуживания систем автомобиля, влияющих на безопасность движения, расход топлива и загрязнение атмосферы по потребности, что может быть достигнуто за счет внедрения ЕТОД (4-й вариант), обеспечивает резкое повышение труда (теоретически на 40 %, а с учетом поправок на отсутствие современных средств виброакустического диагностирования — не менее чем на 25 %). Некоторое увеличение трудоемкости (и снижение производительности труда) имеет место для 5-го варианта, когда плановое ЕТОД дополняется встроенным диагностированием, по результатам которого неизбежен дополнительный текущий ремонт тормозов или доплановое проведение регламента ЕТОД, но именно в результате этого достигается практическое исключение эксплуатации автомобилей с неисправными тормозами.

Что касается 6-го варианта (неплановое ЕТОД по результатам встроенного диагностирования тормозов), то несмотря на самую низкую потребность в выполнении регламентных работ, здесь имеют место резкое увеличение случайного текущего ремонта (за счет нерациональной периодичности диагностирования) и общее повышение трудоемкости ТО и ТР (примерно на 5 % по сравнению с 5-м вариантом). Однако более существенным недостатком этого варианта следует считать ухудшение плановой управляемости работой предприятия. Таким образом, результаты теоретического анализа свидетельствуют о целесообразности сохранения принципов планового предупреждения при организации обслуживания и ремонта автомобилей с диагностированием применительно к АТП общего пользования и о ведущей роли в этом отношении стендового диагностирования.

Что касается встроенного диагностирования, то его «чистое» (неплановое) применение в небольших АТП (7-й вариант) по результатам моделирования практически не выявило повышения производительности труда, хотя здесь можно отметить недоработку модели для более детального исследования этого варианта. В то же время очевидно, что планово-регламентное использование подобной системы (в рамках ЕТОД или ему подобных формах организации обслуживания) приведет к выходным технико-экономическим и социальным показателям, близким к 4-му или 5-му варианту.

Следующая группа показателей (показатели 8... 10) отражает экономические стороны, уже рассмотренные ранее с технических позиций возможных изменений коэффициента использования автомо-

билей и перерасхода ими топлива, а также повышения ресурса шин. За основу экономических последствий от изменения коэффициента использования условно были приняты потери прибыли при вынужденном прекращении транспортного процесса. Преимущество данного показателя по сравнению с другими (в частности, основанными на условии применения дополнительных автомобилей для выполнения транспортной работы взамен восстанавливаемых) заключается в легкости его определения по отчетным данным АТП с учетом конъюнктуры перевозок, а также в большей критичности оценки последствий простоев. Проведенные на основе некоторых отчетных данных перерасчеты показали, что теоретически потери прибыли от простоев почти равны общим затратам на восстановление технического состояния и поэтому могут являться важным и независимым фактором, стимулирующим их снижение.

Еще большую экономическую значимость имеют потери, связанные с перерасходом топлива, которые при современных ценах более чем вдвое превышают возможный экономический эффект от снижения затрат на восстановление и потерь от простоев, вместе взятых. Из сравнения сопоставимых по организации 1-го и 7-го вариантов, а также 4-го и 5-го вариантов можно видеть, что общий экономический эффект от встроенного диагностирования определяется только снижением расхода топлива.

Что касается изменений затрат на шины, то хотя возможный экономический эффект здесь наиболее низкий (не более 5 % от общего), главное значение имеет снижение потребности в этом дефицитном изделии (примерно на 20...28 % по передним колесам). По результатам моделирования около 70 % повышения ресурса шин может быть достигнуто за счет упорядочения операций обслуживания по восстановлению давления, а 30 % — за счет регулирования боковых сил управляемых колес при помощи диагностического стенда.

Суммирование перечисленных ранее переменных экономических показателей свидетельствует о высокой эффективности диагностирования (показатель 10). Если в теоретическом отношении экономическое различие между основными вариантами (3—6-й варианты) организации стендового диагностирования является несущественным (50...80 усл. ед. на один автомобиль в год при его пробеге 50 тыс. км), то по отношению к профилактической системе ТО-1 — ТО-2 (2-й вариант) экономический эффект может достигнуть (без учета потерь от простоев) величины 850 усл. ед. на один автомобиль в год. При применении только встроенного диагностирования теоретический эффект составляет 500 усл. ед. в год. Эта оценка может способствовать независимой от завода-изготовителя установке на автомобили портативных средств встроенного диагностирования, что уже имеет место в зарубежной практике эксплуатации автобусов и большегрузных автомобилей и подтверждает полученные ранее результаты системных исследований.

В России подобным устройством можно считать прибор «Компьютерный диагност» с датчиком расхода топлива, недавно выпущенный на коммерческой основе для автомобилей ВАЗ с электронной системой управления двигателем, а также встраиваемые электронные устройства для контроля состояния контактных систем зажигания. Что касается стационарного диагностирования, то обоим экономический критерий свидетельствует в пользу перехода от двухступенчатой (3-й вариант) к одноступенчатой системе диагностического регламентного обслуживания (4-й и 5-й варианты). Хотя потенциальный экономический эффект здесь относительно невелик (около 100 усл. ед. на один автомобиль в год), он имеет значительную социальную значимость, поскольку в основном связан со снижением трудовых и материальных затрат на обслуживание и ремонт.

Проведенное немного позднее модельное исследование на примере реального парка дизельных автомобилей, выполняющих международные перевозки (на языке GPSS), также подтвердило целесообразность комплексного одноступенчатого диагностирования для организации качественного обслуживания в отношении безопасности движения и загрязнения окружающей среды при одновременной возможности получения дополнительной прибыли от увеличения объема перевозок за счет снижения простоев автомобилей.

Таким образом, общей методологической основой решения проблемы совершенствования системы поддержания работоспособности автомобилей следует признать системный анализ в сочетании с имитационным моделированием. При этом отдельные теоретические методы, такие как определение периодичности обслуживания и диагностирования, которым ранее придавалось ведущее значение, должны применяться для решения частных, а не стратегических задач.

## КОНТРОЛЬНЫЕ ВОПРОСЫ

1. В чем заключается удобство имитационного моделирования при исследовании существующих и перспективных организационно-технологических систем на АТП?
2. Перечислите основные задачи, решаемые в процессе этапа имитационного моделирования вариантов организации ТО и ремонта автомобилей для АТП общего пользования.
3. Каковы особенности использования характеристик надежности и других статистических показателей ТЭА при имитационном моделировании работы автомобилей в АТП?
4. Каковы особенности 1-й стратегии обслуживания и ремонта для единичных изделий (базовых деталей, систем или агрегатов автомобиля), отказы которых обусловлены процессами изнашивания или постепенного накопления неисправностей?

5. Дайте краткую характеристику достоинств и недостатков 2-й стратегии обслуживания и ремонта для единичных изделий автомобиля применительно к условиям работы АТП общего пользования.

6. Каковы особенности 3-й стратегии обслуживания и ремонта для единичных изделий (базовых деталей, систем или агрегатов автомобиля), отказы которых обусловлены процессами изнашивания или постепенного накопления неисправностей?

7. Расскажите о 2-м варианте организации проведения ТО и ремонта автомобилей для АТП общего пользования.

8. Расскажите о 4-м варианте организации проведения ТО и ремонта автомобилей для АТП общего пользования.

9. Расскажите о 7-м варианте организации проведения ТО и ремонта автомобилей для АТП общего пользования.

10. Сделайте основные выводы о перспективах применения стационарного и встроенного диагностирования на автомобильном транспорте, которые следуют из результатов этапа имитационного моделирования.

## ГЛАВА 8

# Приборное и метрологическое обеспечение научных исследований на автотранспортных предприятиях

### 8.1. Основные понятия и определения в области метрологии

В целях обеспечения единства измерений в Российской Федерации, регулирования отношений государственных органов управления Российской Федерации с юридическими и физическими лицами по вопросам изготовления, выпуска, эксплуатации, ремонта, продажи и импорта средств измерений и обеспечения защиты прав и законных интересов граждан, установленного правопорядка и экономики Российской Федерации от отрицательных последствий недостоверных результатов измерений принят Федеральный закон от 26.06.2008 № 102-ФЗ «Об обеспечении единства измерений». Он устанавливает правовые основы обеспечения единства измерений в Российской Федерации; защищает права и законные интересы граждан, общества и государства от отрицательных последствий недостоверных результатов измерений; обеспечивает потребности граждан, общества и государства в получении объективных, достоверных и сопоставимых результатов измерений, используемых в целях защиты жизни и здоровья граждан, охраны окружающей среды, животного и растительного мира, обеспечения обороны и безопасности государства, в том числе экономической безопасности, и способствует развитию экономики Российской Федерации и научно-техническому прогрессу.

В соответствии с этим законом приняты следующие основные понятия и определения в области измерений:

1) аттестация методик (методов) измерений — исследование и подтверждение соответствия методик (методов) измерений установленным метрологическим требованиям к измерениям;

2) ввод в эксплуатацию средства измерений — документально оформленная в установленном порядке готовность средства измерений к использованию по назначению;

3) государственный метрологический надзор — контрольная деятельность в сфере государственного регулирования обеспечения единства измерений, осуществляемая уполномоченными федеральными органами исполнительной власти и заключающаяся в систематической проверке соблюдения установленных законодательством Российской Федерации обязательных требований, а также в применении установленных законодательством Российской Федерации мер за нарушения, выявленные во время надзорных действий;

4) государственный первичный эталон единицы величины — государственной эталон единицы величины, обеспечивающий воспроизведение, хранение и передачу единицы величины с наивысшей в Российской Федерации точностью, утверждаемый в этом качестве в установленном порядке и применяемый в качестве исходного на территории Российской Федерации;

5) государственный эталон единицы величины — эталон единицы величины, находящийся в федеральной собственности;

6) единица величины — фиксированное значение величины, которое принято за единицу данной величины и применяется для количественного выражения однородных с ней величин;

7) единство измерений — состояние измерений, при котором их результаты выражены в допущенных к применению в Российской Федерации единицах величин, а показатели точности измерений не выходят за установленные границы;

8) измерение — совокупность операций, выполняемых для определения количественного значения величины;

9) испытания стандартных образцов или средств измерений в целях утверждения типа — работы по определению метрологических и технических характеристик однотипных стандартных образцов или средств измерений;

10) калибровка средств измерений — совокупность операций, выполняемых в целях определения действительных значений метрологических характеристик средств измерений;

11) методика (метод) измерений — совокупность конкретно описанных операций, выполнение которых обеспечивает получение результатов измерений с установленными показателями точности;

12) метрологическая служба — организующие и (или) выполняющие работы по обеспечению единства измерений и (или) оказывающие услуги по обеспечению единства измерений структурное подразделение центрального аппарата федерального органа исполнительной власти и (или) его территориального органа, юридическое лицо или структурное подразделение юридического лица либо объединения юридических лиц, работники юридического лица, индивидуальный предприниматель;

13) метрологическая экспертиза — анализ и оценка правильности установления и соблюдения метрологических требований применительно к объекту, подвергаемому экспертизе. Метрологическая экс-

пертиза проводится в обязательном (обязательная метрологическая экспертиза) или добровольном порядке;

14) метрологические требования — требования к влияющим на результат и показатели точности измерений характеристикам (параметрам) измерений, эталонов единиц величин, стандартных образцов, средств измерений, а также к условиям, при которых эти характеристики (параметры) должны быть обеспечены;

15) обязательные метрологические требования — метрологические требования, установленные нормативными правовыми актами Российской Федерации и обязательные для соблюдения на территории Российской Федерации;

16) передача единицы величины — приведение единицы величины, хранимой средством измерений, к единице величины, воспроизводимой эталоном данной единицы величины или стандартным образцом;

17) поверка средств измерений (далее — поверка) — совокупность операций, выполняемых в целях подтверждения соответствия средств измерений метрологическим требованиям;

18) прослеживаемость — свойство эталона единицы величины или средства измерений, заключающееся в документально подтвержденном установлении их связи с государственным первичным эталоном соответствующей единицы величины посредством сличения эталонов единиц величин, поверки, калибровки средств измерений;

19) прямое измерение — измерение, при котором искомое значение величины получают непосредственно от средства измерений;

20) сличение эталонов единиц величин — совокупность операций, устанавливающих соотношение между единицами величин, воспроизводимых эталонами единиц величин одного уровня точности и в одинаковых условиях;

21) средство измерений — техническое средство, предназначенное для измерений;

22) стандартный образец — образец вещества (материала) с установленными по результатам испытаний значениями одной и более величин, характеризующих состав или свойство этого вещества (материала);

23) технические системы и устройства с измерительными функциями — технические системы и устройства, которые наряду с их основными функциями выполняют измерительные функции;

24) технические требования к средствам измерений — требования, которые определяют особенности конструкции средств измерений (без ограничения их технического совершенствования) в целях сохранения их метрологических характеристик в процессе эксплуатации средств измерений, достижения достоверности результата измерений, предотвращения несанкционированных настройки и вмешательства, а также требования, обеспечивающие безопасность и электромагнитную совместимость средств измерений;

25) тип средств измерений — совокупность средств измерений, предназначенных для измерений одних и тех же величин, выраженных в одних и тех же единицах величин, основанных на одном и том же принципе действия, имеющих одинаковую конструкцию и изготовленных по одной и той же технической документации;

26) тип стандартных образцов — совокупность стандартных образцов одного и того же назначения, изготавливаемых из одного и того же вещества (материала) по одной и той же технической документации;

27) утверждение типа стандартных образцов или типа средств измерений — документально оформленное в установленном порядке решение о признании соответствия типа стандартных образцов или типа средств измерений метрологическим и техническим требованиям (характеристикам) на основании результатов испытаний стандартных образцов или средств измерений в целях утверждения типа;

28) фасованные товары в упаковках — товары, которые упаковываются в отсутствие покупателя; при этом содержимое упаковки не может быть изменено без ее вскрытия или деформирования, а масса, объем, длина, площадь или иные величины, определяющие количество содержащегося в упаковке товара, должны быть обозначены на упаковке;

29) эталон единицы величины — техническое средство, предназначенное для воспроизведения, хранения и передачи единицы величины.

## 8.2. Метрологическая служба

Государственная метрологическая служба находится в ведении Федерального агентства по техническому регулированию и метрологии (Госстандарт; в 2004—2010 гг. — Ростехрегулирование; с июня 2010 г. — Росстандарт).

Государственная метрологическая служба включает в себя:

- государственные научные метрологические центры;
- органы Государственной метрологической службы на территориях республик в составе Российской Федерации, автономной области, автономных округов, краев, областей, городов Москвы и Санкт-Петербурга. Органы Государственной метрологической службы осуществляют государственный метрологический контроль и надзор на территориях субъектов Российской Федерации;

- головной институт в системе Госстандарта России ВНИИМС.

Государственная метрологическая служба (ГМС) России представляет собой совокупность государственных метрологических органов и создается для управления деятельностью по обеспечению единства измерений.

Общее руководство ГМС осуществляет Росстандарт РФ, на который Федеральным законом от 26.06.2008 № 102-ФЗ «Об обеспечении единства измерений» возложены следующие функции:

- межрегиональная и межотраслевая координация деятельности по обеспечению единства измерений;
- представление Правительству РФ предложений по единицам величин, допускаемым к применению;
- установление правил создания, утверждения, хранения и применения эталонов единиц величин;
- определение общих метрологических требований к средствам, методам и результатам измерений;
- государственный метрологический контроль и надзор;
- контроль за соблюдением условий международных договоров Российской Федерации о признании результатов испытаний и поверки средств измерений;
- руководство деятельностью Государственной метрологической службы и иных государственных служб обеспечения единства измерений;
- участие в деятельности международных организаций по вопросам обеспечения единства измерений;
- утверждение нормативных документов по обеспечению единства измерений;
- утверждение государственных эталонов;
- установление межповерочных интервалов средств измерений;
- отнесение технических устройств к средствам измерений;
- установление порядка разработки и аттестации методик выполнения измерений;
- ведение и координация деятельности Государственных научных метрологических центров (ГНМЦ), Государственной метрологической службы, Государственной службы времени и частоты (ГСВЧ), Государственной службы стандартных образцов (ГССО), Государственной службы стандартных справочных данных (ГСССД);
- аккредитация государственных центров испытаний средств измерений;
- утверждение типа средств измерения;
- ведение Государственного реестра средств измерений;
- аккредитация метрологических служб юридических лиц на право поверки средств измерений;
- утверждение перечней средств измерений, подлежащих поверке;
- установление порядка лицензирования деятельности юридических и физических лиц по изготовлению, ремонту, продаже и прокату средств измерений;
- организация и координация деятельности государственных инспекторов по обеспечению единства измерений;
- организация деятельности и аккредитация метрологических служб юридических лиц на право проведения калибровочных работ;

- планирование и организация выполнения метрологических работ.

В состав ГМС входят семь государственных научных метрологических центров, Всероссийский научно-исследовательский институт метрологической службы (ВНИИМС) и около 100 центров стандартизации и метрологии. Наиболее крупные среди научных центров — НПО «ВНИИ метрологии имени Д. И. Менделеева» (ВНИИМ, Санкт-Петербург), НПО «ВНИИ физико-технических и радиотехнических измерений» (ВНИИФТРИ, Московская область), Сибирский государственный научно-исследовательский институт метрологии (СНИИМ, Новосибирск), Уральский научно-исследовательский институт метрологии (УНИИМ, Екатеринбург). Научные центры являются держателями государственных эталонов, а также проводят исследования по теории измерений, принципам и методам высокоточных измерений, разработке научно-методических основ совершенствования российской системы измерений.

### 8.3. Метрологическое обеспечение научных исследований

*Метрологическое обеспечение* — установление и применение научных и организационных основ, технических средств, правил и норм, необходимых для достижения единства и требуемой точности измерения.

Метрологическое обеспечение базируется на научной, технической и организационной основе.

Научной основой метрологического обеспечения (МО) является *метрология* — наука об измерениях, методах и средствах обеспечения их единства и способах достижения требуемой точности.

Технической основой МО являются:

- воспроизведение на предприятии (учреждении) единиц физических величин с помощью рабочих эталонов или образцовых средств измерений;
- передача размеров единиц физических величин рабочим средствам измерений путем осуществления обязательной государственной или ведомственной поверки средств измерений в целях достижения единообразия парка средств измерений при их эксплуатации и ремонте;
- использование стандартных справочных данных о физических константах и свойствах веществ и материалов, предназначенных для получения достоверных результатов при проведении научных исследований;
- разработка и обязательная метрологическая аттестация нестандартных средств измерений.

Организационной основой МО является служба стандартизации и метрологии предприятия, состоящая из отдела стандартизации и

метрологии, экспертов-метрологов и ответственных за состояние средств измерений в подразделениях.

Основная задача МО — постоянное поддержание средств измерений в нормальном эксплуатационном состоянии. Ее решение обеспечивается путем:

- проведения входного контроля новых средств измерения;
- постановки средств измерений на технический учет в отделе стандартизации и метрологии;
- назначения межповерочных интервалов средств измерений;
- планирования и организации поверки средств измерений;
- разработки, аттестации и эксплуатации нестандартизованных средств измерений;
- метрологического контроля за эксплуатацией средств измерений в подразделениях;
- ремонта и наладки средств измерений;
- сбора и анализа информации о качестве средств измерений в эксплуатации.

Поверку средств измерения (СИ) проводят для установления их пригодности к применению.

Пригодными к применению в течение установленного межповерочного интервала признают СИ, поверка которых, выполненная в соответствии с требованиями нормативно-технических документов, подтверждает их соответствие метрологическим и техническим требованиям к данному СИ.

В соответствии с Правилами по метрологии «Государственная система обеспечения единства измерений. Порядок проведения поверки средств измерений» (ПР 50.2.006-94) средства измерений подвергаются первичной, периодической, внеочередной и инспекционной поверкам.

Первичной поверке подлежат средства измерений утвержденных типов при выпуске из производства и ремонта, при ввозе по импорту. Первичной поверке могут не подвергаться средства измерений при ввозе по импорту на основании заключенных международных соглашений (договоров) о признании результатов поверки, произведенной в зарубежных странах.

Первичной поверке подлежит, как правило, каждый экземпляр средств измерений. Допускается выборочная поверка.

Первичную поверку органы Государственной метрологической службы могут производить на контрольно-поверочных пунктах, организуемых юридическими лицами, выпускающими и ремонтирующими средства измерений.

Периодической поверке подлежат средства измерений, находящиеся в эксплуатации или на хранении, через определенные межповерочные интервалы.

Конкретные перечни средств измерений, подлежащих поверке, составляют юридические и физические лица — владельцы средств



измерений. Перечни средств измерений, подлежащих поверке, направляют в органы Государственной метрологической службы. Органы Государственной метрологической службы в процессе осуществления государственного надзора за соблюдением метрологических правил и норм контролируют правильность составления перечней средств измерений, подлежащих поверке.

Периодическую поверку должен проходить каждый экземпляр средств измерений. Периодической поверке могут не подвергаться средства измерений, находящиеся на длительном хранении. Периодическую поверку средств измерений, предназначенных для измерений (воспроизведения) нескольких величин или имеющих несколько диапазонов измерений, но используемых для измерений (воспроизведения) меньшего числа величин или на меньшем числе диапазонов измерений, допускается на основании решения главного метролога или руководителя юридического лица производить только по тем требованиям нормативных документов по поверке, которые определяют пригодность средств измерений для применяемого числа величин и применяемых диапазонов измерений. Соответствующая запись должна быть сделана в эксплуатационных документах.

Результаты периодической поверки действительны в течение межповерочного интервала. Первый межповерочный интервал устанавливается при утверждении типа. Органы Государственной метрологической службы и юридические лица обязаны вести учет результатов периодических поверок и разрабатывать рекомендации по корректировке межповерочных интервалов с учетом специфики их применения. Корректировка межповерочных интервалов проводится органом Государственной метрологической службы по согласованию с метрологической службой юридического лица. Периодическая поверка может производиться на территории пользователя, органа Государственной метрологической службы или юридического лица, аккредитованного на право поверки. Место поверки выбирает пользователь средств измерений исходя из экономических факторов и возможности транспортировки поверяемых средств измерений и эталонов.

Внеочередную поверку производят при эксплуатации (хранении) средств измерений:

- при повреждении знака поверительного клейма, а также в случае утраты свидетельства о поверке;
- вводе в эксплуатацию средств измерений после длительного хранения (более одного межповерочного интервала);
- проведении повторной юстировки или настройки, известном или предполагаемом ударном воздействии на средство измерений или неудовлетворительной работе прибора.

Инспекционную поверку производят для выявления пригодности к применению средств измерений при осуществлении государственного метрологического надзора. Инспекционную поверку можно производить не в полном объеме, предусмотренном методикой по-

верки. Результаты инспекционной поверки отражают в акте поверки. Инспекционную поверку производят в присутствии представителя проверяемого юридического или физического лица.

Поверка в рамках метрологической экспертизы, производимой по поручению органов суда, прокуратуры, арбитражного суда и федеральных органов исполнительной власти, проводится по их письменному требованию. По результатам поверки составляют заключение, которое утверждает руководитель органа Государственной метрологической службы, и направляют его заявителю. Один экземпляр заключения должен храниться в органе Государственной метрологической службы, проводившем поверку.

Юридические и физические лица, выпускающие средства измерений из производства или ремонта, ввозящие средства измерений и использующие их в целях эксплуатации, проката или продажи, обязаны своевременно представлять средства измерений на поверку.

Органы Государственной метрологической службы осуществляют поверку средств измерений на основании графиков поверки, составляемых юридическими и физическими лицами. Графики поверки составляются на срок, устанавливаемый владельцами средств измерений. Сроки представления графиков поверки устанавливают органы Государственной метрологической службы. Графики поверки могут быть скорректированы в зависимости от изменения номенклатуры и количества средств измерений.

Графики поверки направляются в орган Государственной метрологической службы, на обслуживаемой территории которого находятся владельцы средств измерений. Графики поверки составляются в трех экземплярах. В течение 10 дней с момента поступления графиков поверки средств измерений орган Государственной метрологической службы проводит их рассмотрение. Порядок рассмотрения и согласования графиков поверки устанавливает руководитель органа Государственной метрологической службы. При рассмотрении графиков поверки определяют средства измерений, поверка которых проводится в органе Государственной метрологической службы.

Данные средства измерений отмечаются в третьем экземпляре, который возвращается для сведения заявителю. В ответе могут быть указаны другие органы Государственной метрологической службы или юридические лица, которые могут обеспечить поверку средств измерений, не обеспеченных поверкой в данном органе Государственной метрологической службы. Заявитель повторно направляет графики поверки в другой орган Государственной метрологической службы или юридическое лицо по своему выбору, который их согласовывает. При согласовании графиков поверки проверяют полноту информации о средствах измерений, представляемых на поверку, уточняют место, сроки, объем поверки, а также оплату. Первый экземпляр согласованных графиков поверки и подписанных руководителем органа Государственной метрологической службы направляется заявителю.



Доставку средств измерений на поверку обеспечивают юридические и физические лица — владельцы средств измерений.

Средства измерений сдаются на поверку в органы Государственной метрологической службы под расписку. Ответственность за сохранность средств измерений несет орган Государственной метрологической службы в соответствии с действующим законодательством.

#### 8.4. Нормирование метрологических характеристик

Нормированные метрологические характеристики СИ устанавливаются государственными стандартами в зависимости от их специфики и назначения. К ним относятся цена деления равномерной шкалы СИ, минимальная цена деления неравномерной шкалы прибора, пределы шкалы, погрешности СИ, время установления показаний. Полный перечень нормируемых метрологических характеристик конкретного СИ регламентирован ГОСТ 8.009—84 «Государственная система обеспечения единства измерений. Нормируемые метрологические характеристики средств измерений».

Пределы допустимых погрешностей обычно выражают в форме приведенных, относительных или абсолютных погрешностей в зависимости от типа СИ, их назначения, условий применения. Так, для определения допустимой погрешности СИ длины и массы, когда она выражается в единицах измеряемой величины (в миллиметрах и граммах) используют абсолютную погрешность.

Пределы допустимой абсолютной погрешности устанавливаются по формуле

$$\Delta = \pm(a + bx),$$

где  $\Delta$  — пределы допустимой абсолютной основной погрешности, выраженной в единицах измеряемой величины или в делениях шкалы прибора;  $a$ ,  $b$  — положительные числа, не зависящие от  $x$ ;  $x$  — значение измеряемой величины на входе (выходе) средств измерений или число делений шкалы прибора (СИ).

Конкретные значения  $a$  и  $b$  для СИ регламентированы ГОСТ 8.401—80 «Государственная система обеспечения единства измерений. Классы точности средств измерений. Общие требования».

В том случае, если границы погрешностей СИ неизменны в пределах диапазона измерений, пределы допустимых погрешностей выражают в форме приведенных погрешностей, а если указанные границы не постоянны по диапазону измерений, то относительными погрешностями.

Пределы допустимой приведенной погрешности устанавливают по формуле

$$\gamma = \frac{\Delta}{x_N} = \pm P,$$

где  $x_N$  — больший из пределов измерений СИ или больший из модулей пределов измерений, если нулевое значение находится внутри диапазона измерений;  $P$  — отвлеченное положительное число, выбираемое, из ряда  $1 \times 10$ ;  $1,5 \times 10$ ;  $1,6 \times 10$ ;  $2 \times 10$ ;  $2,5 \times 10$ ;  $4 \times 10$ ;  $5 \times 10$ ;  $6 \times 10$  (1; 0; -1; -2 и т.д.).

Пределы допустимой относительной основной погрешности определяют по формуле

$$\delta = \frac{\Delta}{X} = \pm q,$$

где  $X$  — значение измеряемой величины на входе (выходе) или число делений шкалы СИ;  $q$  — относительное положительное число, которое выбирается из ряда, приведенного для числа  $P$ .

#### 8.5. Измерение физических величин, источники ошибок

*Физическая величина* — свойство, общее в качественном отношении многим объектам (системам и т.д.), но в количественном отношении индивидуальное для каждого объекта.

Термин «величина» применяется в отношении свойств или их характеристик, которые мы умеем оценивать количественно, т.е. измерять. Правда, существуют и такие свойства и характеристики, которые нельзя оценить количественно, но можно найти способ их количественной оценки (например, запах и т.д.). Принятая в настоящее время терминология дает следующее определение понятия «измерение».

*Измерение* — нахождение значения физической величины опытным путем с помощью специальных технических средств. По способу получения числового значения измеряемой величины различают прямые, косвенные, совокупные и совместные измерения. Прямыми называют измерения, при которых искомое значение величины находится непосредственно из опытных данных (измерение длины линейкой, температуры — термометром и т.д.). Косвенными считают измерения, результат которых определяется путем прямых измерений величин, связанных с измеряемой величиной известной зависимостью (скорость автомобиля по результатам замера пройденного пути и затраченного на это времени и т.д.). Совокупными называются измерения, в которых значения измеряемых величин находят по данным повторных измерений одной или нескольких одноименных величин при различных сочетаниях этих

величин. Результаты совокупных измерений находят путем решения системы уравнений, полученных по результатам прямых измерений. Совместными и называются производимые одновременно (прямые или косвенные) измерения двух или нескольких неодновременных величин. Их цель — нахождение функциональной зависимости между величинами (зависимость длины седла от температуры и т.д.).

Для измерения физических величин в исследованиях применяют мерительный инструмент или технические средства, как переносные (деселерометры, стробоскопы, осциллографы и т.д.), так и стационарные (стенды тормозных, тяговых качеств, мотор-тестеры и т.д.).

В зависимости от физической основы средства измерения подразделяются на механические (пружинные деселерометры и т.д.); оптические (дымометры, спектрометры и т.д.); пневматические (пневмомикрометры и т.д.); гидравлические (мессдозы и т.д.); акустические (шумомеры и т.д.); электрические (вольтметры, амперметры и т.д.); специальные (использующие при измерениях некоторые физические эффекты — радиоизотопные методы и т.д.) и комбинированные (большинство современных стендов) средства измерений.

В настоящее время получили широкое применение электрические и комбинированные измерительные приборы и стенды. С их помощью измеряются практически любые параметры автомобиля, его агрегатов и систем.

При эксплуатации автомобильного транспорта приходится проводить различные виды измерений:

- износы деталей, т.е. линейные размеры и массу — при исследованиях надежности автомобилей;
- перемещения, скорости и ускорения, усилия в отдельных узлах и механизмах, напряжения давления — при технической диагностике;
- хронометраж, расходы эксплуатационных материалов при ТО и ТР автомобилей.

Исходя из этого на АТП применяется разнообразное измерительное оборудование. Надежность результатов исследований в значительной степени зависит от точности измерений.

Под *точностью измерений* следует понимать степень соответствия результата измерения действительному значению измеряемой величины.

Наличие такого несоответствия свидетельствует об ошибке измерения, имеющей определенные источники.

Различают четыре основных источника ошибок:

- 1) датчик, который неправильно реагирует на измеряемую величину (неправильная установка, крепление датчика и т.д.);
- 2) измерительное устройство, а именно погрешности из-за неверного функционирования измерительного устройства (наличие износов его механических частей, их несносность, падение напряжения в электрических составляющих и т.д.);

3) сам исследователь, который в силу определенных причин (неопытность, усталость и т.д.) неправильно считывает показания прибора или неверно обрабатывает результаты измерений;

4) внешняя среда (температура, атмосферное давление, влажность, магнитные и электрические поля и т.д.).

## 8.6. Виды ошибок

Различают три класса (вида) ошибок: систематические, случайные и грубые.

**Систематические ошибки.** Они остаются постоянными при повторных измерениях или изменяются по определенному закону (рис. 8.1).

Погрешность меры длины при отклонении температуры от номинала определяется по формуле

$$\Delta l_t = (\alpha \Delta t + \beta \Delta t^2) l_n,$$

где  $\alpha$ ,  $\beta$  — коэффициенты, определяемые совместными измерениями;  $\Delta t$  — отклонение температуры от нормальной;  $l_n$  — длина меры при нормальной температуре.

Определяют систематические ошибки путем эксперимента или вычисляют на основании характеристик измерительных средств, затем полученный результат корректируют.

Различают следующие группы систематических ошибок:

- инструментальные погрешности (равноплечие весы не могут быть равноплечими, так как имеется трение в сочленениях деталей приборов);
- погрешности в результате неправильной установки измерительного устройства;

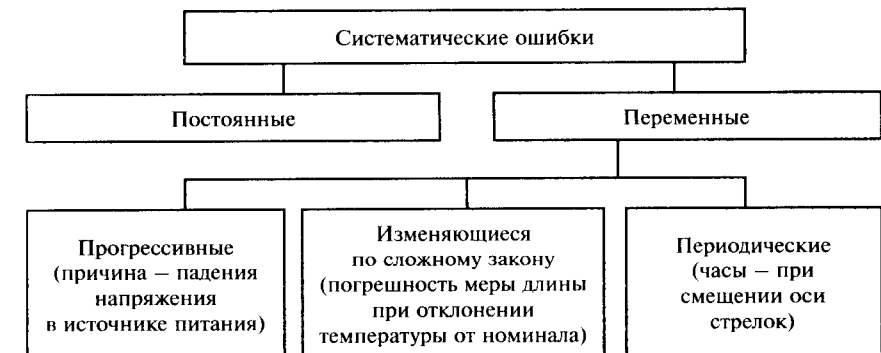


Рис. 8.1. Классификация систематических ошибок (и примеры проявления на практике)

- погрешности, возникшие вследствие внешних влияний;
- погрешности метода измерения (теоретические погрешности);
- субъективные погрешности (погрешности градуировки для шкальных приборов; погрешности в результате старения, износа или неисправности средств измерения (гири — износ)).

Для установки в правильное положение измерительные средства снабжаются уровнями (отвесами, ватерпасами и т. д.), что приводит к ошибкам измерений; внешнее влияние может быть обусловлено тепловыми и воздушными потоками, магнитными и электрическими полями, изменениями атмосферного давления и влажности воздуха и т. д. Особенно опасно, когда эти факторы действуют на протяжении всего периода измерения.

Способы исключения систематических погрешностей:

- 1) профилактика погрешностей (изоляция средств измерения, их ремонт, регулировка и т. д.);
- 2) исключение погрешностей в процессе измерения (способы замещения, компенсации погрешности по знаку, противопоставления, симметричных наблюдений);
- 3) внесение поправок в результат измерения;
- 4) оценка границ систематических погрешностей, если их нельзя исключить (для электрических счетчиков).

**Случайные ошибки.** Случайным называется событие, которое при осуществлении определенного комплекса условий может произойти или может не произойти. При проведении повторных измерений из множества возможных причин отклонение результата измерения от истинного значения также может иметь место или не иметь.

Таким образом, абсолютная случайная погрешность

$$\delta_i = x_i - A,$$

где  $x_i$  — результат единичного измерения;  $A$  — истинное значение измеряемой величины.

Поскольку истинное значение измеряемой величины не всегда известно, то за величину  $A$  можно принять среднее арифметическое:

$$A = \bar{X}^* = \frac{1}{N} \sum_{i=1}^n x_i,$$

где  $N$  — количество измерений.

Тогда абсолютная ошибка единичного измерения или разбора будет равна:

$$\delta_i = x_i - \bar{X}^*.$$

Относительная ошибка измерений определяется отношением абсолютной ошибки к действительному значению измеряемой величины

$$\delta_o = \pm \delta_i / \bar{X}^*.$$

По безразмерной величине относительной ошибки оценивают точность измерения, %:

$$\delta_o = \pm \frac{\delta_i}{\bar{X}^*} 100.$$

Большая часть случайных погрешностей измерений подчиняется нормальному закону распределения.

В теории ошибок есть понятие среднеквадратической ошибки  $S^2$ . Чем меньше среднеквадратическая ошибка, тем больше сходимость результатов измерений:

$$S^2 = \frac{1}{N-1} \sum \delta_i^2 = \sigma^2.$$

Доверительный интервал (интервал оценивания) определяется по формуле

$$E = \pm tS = \pm t \sqrt{\frac{\sum (x_i - \bar{X}^*)^2}{N-1}},$$

где  $t$  — обращенное значение функции Лапласа.

Если случайные погрешности результатов отдельных наблюдений подчиняются нормальному закону распределения, то и погрешности средних значений этих наблюдений также подчиняются этому распределению, но с другим рассеянием.

Теория дает следующее выражение среднеквадратического отклонения  $\sigma_o$  для среднего значения:

$$\sigma_o = \sigma / \sqrt{N}.$$

При статистическом методе

$$S_o = \frac{S}{\sqrt{N}} = \sqrt{\frac{\sum (x_i - \bar{X}^*)^2}{N(N-1)}}.$$

Доверительный интервал для среднего определяется по формуле

$$E_o = \pm t \sqrt{\frac{\sum (x_i - \bar{X}^*)^2}{N(N-1)}}.$$

#### Вероятность попадания случайных погрешностей в заданный интервал

Величина ошибки в долях $\sigma$ .....	0,5	0,675	1,0	2,0	3,0
Вероятность того, что случайные погрешности не выйдут за указанные интервалы .....	0,379	0,500	0,682	0,95	0,999

Практически случайные ошибки наблюдений устраняют путем увеличения числа опытов (измерений), при этом расчеты проводят по среднему.

**Грубые ошибки.** Грубые погрешности (промахи) существенно превышают систематические и случайные ошибки и являются, как правило, следствием ошибок экспериментатора. В расчеты данный вид ошибок не принимается.

Для исключения грубых ошибок используют критерий Стьюдента:

$$\frac{Y_{\min}^{\max} - \bar{Y}}{S_y} = t \begin{cases} \text{Если } t \geq t_{\alpha} \left( \begin{matrix} \alpha \\ f = N - 1 \end{matrix} \right), \text{ то результат исключается;} \\ \text{Если } t \leq t_{\alpha} \left( \begin{matrix} \alpha \\ f = N - 1 \end{matrix} \right), \text{ то результат не исключается,} \end{cases}$$

где  $Y_{\min}^{\max}$  — результат, который ставится под сомнение;  $\bar{Y}$  — среднее значение, вычисляемое без  $Y_{\min}^{\max}$ .

Среднеквадратическое отклонение  $S_y$  вычисляется без  $Y_{\min}^{\max}$  по формуле

$$S_y = \sqrt{\frac{\sum (Y_i - \bar{Y})^2}{N - 1}}.$$

### КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Дайте определение понятий «физической величины», «измерения».
2. Расскажите о технических средствах измерений и приведите их классификацию.
3. Назовите виды измерений, проводящихся в области эксплуатации автомобильного транспорта.
4. Дайте определение понятия «точность измерения», перечислите основные источники ошибок измерений.
5. Дайте характеристику систематических ошибок намерений, приведите их классификацию.
6. Назовите способы исключения систематических ошибок.
7. Дайте характеристику случайных ошибок измерений, назовите способы их устранения.
8. Дайте определения средней квадратической ошибки и интервала оценивания.
9. Опишите процедуру определения случайной ошибки по среднему.
10. Дайте определение понятия «грубая ошибка», приведите расчетные формы.

11. Что такое метрологическое обеспечение? Какова его научная, техническая и организационная основа.

12. Назовите пути обеспечения успешного решения основной задачи метрологического обеспечения.

13. Дайте определение понятия «поверка СИ», приведите классификацию способов поверки СИ.

14. Каковы сроки межповерочных интервалов СИ?

15. Дайте краткую характеристику процедуры нормирования метрологических характеристик СИ.

## ЗАКЛЮЧЕНИЕ

В данном учебнике рассмотрены основы и приведены примеры проведения работ в условиях пассивного и активного экспериментов в области технической эксплуатации автомобилей, достаточно широко представлены отдельные вопросы подготовки и обработки результатов научных исследований с возможностью использования популярной программы STATISTICA для среды WINDOWS. Эти методы и приемы в дальнейшем будут развиваться при изучении общеинженерных и профессиональных дисциплин.

Для изучающих изложенный курс важно освоить особенности и принципиальные подходы, которые необходимо использовать при работе со случайными величинами. Это послужит фундаментом развития полученных знаний и поможет современному инженеру в области технической эксплуатации автомобильного транспорта в дальнейшем эффективно применять в практической деятельности современные математические методы, такие как корреляционный, регрессионный, дисперсионный и компонентный анализы. Освоение указанных методов, в свою очередь, является ступенью к изучению и применению на практике факторного анализа, динамического программирования, методов исследования операций, сетевого планирования и управления и других перспективных математических инструментариев, способствующих выработке и принятию рациональных инженерных решений.

Данный учебник базируется на основах системного подхода и системного анализа и способствует лучшему пониманию изучаемых процессов и явлений.

## ПРИЛОЖЕНИЯ

### Приложение 1

$$\text{Нормированная функция Лапласа } \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{t^2}{2}} dt$$

z	Сотые доли для z									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	040	080	120	160	199	239	279	319	359
0,1	398	438	478	517	557	596	636	675	714	753
0,2	793	832	871	910	948	987	•026	•064	•103	•141
0,3	0,1179	217	255	293	331	368	406	443	480	517
0,4	554	591	628	664	700	736	772	808	844	879
0,5	915	950	985	•019	•054	•088	•123	•157	•190	•224
0,6	0,2257	291	324	357	389	422	454	486	517	549
0,7	580	611	642	673	703	734	764	794	823	852
0,8	881	910	939	967	995	•023	•051	•078	•106	•133
0,9	0,3159	186	212	238	261	289	315	340	365	389
1,0	413	437	461	485	508	583	554	577	599	621
1,1	643	665	686	708	729	749	770	790	810	830
1,2	849	869	888	907	925	944	962	980	997	•015
1,3	0,4032	049	066	082	099	115	131	147	162	177
1,4	192	207	222	236	251	265	279	292	306	319
1,5	332	345	357	370	382	394	406	418	429	441
1,6	452	463	474	484	495	505	515	525	535	545
1,7	554	564	573	582	591	599	608	616	625	633
1,8	641	649	656	664	671	678	686	693	699	706
1,9	713	719	726	732	738	744	750	756	761	767
2,0	772	778	783	788	793	798	803	808	812	817

z	Сотые доли для z									
	0	1	2	3	4	5	6	7	8	9
2,1	821	82,6	830	834	838	842	846	850	854	857
2,2	860	864	867	871	874	877	880	883	886	889
2,3	892	895	898	900	903	906	908	911	913	915
2,4	918	920	922	924	926	928	930	932	934	936
2,5	937	939	941	942	944	946	947	949	950	952
2,6	953	955	956	957	958	959	960	962	963	964
2,7	965	966	967	968	969	970	971	971	972	973
2,8	974	975	975	976	977	978	978	979	980	980
2,9	981	981	982	983	983	984	984	985	985	986
3,0	986	986	987	987	988	988	988	989	989	989

Значение верхнего  $q$  предела  $\chi_q^2$  в зависимости от вероятности

$$P(\chi^2 > \chi_q^2) = \frac{1}{\Gamma\left(\frac{r}{2}\right) 2^{\frac{r}{2}} x_q^{\frac{r}{2}}} \int_{x_q^2}^{\infty} x^{\frac{r}{2}-1} e^{-\frac{x}{2}} dx$$

и числа  $r$  степеней свободы  $\chi^2$ -распределения  
(распределения Пирсона)

Число степеней свободы $r$	Вероятность $P(\chi^2 > \chi_q^2)$													
	0,975	0,95	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10	0,05	0,025	
1	0,001	0,004	0,016	0,064	0,148	0,275	0,455	0,708	1,07	1,64	2,7	3,8	5,0	
2	0,051	0,103	0,211	0,446	0,713	1,023	1,386	1,84	2,41	3,22	4,6	6,0	7,4	
3	0,213	0,352	0,584	1,005	1,424	1,87	2,37	2,94	3,66	4,64	6,3	7,8	9,3	
4	0,484	0,71	1,06	1,65	2,19	2,76	3,36	4,04	4,9	6,0	7,8	9,5	11,1	
5	0,83	1,14	1,61	2,34	3,00	3,66	4,35	5,1	6,1	7,3	9,2	11,1	12,8	
6	1,24	1,63	2,20	3,07	3,83	4,56	5,35	6,2	7,2	8,6	10,6	12,6	14,4	
7	1,69	2,17	2,83	3,82	4,67	5,49	6,35	7,3	8,4	9,8	12,0	14,1	16,0	
8	2,18	2,73	3,49	4,59	5,53	6,43	7,34	8,4	9,5	11,0	13,4	15,5	17,5	
9	2,70	3,32	4,17	5,38	6,39	7,34	8,34	9,4	10,7	12,2	14,7	16,9	19,0	
10	3,25	3,94	4,86	6,18	7,27	8,29	9,34	10,5	11,8	13,4	16,0	18,3	20,5	
11	3,81	4,6	5,6	7,0	8,1	9,24	10,3	11,5	12,9	14,6	17,3	19,7	21,9	
12	4,40	5,2	6,3	7,8	9,0	10,2	11,3	12,6	14,0	15,8	18,5	21,0	23,3	
13	5,0	5,9	7,0	8,6	9,9	11,1	12,3	13,6	15,1	17,0	19,8	22,4	24,7	
14	5,6	6,6	7,8	9,5	10,8	12,1	13,3	14,7	16,2	18,2	21,1	23,7	26,1	
15	6,3	7,3	8,5	10,3	11,7	13,0	14,3	15,7	17,3	19,3	22,3	25,0	27,5	
16	6,9	8,0	9,3	11,2	12,6	14,0	15,3	16,8	18,4	20,5	23,5	26,3	28,8	
17	7,6	8,7	10,1	12,0	13,5	14,9	16,3	17,8	19,5	21,6	24,8	27,6	30,2	
18	8,2	9,4	10,9	12,9	14,4	15,9	17,3	18,9	20,6	22,8	26,0	28,9	31,5	
19	8,9	10,1	11,7	13,7	15,4	16,9	18,3	19,9	21,7	23,9	27,2	30,1	32,9	

Продолжение таблицы

Число степеней свободы $r$	Вероятность $P(\chi^2 > \chi^2_{\alpha})$												
	0,975	0,95	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10	0,05	0,025
20	9,6	10,9	12,4	14,6	16,3	17,8	19,3	20,9	22,8	25,0	28,4	31,4	34,2
21	10,3	11,6	13,2	15,4	17,2	18,8	20,3	22,0	23,9	26,2	29,6	32,7	35,5
22	11,0	12,3	14,0	16,3	18,1	19,7	21,3	23,0	24,9	27,3	30,8	33,9	36,8
23	11,7	13,1	14,8	17,2	19,0	20,7	22,3	24,1	26,0	28,4	32,0	35,2	38,1
24	12,4	13,8	15,7	18,1	19,9	21,7	23,3	25,1	27,1	29,6	33,2	36,4	39,4
25	13,1	14,6	16,5	18,9	20,9	22,6	24,3	26,1	28,1	30,7	34,4	37,7	40,6
26	13,8	15,4	17,3	19,8	21,8	23,6	25,3	27,2	29,3	31,8	35,6	38,9	41,9
27	14,6	16,2	18,1	20,7	22,7	24,5	26,3	28,2	30,3	32,9	36,7	40,1	43,2
28	15,3	16,9	18,9	21,6	23,6	25,5	27,3	29,3	31,4	34,0	37,9	41,3	44,5
29	16,0	17,7	19,8	22,5	24,6	26,5	28,3	30,3	32,5	35,1	39,1	42,6	45,7
30	16,8	18,5	20,6	23,4	25,5	27,4	29,3	31,3	33,5	36,3	40,3	43,8	47,0
32	18,3	20,1	21,3	25,1	27,4	29,4	31,3	33,4	35,7	38,5	42,6	46,2	49,5
34	19,8	21,7	24,0	26,9	29,2	31,3	33,3	35,4	37,8	40,7	44,9	48,6	52,0
36	21,3	23,3	25,6	28,7	31,1	33,2	35,3	37,5	39,9	42,6	47,2	51,0	54,4
38	22,9	24,9	27,3	30,5	33,0	35,2	37,3	39,6	42,0	45,1	49,5	53,4	56,9
40	24,4	26,5	29,1	32,4	34,9	37,1	39,3	41,6	44,2	47,3	51,8	55,8	59,3
42	26,0	28,1	30,8	34,2	36,8	39,1	41,3	43,7	46,2	49,5	54,1	58,1	61,8
44	27,6	29,8	32,5	36,0	38,6	41,0	43,3	45,7	48,4	51,6	56,4	60,5	64,2
46	29,2	31,4	34,2	37,8	40,5	43,0	45,3	47,8	50,5	52,8	58,6	62,8	66,6
48	30,8	33,1	36,0	39,6	42,4	44,9	46,3	49,8	52,6	56,0	60,9	65,2	69,0
50	32,4	34,8	37,7	41,4	44,3	46,9	49,3	51,9	54,7	58,2	63,2	67,5	71,4
52	34,0	36,4	39,4	43,3	46,2	48,8	51,3	53,9	56,8	60,3	65,4	69,8	73,8
54	35,6	38,1	41,2	45,1	48,1	50,8	53,3	56,0	58,9	62,5	67,7	72,2	76,2
56	37,2	39,8	42,9	47,0	50,0	52,7	55,3	58,0	61,0	64,7	69,9	74,5	78,6
58	38,8	41,5	44,7	48,8	51,9	54,7	57,3	60,1	63,1	66,8	72,2	76,8	80,1

Продолжение таблицы

Число степеней свободы $r$	Вероятность $P(\chi^2 > \chi^2_{\alpha})$												
	0,975	0,95	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10	0,05	0,025
60	40,5	43,2	46,5	50,6	53,8	56,2	59,3	62,1	65,2	69,0	74,4	79,1	83,3
62	42,1	44,9	48,2	52,5	55,7	58,6	61,3	64,2	67,3	71,1	76,6	81,4	85,7
64	43,8	46,6	50,0	54,3	57,6	60,5	63,3	66,2	69,4	73,3	78,9	83,7	88,0
66	45,4	48,3	51,8	56,2	59,5	62,5	65,3	68,3	71,5	75,4	81,1	86,0	90,3
68	47,1	50,0	53,5	58,0	61,4	64,4	67,4	70,3	73,8	77,6	83,8	88,2	92,7
70	48,8	51,7	55,3	59,9	63,3	66,4	69,3	72,4	75,7	79,7	85,5	90,5	95,0
72	50,4	53,5	57,1	61,8	65,3	68,3	71,3	74,4	77,8	81,9	87,7	92,8	97,4
74	52,1	55,2	58,9	63,6	67,2	70,3	73,3	76,4	79,9	84,0	90,0	95,7	100,0
76	53,9	56,9	60,7	65,5	69,1	72,3	75,3	78,5	81,9	85,8	92,2	97,4	102,0
78	55,5	58,7	62,5	67,3	71,0	74,2	77,3	80,5	84,0	88,3	94,4	99,6	104,3
80	57,2	60,4	64,3	69,2	72,9	76,2	79,3	82,6	86,2	90,4	96,6	101,9	106,6
82	58,8	62,1	66,1	71,1	74,8	78,2	81,3	84,8	88,2	92,5	98,8	104,1	108,9
84	60,5	63,9	67,9	72,9	76,8	80,1	83,3	86,7	90,3	94,7	101,0	106,4	111,2
86	62,2	65,7	69,7	74,8	78,7	82,1	85,3	88,7	92,4	96,8	103,2	108,7	113,5
88	63,9	67,4	71,5	76,7	80,6	84,0	87,3	90,7	94,4	98,9	105,4	110,9	115,5
90	65,6	69,1	73,3	78,6	82,5	86,0	89,3	92,8	96,5	101,1	107,6	113,2	118,1
92	67,4	70,9	75,1	80,4	84,4	88,0	91,3	94,8	98,6	103,2	109,8	115,4	120,4
94	69,1	72,6	76,9	82,3	86,4	89,9	93,3	96,8	100,7	105,3	111,9	117,6	122,7
96	70,8	74,7	78,7	84,2	88,3	91,9	95,3	98,9	102,8	107,4	114,1	119,9	125,0
98	72,5	76,2	80,5	86,1	90,2	93,8	97,3	100,9	104,8	109,6	116,3	122,1	127,3
100	74,2	77,9	82,4	87,9	92,1	95,8	99,3	103,0	106,9	111,7	118,5	124,3	129,6
110	82,9	86,8	91,5	97,4	101,8	105,6	109,3	113,1	117,3	121,9	129,4	135,5	140,9
120	91,6	95,7	100,6	106,8	111,4	115,5	119,3	123,3	127,6	132,8	140,2	146,6	152,2
130	100,3	101,7	109,8	116,3	121,1	125,3	129,3	133,6	137,9	151,1	143,3	157,6	163,5
140	109,1	113,7	119,0	125,8	130,8	135,1	139,3	143,6	148,3	153,9	161,8	168,6	174,6

Число степеней свободы $r$	Вероятность $P(\chi^2 > \chi_q^2)$												
	0,975	0,95	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10	0,05	0,025
150	118,0	122,7	128,3	135,2	140,5	144,1	149,3	153,7	158,6	164,4	172,6	179,6	185,8
200	162,7	168,3	174,8	183,0	189,0	194,3	199,3	204,4	210,0	216,6	226,0	233,9	241,1
300	253,9	260,9	260,1	279,2	288,7	293,2	299,3	305,6	312,3	320,4	331,8	341,4	349,5
400	346,5	354,6	364,2	376,0	384,7	392,2	399,3	406,5	414,3	423,6	436,7	447,6	457,3
500	439,9	449,2	459,9	472,2	483,0	491,4	499,3	507,4	516,1	526,4	540,9	553,1	563,8
600	534,0	544,2	556,1	570,7	581,4	590,6	599,3	608,1	617,7	628,9	644,8	658,1	669,8
800	723,5	735,4	749,5	766,2	778,6	789,3	799,3	809,5	820,5	833,5	851,7	866,9	880,3
1000	914,3	927,6	943,1	962,2	976,1	988,0	999,3	1010	1023	1037	1056	1075	1090

Значение  $q$  процентных пределов  $t_{k,q}$  (для уровня значимости  $q$ ) в зависимости от  $k$  степеней свободы и от вероятности

$$\frac{q}{100} = 2 \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right) \sqrt{\pi k}} \int_{t_{q,k}}^{\infty} \left(1 + \frac{t^2}{k}\right)^{-\frac{k+1}{2}} dt$$

и их соответствие квантилям односторонней доверительной

вероятности  $P = \left(1 - \frac{q}{2 \cdot 100}\right)$  для распределения Стьюдента

$k$	Значение $t_{q,k}$ при уровне значимости $q$ , %							
	20,0	10,0	5,0	2,0	1,0	0,5	0,2	0,1
	Квантили $t$ -распределения при $P$							
	0,90	0,95	0,975	0,990	0,995	0,9975	0,999	0,9995
1	3,0777	6,3138	12,7062	31,8205	63,6567	127,3213	318,3088	636,6192
2	1,8856	2,9200	4,3027	6,9646	9,9248	14,0890	22,3271	31,5991
3	1,6377	1,3534	3,1824	4,5407	5,8409	7,4533	10,2145	12,9240
4	1,5332	1,1318	2,7764	3,7469	4,6041	5,5976	7,1732	8,6103
5	1,4759	1,0150	2,5706	3,3649	4,0321	4,7733	5,8934	6,8688
6	1,4398	1,9432	2,4469	3,1427	3,7074	4,3168	5,2076	5,9588
7	1,4149	1,8946	2,3646	2,9980	3,4995	4,0293	4,7863	5,4079
8	1,3968	1,8595	2,3060	2,8965	3,3554	3,8325	4,5008	5,0413
9	1,3830	1,8331	2,2622	2,8214	3,2498	3,6897	4,2968	4,7809
10	1,3722	1,8125	2,2281	2,7638	3,1693	3,5814	4,1437	4,5869
11	1,3634	1,7959	2,2010	2,7181	3,1058	3,4966	4,0247	4,4370
12	1,3562	1,7823	2,1788	2,6810	3,0545	3,4284	3,9296	4,3178
13	1,3502	1,7709	2,1604	2,6503	3,0123	3,3725	3,8520	4,2208
14	1,3450	1,7613	2,1448	2,6245	2,9768	3,3257	3,7874	4,1405
15	1,3406	1,7530	2,1314	2,6025	2,9467	3,2860	3,7328	4,0728
16	1,3368	1,7459	2,1199	2,5835	2,9208	3,2520	3,6862	4,0150
17	1,3334	1,7396	2,1098	2,5669	2,8982	3,2224	3,6458	3,9651



Продолжение таблицы

k	Значение $t_{q,k}$ при уровне значимости $q$ , %							
	20,0	10,0	5,0	2,0	1,0	0,5	0,2	0,1
	Квантили $t$ -распределения при $P$							
	0,90	0,95	0,975	0,990	0,995	0,9975	0,999	0,9995
18	1,3304	1,7341	2,1009	2,5524	2,8784	3,1966	3,6105	3,9216
19	1,3277	1,7291	2,0930	2,5395	2,8609	3,1737	3,5794	3,8834
20	1,3253	1,7247	2,0860	2,5280	2,8453	3,1534	3,5518	3,8495
21	1,3232	1,7207	2,0796	2,5176	2,8314	3,1352	3,5272	3,8193
22	1,3212	1,7171	2,0739	2,5083	2,8188	3,1188	3,5050	3,7921
23	1,3195	1,7139	2,0687	2,4999	2,8073	3,1040	3,4850	3,7676
24	1,3178	1,7109	2,0639	2,4922	2,7969	3,0905	3,4668	3,7455
25	1,3163	1,7081	2,0595	2,4851	2,7874	3,0782	3,4502	3,7251
26	1,3150	1,7056	2,0555	2,4786	2,7787	3,0669	3,4350	3,7066
27	1,3137	1,7033	2,0518	2,4727	2,7707	3,0565	3,4210	3,6896
28	1,3125	1,7011	2,0484	2,4671	2,7633	3,0469	3,4082	3,6739
29	1,3114	1,6991	2,04521	2,4620	2,7564	3,0380	3,3962	3,6594
30	1,3104	1,6973	2,0423	2,4573	2,7500	3,0298	3,3852	3,6400
32	1,3086	1,6939	2,0369	2,4487	2,7385	3,0149	3,3653	3,6218
34	1,3070	1,6909	2,0322	2,4411	2,7284	3,0020	3,3479	3,6007
36	1,3055	1,6883	2,0281	2,4345	2,7195	2,9905	3,3326	3,5821
38	1,3042	1,6860	2,0244	2,4286	2,7116	2,9803	3,3190	3,5657
40	1,3031	1,6839	2,0211	2,4233	2,7045	2,9712	3,3069	3,5510
42	1,3020	1,6820	2,0181	2,4185	2,6981	2,9630	3,2960	3,5377
44	1,3011	1,6802	2,0154	2,4141	2,6923	2,9555	3,2861	3,5258
46	1,3002	1,6787	2,0129	2,4102	2,6870	2,9488	3,2771	3,5150
48	1,2994	1,6772	2,0106	2,4066	2,6822	2,9426	3,2689	3,5051
50	1,2987	1,6759	2,0086	2,4033	2,6778	2,9370	3,2614	3,4960
55	1,2971	1,6730	2,0040	2,3961	2,6682	2,9247	3,2561	3,4764
60	1,2958	1,6706	2,0003	2,3901	2,6603	2,9146	3,2317	3,4602

Окончание таблицы

k	Значение $t_{q,k}$ при уровне значимости $q$ , %							
	20,0	10,0	5,0	2,0	1,0	0,5	0,2	0,1
	Квантили $t$ -распределения при $P$							
	0,90	0,95	0,975	0,990	0,995	0,9975	0,999	0,9995
65	1,2947	1,6686	1,9971	2,3851	2,6536	2,9060	3,2204	3,4466
70	1,2938	1,6669	1,9944	2,3808	2,6479	2,8987	3,2108	3,4350
80	1,2922	1,6641	1,9901	2,3739	2,6387	2,8870	3,1953	3,4163
90	1,2910	1,6620	1,9867	2,3685	2,6316	2,8779	3,1833	3,4019
100	1,2901	1,6602	1,9840	2,3642	2,6259	2,8707	3,1737	3,3905
120	1,2886	1,6577	1,9799	2,3578	2,6174	2,8599	3,1595	3,1735
150	1,2872	1,6551	1,9759	2,3515	2,6090	2,8492	3,1455	3,3566
200	1,2858	1,6525	1,9719	2,3451	2,6006	2,8385	3,1315	3,3398
250	1,2849	1,6510	1,9695	2,3414	2,5956	2,8322	3,1232	3,3299
300	1,2844	1,6499	1,9679	2,3388	2,5923	2,8279	3,1176	3,3233
400	1,2837	1,6487	1,9659	2,3357	2,5882	2,8227	3,1107	3,3150
500	1,2832	1,6479	1,9647	2,3338	2,5857	2,8195	3,1066	3,3101
$\infty$	0,2816	1,6449	1,9600	2,3263	2,5758	2,8070	3,0902	3,2905

Значения пяти- и однопроцентных верхних пределов уклонения  
свободы

$k_2$ — степени свободы для меньшей дисперсии	$k_1$ — степени свободы											
	1	2	3	4	5	6	7	8	9	10	11	12
1	161	200	216	225	230	234	237	239	241	242	243	244
	<b>4052</b>	<b>4999</b>	<b>5403</b>	<b>5625</b>	<b>5764</b>	<b>5889</b>	<b>5928</b>	<b>5981</b>	<b>6022</b>	<b>6056</b>	<b>6082</b>	<b>6106</b>
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39	19,40	19,41
	<b>98,49</b>	<b>99,01</b>	<b>99,17</b>	<b>99,25</b>	<b>99,30</b>	<b>99,33</b>	<b>99,34</b>	<b>99,36</b>	<b>99,38</b>	<b>99,40</b>	<b>99,41</b>	<b>99,42</b>
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,76	8,74
	<b>34,12</b>	<b>30,81</b>	<b>29,46</b>	<b>28,71</b>	<b>28,24</b>	<b>27,91</b>	<b>27,67</b>	<b>27,49</b>	<b>27,34</b>	<b>27,23</b>	<b>27,13</b>	<b>27,05</b>
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93	5,91
	<b>21,20</b>	<b>18,00</b>	<b>16,69</b>	<b>15,98</b>	<b>15,52</b>	<b>15,21</b>	<b>14,98</b>	<b>14,80</b>	<b>14,66</b>	<b>14,54</b>	<b>14,45</b>	<b>14,37</b>
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,70	4,68
	<b>16,26</b>	<b>13,27</b>	<b>12,06</b>	<b>11,39</b>	<b>10,97</b>	<b>10,67</b>	<b>10,45</b>	<b>10,27</b>	<b>10,15</b>	<b>10,05</b>	<b>9,96</b>	<b>9,89</b>
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00
	<b>13,7</b>	<b>10,92</b>	<b>9,78</b>	<b>9,15</b>	<b>8,75</b>	<b>8,47</b>	<b>8,26</b>	<b>8,10</b>	<b>7,98</b>	<b>7,87</b>	<b>7,79</b>	<b>7,72</b>
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,60	3,57
	<b>12,25</b>	<b>9,55</b>	<b>8,45</b>	<b>7,85</b>	<b>7,46</b>	<b>7,19</b>	<b>7,00</b>	<b>6,84</b>	<b>6,71</b>	<b>6,62</b>	<b>6,54</b>	<b>6,47</b>
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,31	3,28
	<b>11,26</b>	<b>8,65</b>	<b>7,59</b>	<b>7,01</b>	<b>6,63</b>	<b>6,37</b>	<b>6,19</b>	<b>6,03</b>	<b>5,91</b>	<b>5,82</b>	<b>5,74</b>	<b>5,67</b>
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,10	3,07
	<b>10,56</b>	<b>8,02</b>	<b>6,99</b>	<b>6,42</b>	<b>6,06</b>	<b>5,80</b>	<b>5,62</b>	<b>5,47</b>	<b>5,35</b>	<b>5,26</b>	<b>5,18</b>	<b>5,11</b>
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,94	2,91
	<b>10,04</b>	<b>7,56</b>	<b>6,55</b>	<b>5,99</b>	<b>5,64</b>	<b>5,39</b>	<b>5,21</b>	<b>5,06</b>	<b>4,95</b>	<b>4,85</b>	<b>4,78</b>	<b>4,71</b>

Приложение 4

величины  $F$  (распределение Фишера) в зависимости от степеней  
 $k_1$  и  $k_2^*$

для большей дисперсии											
14	16	20	24	30	40	50	75	100	200	500	$\infty$
245	246	248	249	250	251	252	253	253	254	254	254
<b>6142</b>	<b>6169</b>	<b>6208</b>	<b>6234</b>	<b>6258</b>	<b>6286</b>	<b>6302</b>	<b>6323</b>	<b>6334</b>	<b>6352</b>	<b>6361</b>	<b>6366</b>
19,42	19,43	19,44	19,45	19,46	19,47	19,47	19,48	19,49	19,49	19,50	19,50
<b>99,43</b>	<b>99,44</b>	<b>99,45</b>	<b>99,46</b>	<b>99,47</b>	<b>99,48</b>	<b>99,48</b>	<b>99,49</b>	<b>99,49</b>	<b>99,49</b>	<b>99,50</b>	<b>99,50</b>
8,712	8,69	8,66	8,64	8,62	8,60	8,58	8,57	8,56	8,54	8,54	8,53
<b>6,92</b>	<b>26,83</b>	<b>26,69</b>	<b>26,60</b>	<b>26,50</b>	<b>26,41</b>	<b>26,35</b>	<b>26,27</b>	<b>26,23</b>	<b>26,18</b>	<b>26,14</b>	<b>26,12</b>
5,871	5,84	5,80	5,77	5,74	5,71	5,70	5,68	5,66	5,65	5,64	5,63
<b>4,24</b>	<b>14,15</b>	<b>14,02</b>	<b>13,93</b>	<b>13,83</b>	<b>13,74</b>	<b>13,69</b>	<b>13,61</b>	<b>13,57</b>	<b>13,52</b>	<b>13,48</b>	<b>13,46</b>
4,64	4,60	4,56	4,53	4,50	4,46	4,44	4,42	4,40	4,38	4,37	4,36
<b>9,77</b>	<b>9,68</b>	<b>9,55</b>	<b>9,47</b>	<b>9,38</b>	<b>9,29</b>	<b>9,24</b>	<b>9,17</b>	<b>9,13</b>	<b>9,07</b>	<b>9,04</b>	<b>9,02</b>
3,96	3,92	3,87	3,84	3,81	3,77	3,75	3,72	3,71	3,69	3,68	3,67
<b>7,60</b>	<b>7,52</b>	<b>7,39</b>	<b>7,31</b>	<b>7,23</b>	<b>7,14</b>	<b>7,09</b>	<b>7,02</b>	<b>6,99</b>	<b>6,94</b>	<b>6,90</b>	<b>6,88</b>
3,52	3,49	3,44	3,41	3,38	3,34	3,32	3,29	3,28	3,25	3,24	3,23
<b>6,35</b>	<b>6,27</b>	<b>6,15</b>	<b>6,07</b>	<b>5,98</b>	<b>5,90</b>	<b>5,85</b>	<b>5,78</b>	<b>5,75</b>	<b>5,70</b>	<b>5,67</b>	<b>5,65</b>
3,23	3,20	3,15	3,12	3,08	3,05	3,03	3,00	2,98	2,96	2,94	2,93
<b>5,56</b>	<b>5,48</b>	<b>5,36</b>	<b>5,28</b>	<b>5,20</b>	<b>5,11</b>	<b>5,06</b>	<b>5,00</b>	<b>4,96</b>	<b>4,91</b>	<b>4,88</b>	<b>4,86</b>
3,02	2,98	2,93	2,90	2,86	2,82	2,80	2,77	2,76	2,73	2,72	2,71
<b>5,00</b>	<b>4,92</b>	<b>4,80</b>	<b>4,73</b>	<b>4,64</b>	<b>4,56</b>	<b>4,51</b>	<b>4,45</b>	<b>4,41</b>	<b>4,36</b>	<b>4,33</b>	<b>4,31</b>
2,86	2,82	2,77	2,74	2,70	2,67	2,64	2,61	2,59	2,56	2,55	2,54
<b>4,60</b>	<b>4,52</b>	<b>4,41</b>	<b>4,33</b>	<b>4,25</b>	<b>4,17</b>	<b>4,12</b>	<b>4,05</b>	<b>4,01</b>	<b>3,96</b>	<b>3,93</b>	<b>3,91</b>

$k_2$ — степени свободы для меньшей дисперсии	$k_1$ — степени свободы											
	1	2	3	4	5	6	7	8	9	10	11	12
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,82	2,79
	<b>9,85</b>	<b>7,20</b>	<b>6,22</b>	<b>5,67</b>	<b>5,32</b>	<b>5,07</b>	<b>4,88</b>	<b>4,74</b>	<b>4,63</b>	<b>4,54</b>	<b>4,46</b>	<b>4,40</b>
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,72	2,69
	<b>9,33</b>	<b>6,93</b>	<b>5,95</b>	<b>5,41</b>	<b>5,06</b>	<b>4,82</b>	<b>4,65</b>	<b>4,50</b>	<b>4,39</b>	<b>4,30</b>	<b>4,22</b>	<b>4,16</b>
13	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,63	2,60
	<b>9,07</b>	<b>6,70</b>	<b>5,74</b>	<b>5,20</b>	<b>4,86</b>	<b>4,62</b>	<b>4,44</b>	<b>4,30</b>	<b>4,19</b>	<b>4,10</b>	<b>4,02</b>	<b>3,96</b>
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,56	2,53
	<b>8,86</b>	<b>6,51</b>	<b>5,56</b>	<b>5,03</b>	<b>4,69</b>	<b>4,46</b>	<b>4,28</b>	<b>4,14</b>	<b>4,03</b>	<b>3,94</b>	<b>3,86</b>	<b>3,80</b>
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,51	2,48
	<b>8,68</b>	<b>6,36</b>	<b>5,42</b>	<b>4,89</b>	<b>4,56</b>	<b>4,32</b>	<b>4,14</b>	<b>4,00</b>	<b>3,89</b>	<b>3,80</b>	<b>3,73</b>	<b>3,67</b>
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42
	<b>8,53</b>	<b>6,23</b>	<b>5,29</b>	<b>4,77</b>	<b>4,44</b>	<b>4,20</b>	<b>4,03</b>	<b>3,89</b>	<b>3,78</b>	<b>3,69</b>	<b>3,61</b>	<b>3,55</b>
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,41	2,38
	<b>8,40</b>	<b>6,11</b>	<b>5,18</b>	<b>4,67</b>	<b>4,34</b>	<b>4,10</b>	<b>3,93</b>	<b>3,79</b>	<b>3,68</b>	<b>3,59</b>	<b>3,52</b>	<b>3,45</b>
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34
	<b>8,28</b>	<b>6,01</b>	<b>5,09</b>	<b>4,58</b>	<b>4,25</b>	<b>4,01</b>	<b>3,85</b>	<b>3,71</b>	<b>3,60</b>	<b>3,51</b>	<b>3,44</b>	<b>3,37</b>
19	4,38	3,52	3,13	2,90	2,74	2,63	2,55	2,48	2,43	2,38	2,34	2,31
	<b>8,18</b>	<b>5,93</b>	<b>5,01</b>	<b>4,50</b>	<b>4,17</b>	<b>3,94</b>	<b>3,77</b>	<b>3,63</b>	<b>3,52</b>	<b>3,43</b>	<b>3,36</b>	<b>3,30</b>
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	2,31	2,28
	<b>8,10</b>	<b>5,85</b>	<b>4,94</b>	<b>4,43</b>	<b>4,10</b>	<b>3,87</b>	<b>3,71</b>	<b>3,56</b>	<b>3,45</b>	<b>3,37</b>	<b>3,30</b>	<b>3,23</b>
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28	2,25
	<b>8,02</b>	<b>5,78</b>	<b>4,87</b>	<b>4,37</b>	<b>4,04</b>	<b>3,81</b>	<b>3,65</b>	<b>3,51</b>	<b>3,40</b>	<b>3,31</b>	<b>3,24</b>	<b>3,17</b>

для большей дисперсии											
14	16	20	24	30	40	50	75	100	200	500	$\infty$
2,74	2,70	2,65	2,61	2,57	2,53	2,50	2,47	2,45	2,42	2,41	2,40
<b>4,29</b>	<b>4,21</b>	<b>4,10</b>	<b>4,02</b>	<b>3,94</b>	<b>3,86</b>	<b>3,80</b>	<b>3,74</b>	<b>3,70</b>	<b>3,66</b>	<b>3,62</b>	<b>3,60</b>
2,64	2,60	2,54	2,50	2,46	2,42	2,40	2,36	2,35	2,32	2,31	2,30
<b>4,05</b>	<b>3,98</b>	<b>3,86</b>	<b>3,78</b>	<b>3,70</b>	<b>3,61</b>	<b>3,56</b>	<b>3,49</b>	<b>3,46</b>	<b>3,41</b>	<b>3,38</b>	<b>3,36</b>
2,55	2,51	2,46	2,42	2,38	2,34	2,32	2,28	2,26	2,24	2,22	2,21
<b>3,85</b>	<b>3,78</b>	<b>3,67</b>	<b>3,59</b>	<b>3,51</b>	<b>3,42</b>	<b>3,37</b>	<b>3,30</b>	<b>3,27</b>	<b>3,21</b>	<b>3,18</b>	<b>3,16</b>
2,48	2,44	2,39	2,35	2,31	2,27	2,24	2,21	2,19	2,16	2,14	2,13
<b>3,70</b>	<b>3,62</b>	<b>3,51</b>	<b>3,43</b>	<b>3,34</b>	<b>3,26</b>	<b>3,21</b>	<b>3,14</b>	<b>3,11</b>	<b>3,06</b>	<b>3,02</b>	<b>3,00</b>
2,43	2,39	2,33	2,29	2,25	2,21	2,18	2,15	2,12	2,10	2,08	2,07
<b>3,56</b>	<b>3,48</b>	<b>3,36</b>	<b>3,29</b>	<b>3,20</b>	<b>3,12</b>	<b>3,07</b>	<b>3,00</b>	<b>2,97</b>	<b>2,92</b>	<b>2,89</b>	<b>2,87</b>
2,37	2,33	2,28	2,24	2,20	2,16	2,13	2,09	2,07	2,04	2,02	2,01
<b>3,45</b>	<b>3,37</b>	<b>3,25</b>	<b>3,18</b>	<b>3,10</b>	<b>3,01</b>	<b>2,96</b>	<b>2,89</b>	<b>2,86</b>	<b>2,80</b>	<b>2,77</b>	<b>2,75</b>
2,33	2,29	2,23	2,19	2,15	2,11	2,08	2,04	2,02	1,99	1,97	1,96
<b>3,35</b>	<b>3,27</b>	<b>3,16</b>	<b>3,08</b>	<b>3,00</b>	<b>2,92</b>	<b>2,86</b>	<b>2,79</b>	<b>2,76</b>	<b>2,70</b>	<b>2,67</b>	<b>2,65</b>
2,29	2,25	2,19	2,15	2,11	2,07	2,04	2,00	1,98	1,95	1,93	1,92
<b>3,27</b>	<b>3,19</b>	<b>3,07</b>	<b>3,00</b>	<b>2,91</b>	<b>2,83</b>	<b>2,78</b>	<b>2,71</b>	<b>2,68</b>	<b>2,62</b>	<b>2,59</b>	<b>2,57</b>
2,26	2,21	2,15	2,11	2,07	2,02	2,00	1,96	1,94	1,91	1,90	1,88
<b>3,19</b>	<b>3,12</b>	<b>3,00</b>	<b>2,92</b>	<b>2,84</b>	<b>2,76</b>	<b>2,70</b>	<b>2,63</b>	<b>2,60</b>	<b>2,54</b>	<b>2,51</b>	<b>2,49</b>
2,23	2,18	2,12	2,08	2,04	1,99	1,96	1,92	1,90	1,87	1,85	1,84
<b>3,13</b>	<b>3,05</b>	<b>2,94</b>	<b>2,86</b>	<b>2,77</b>	<b>2,69</b>	<b>2,63</b>	<b>2,56</b>	<b>2,53</b>	<b>2,47</b>	<b>2,44</b>	<b>2,42</b>
2,20	2,15	2,09	2,05	2,00	1,96	1,93	1,89	1,87	1,84	1,82	1,81
<b>3,07</b>	<b>2,99</b>	<b>2,88</b>	<b>2,80</b>	<b>2,72</b>	<b>2,63</b>	<b>2,58</b>	<b>2,51</b>	<b>2,47</b>	<b>2,42</b>	<b>2,38</b>	<b>2,36</b>

$k_2$ — степени свободы для меньшей дисперсии	$k_1$ — степени свободы											
	1	2	3	4	5	6	7	8	9	10	11	12
22	4,30	3,44	3,05	2,82	2,66	2,55	2,47	2,40	2,35	2,30	2,26	2,23
	<b>7,94</b>	<b>5,72</b>	<b>4,82</b>	<b>4,31</b>	<b>3,99</b>	<b>3,76</b>	<b>3,59</b>	<b>3,45</b>	<b>3,35</b>	<b>3,26</b>	<b>3,18</b>	<b>3,12</b>
23	4,28	3,42	3,03	2,80	2,64	2,53	2,45	2,38	2,32	2,28	2,24	2,20
	<b>7,88</b>	<b>5,66</b>	<b>4,76</b>	<b>4,26</b>	<b>3,94</b>	<b>3,71</b>	<b>3,54</b>	<b>3,41</b>	<b>3,30</b>	<b>3,21</b>	<b>3,14</b>	<b>3,07</b>
24	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30	2,26	2,22	2,18
	<b>7,82</b>	<b>5,61</b>	<b>4,72</b>	<b>4,22</b>	<b>3,90</b>	<b>3,67</b>	<b>3,50</b>	<b>3,36</b>	<b>3,25</b>	<b>3,17</b>	<b>3,09</b>	<b>3,03</b>
25	4,24	3,38	2,99	2,76	2,60	2,49	2,41	2,34	2,28	2,24	2,20	2,16
	<b>7,77</b>	<b>5,57</b>	<b>4,68</b>	<b>4,18</b>	<b>3,86</b>	<b>3,63</b>	<b>3,46</b>	<b>3,32</b>	<b>3,21</b>	<b>3,13</b>	<b>3,05</b>	<b>2,99</b>
26	4,22	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,18	2,15
	<b>7,72</b>	<b>5,53</b>	<b>4,64</b>	<b>4,14</b>	<b>3,82</b>	<b>3,59</b>	<b>3,42</b>	<b>3,29</b>	<b>3,17</b>	<b>3,09</b>	<b>3,02</b>	<b>2,96</b>
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,30	2,25	2,20	2,16	2,13
	<b>7,68</b>	<b>5,49</b>	<b>4,60</b>	<b>4,11</b>	<b>3,79</b>	<b>3,56</b>	<b>3,39</b>	<b>3,26</b>	<b>3,14</b>	<b>3,06</b>	<b>2,98</b>	<b>2,93</b>
28	4,20	3,34	2,95	2,71	2,56	2,44	2,36	2,29	2,24	2,19	2,15	2,12
	<b>7,64</b>	<b>5,45</b>	<b>4,57</b>	<b>4,07</b>	<b>3,76</b>	<b>3,53</b>	<b>3,36</b>	<b>3,23</b>	<b>3,11</b>	<b>3,03</b>	<b>2,95</b>	<b>2,90</b>
29	4,18	3,33	2,93	2,70	2,54	2,43	2,35	2,28	2,22	2,18	2,14	2,10
	<b>7,60</b>	<b>5,42</b>	<b>4,54</b>	<b>4,04</b>	<b>3,73</b>	<b>3,50</b>	<b>3,33</b>	<b>3,20</b>	<b>3,08</b>	<b>3,00</b>	<b>2,92</b>	<b>2,87</b>
30	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21	2,16	2,12	2,09
	<b>7,56</b>	<b>5,39</b>	<b>4,51</b>	<b>4,02</b>	<b>3,70</b>	<b>3,47</b>	<b>3,30</b>	<b>3,17</b>	<b>3,06</b>	<b>2,98</b>	<b>2,90</b>	<b>2,84</b>
32	4,15	3,30	2,90	2,67	2,51	2,40	2,32	2,25	2,19	2,14	2,10	2,07
	<b>7,50</b>	<b>5,34</b>	<b>4,46</b>	<b>3,97</b>	<b>3,66</b>	<b>3,42</b>	<b>3,25</b>	<b>3,12</b>	<b>3,01</b>	<b>2,94</b>	<b>2,86</b>	<b>2,80</b>
34	4,13	3,28	2,88	2,65	2,49	2,38	2,30	2,23	2,17	2,12	2,08	2,05
	<b>7,44</b>	<b>5,29</b>	<b>4,42</b>	<b>3,93</b>	<b>3,61</b>	<b>3,38</b>	<b>3,21</b>	<b>3,08</b>	<b>2,97</b>	<b>2,89</b>	<b>2,82</b>	<b>2,76</b>
36	4,11	3,26	2,86	2,63	2,48	2,36	2,28	2,21	2,15	2,10	2,06	2,03
	<b>7,39</b>	<b>5,25</b>	<b>4,38</b>	<b>3,89</b>	<b>3,58</b>	<b>3,35</b>	<b>3,18</b>	<b>3,04</b>	<b>2,94</b>	<b>2,86</b>	<b>2,78</b>	<b>2,72</b>

для большей дисперсии											
14	16	20	24	30	40	50	75	100	200	500	$\infty$
2,18	2,13	2,07	2,03	1,98	1,93	1,91	1,87	1,84	1,81	1,80	1,78
<b>3,02</b>	<b>2,94</b>	<b>2,83</b>	<b>2,75</b>	<b>2,67</b>	<b>2,58</b>	<b>2,53</b>	<b>2,46</b>	<b>2,42</b>	<b>2,37</b>	<b>2,33</b>	<b>2,31</b>
2,14	2,10	2,04	2,00	1,96	1,91	1,88	1,84	1,82	1,79	1,77	1,76
<b>2,97</b>	<b>2,89</b>	<b>2,78</b>	<b>2,70</b>	<b>2,62</b>	<b>2,53</b>	<b>2,48</b>	<b>2,41</b>	<b>2,37</b>	<b>2,32</b>	<b>2,28</b>	<b>2,26</b>
2,13	2,09	2,02	1,98	1,94	1,89	1,86	1,82	1,80	1,76	1,74	1,73
<b>2,93</b>	<b>2,85</b>	<b>2,74</b>	<b>2,66</b>	<b>2,58</b>	<b>2,49</b>	<b>2,44</b>	<b>2,36</b>	<b>2,33</b>	<b>2,27</b>	<b>2,23</b>	<b>2,21</b>
2,11	2,06	2,00	1,96	1,92	1,87	1,84	1,80	1,77	1,74	1,72	1,71
<b>2,89</b>	<b>2,81</b>	<b>2,70</b>	<b>2,62</b>	<b>2,54</b>	<b>2,45</b>	<b>2,40</b>	<b>2,32</b>	<b>2,29</b>	<b>2,23</b>	<b>2,19</b>	<b>2,17</b>
2,10	2,05	1,99	1,95	1,90	1,85	1,82	1,78	1,76	1,72	1,70	1,69
<b>2,86</b>	<b>2,77</b>	<b>2,66</b>	<b>2,58</b>	<b>2,50</b>	<b>2,41</b>	<b>2,36</b>	<b>2,28</b>	<b>2,25</b>	<b>2,19</b>	<b>2,15</b>	<b>2,13</b>
2,08	2,03	1,97	1,93	1,88	1,84	1,80	1,76	1,74	1,71	1,68	1,67
<b>2,83</b>	<b>2,74</b>	<b>2,63</b>	<b>2,55</b>	<b>2,47</b>	<b>2,38</b>	<b>2,33</b>	<b>2,25</b>	<b>2,21</b>	<b>2,16</b>	<b>2,12</b>	<b>2,10</b>
2,06	2,02	1,96	1,91	1,87	1,81	1,78	1,75	1,72	1,69	1,67	1,65
<b>2,80</b>	<b>2,71</b>	<b>2,60</b>	<b>2,52</b>	<b>2,44</b>	<b>2,35</b>	<b>2,30</b>	<b>2,22</b>	<b>2,18</b>	<b>2,13</b>	<b>2,09</b>	<b>2,06</b>
2,05	2,00	1,94	1,90	1,85	1,80	1,77	1,73	1,71	1,68	1,65	1,64
<b>2,77</b>	<b>2,68</b>	<b>2,57</b>	<b>2,49</b>	<b>2,41</b>	<b>2,32</b>	<b>2,27</b>	<b>2,19</b>	<b>2,15</b>	<b>2,10</b>	<b>2,06</b>	<b>2,03</b>
2,04	1,99	1,93	1,89	1,84	1,79	1,76	1,72	1,69	1,66	1,64	1,62
<b>2,74</b>	<b>2,66</b>	<b>2,55</b>	<b>2,47</b>	<b>2,38</b>	<b>2,29</b>	<b>2,24</b>	<b>2,16</b>	<b>2,13</b>	<b>2,07</b>	<b>2,03</b>	<b>2,01</b>
2,02	1,97	1,91	1,86	1,82	1,76	1,74	1,69	1,67	1,64	1,61	1,59
<b>2,70</b>	<b>2,62</b>	<b>2,51</b>	<b>2,42</b>	<b>2,34</b>	<b>2,25</b>	<b>2,20</b>	<b>2,12</b>	<b>2,08</b>	<b>2,02</b>	<b>1,98</b>	<b>1,98</b>
2,00	1,95	1,89	1,84	1,80	1,74	1,71	1,67	1,64	1,61	1,59	1,57
<b>2,66</b>	<b>2,58</b>	<b>2,47</b>	<b>2,38</b>	<b>2,30</b>	<b>2,21</b>	<b>2,15</b>	<b>2,08</b>	<b>2,04</b>	<b>1,98</b>	<b>1,94</b>	<b>1,91</b>
1,98	1,93	1,87	1,82	1,78	1,72	1,69	1,65	1,62	1,59	1,56	1,55
<b>2,62</b>	<b>2,54</b>	<b>2,43</b>	<b>2,35</b>	<b>2,26</b>	<b>2,17</b>	<b>2,12</b>	<b>2,04</b>	<b>2,00</b>	<b>1,94</b>	<b>1,90</b>	<b>1,87</b>

$k_2$ — степени свободы для меньшей дисперсии	$k_1$ — степени свободы											
	1	2	3	4	5	6	7	8	9	10	11	12
38	4,10	3,25	2,85	2,65	2,46	2,35	2,26	2,19	2,14	2,09	2,05	2,02
	<b>7,35</b>	<b>5,21</b>	<b>4,34</b>	<b>3,86</b>	<b>3,54</b>	<b>3,32</b>	<b>3,15</b>	<b>3,02</b>	<b>2,91</b>	<b>2,82</b>	<b>2,75</b>	<b>2,69</b>
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,07	2,04	2,00
	<b>7,31</b>	<b>5,18</b>	<b>4,31</b>	<b>3,83</b>	<b>3,51</b>	<b>3,29</b>	<b>3,12</b>	<b>2,99</b>	<b>2,88</b>	<b>2,80</b>	<b>2,73</b>	<b>2,66</b>
42	4,07	3,22	2,83	2,59	2,44	2,32	2,24	2,17	2,11	2,06	2,02	1,99
	<b>7,27</b>	<b>5,15</b>	<b>4,29</b>	<b>3,80</b>	<b>3,49</b>	<b>3,26</b>	<b>3,10</b>	<b>2,96</b>	<b>2,86</b>	<b>2,77</b>	<b>2,70</b>	<b>2,64</b>
44	4,06	3,21	2,82	2,58	2,43	2,31	2,23	2,16	2,10	2,05	2,01	1,98
	<b>7,24</b>	<b>5,12</b>	<b>4,26</b>	<b>3,78</b>	<b>3,46</b>	<b>3,24</b>	<b>3,07</b>	<b>2,94</b>	<b>2,84</b>	<b>2,75</b>	<b>2,68</b>	<b>2,62</b>
46	4,05	3,20	2,81	2,57	2,42	2,30	2,22	2,14	2,09	2,04	2,00	1,97
	<b>7,21</b>	<b>5,10</b>	<b>4,24</b>	<b>3,76</b>	<b>3,44</b>	<b>3,22</b>	<b>3,05</b>	<b>2,92</b>	<b>2,82</b>	<b>2,73</b>	<b>2,66</b>	<b>2,60</b>
48	4,04	3,19	2,80	2,56	2,41	2,30	2,21	2,14	2,08	2,03	1,99	1,96
	<b>7,19</b>	<b>5,08</b>	<b>4,22</b>	<b>3,74</b>	<b>3,42</b>	<b>3,20</b>	<b>3,04</b>	<b>2,90</b>	<b>2,80</b>	<b>2,71</b>	<b>2,64</b>	<b>2,58</b>
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,02	1,98	1,95
	<b>7,17</b>	<b>5,06</b>	<b>4,20</b>	<b>3,72</b>	<b>3,41</b>	<b>3,18</b>	<b>3,02</b>	<b>2,88</b>	<b>2,78</b>	<b>2,70</b>	<b>2,62</b>	<b>2,56</b>
55	4,02	3,17	2,78	2,54	2,38	2,27	2,18	2,11	2,05	2,00	1,97	1,93
	<b>7,12</b>	<b>5,01</b>	<b>4,16</b>	<b>3,68</b>	<b>3,37</b>	<b>3,15</b>	<b>2,98</b>	<b>2,85</b>	<b>2,75</b>	<b>2,66</b>	<b>2,59</b>	<b>2,53</b>
60	4,00	3,15	2,76	2,52	2,37	2,25	2,17	2,10	2,04	1,99	1,95	1,92
	<b>7,08</b>	<b>4,98</b>	<b>4,13</b>	<b>3,65</b>	<b>3,34</b>	<b>3,12</b>	<b>2,95</b>	<b>2,82</b>	<b>2,72</b>	<b>2,63</b>	<b>2,56</b>	<b>2,50</b>
65	3,99	3,14	2,75	2,51	2,36	2,24	2,15	2,08	2,02	1,98	1,94	1,90
	<b>7,04</b>	<b>4,95</b>	<b>4,10</b>	<b>3,62</b>	<b>3,31</b>	<b>3,09</b>	<b>2,93</b>	<b>2,79</b>	<b>2,70</b>	<b>2,61</b>	<b>2,54</b>	<b>2,47</b>
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,01	1,97	1,93	1,89
	<b>7,01</b>	<b>4,92</b>	<b>4,08</b>	<b>3,60</b>	<b>3,29</b>	<b>3,07</b>	<b>2,91</b>	<b>2,77</b>	<b>2,67</b>	<b>2,59</b>	<b>2,51</b>	<b>2,45</b>
80	3,96	3,11	2,72	2,48	2,33	2,21	2,12	2,05	1,99	1,95	1,91	1,89
	<b>6,96</b>	<b>4,88</b>	<b>4,04</b>	<b>3,56</b>	<b>3,25</b>	<b>3,04</b>	<b>2,87</b>	<b>2,74</b>	<b>2,64</b>	<b>2,55</b>	<b>2,48</b>	<b>2,41</b>

для большей дисперсии											
14	16	20	24	30	40	50	75	100	200	500	$\infty$
1,96	1,92	1,85	1,80	1,76	1,71	1,67	1,63	1,60	1,57	1,54	1,53
<b>2,59</b>	<b>2,51</b>	<b>2,40</b>	<b>2,32</b>	<b>2,22</b>	<b>2,14</b>	<b>2,08</b>	<b>2,00</b>	<b>1,97</b>	<b>1,90</b>	<b>1,86</b>	<b>1,84</b>
1,95	1,90	1,84	1,79	1,74	1,69	1,66	1,61	1,59	1,55	1,53	1,51
<b>2,56</b>	<b>2,49</b>	<b>2,37</b>	<b>2,29</b>	<b>2,20</b>	<b>2,11</b>	<b>2,05</b>	<b>1,97</b>	<b>1,94</b>	<b>1,88</b>	<b>1,84</b>	<b>1,81</b>
1,94	1,89	1,82	1,78	1,73	1,68	1,64	1,60	1,57	1,54	1,51	1,49
<b>2,54</b>	<b>2,46</b>	<b>2,35</b>	<b>2,26</b>	<b>2,17</b>	<b>2,08</b>	<b>2,02</b>	<b>1,94</b>	<b>1,91</b>	<b>1,85</b>	<b>1,80</b>	<b>1,78</b>
1,92	1,88	1,81	1,76	1,72	1,66	1,63	1,58	1,56	1,52	1,50	1,48
<b>2,52</b>	<b>2,44</b>	<b>2,32</b>	<b>2,24</b>	<b>2,15</b>	<b>2,06</b>	<b>2,00</b>	<b>1,92</b>	<b>1,88</b>	<b>1,82</b>	<b>1,78</b>	<b>1,75</b>
1,91	1,87	1,80	1,75	1,71	1,65	1,62	1,57	1,54	1,51	1,48	1,46
<b>2,50</b>	<b>2,42</b>	<b>2,30</b>	<b>2,22</b>	<b>2,13</b>	<b>2,04</b>	<b>1,98</b>	<b>1,90</b>	<b>1,86</b>	<b>1,80</b>	<b>1,76</b>	<b>1,72</b>
1,90	1,86	1,79	1,74	1,70	1,64	1,61	1,56	1,53	1,50	1,47	1,45
<b>2,48</b>	<b>2,40</b>	<b>2,28</b>	<b>2,20</b>	<b>2,11</b>	<b>2,02</b>	<b>1,96</b>	<b>1,88</b>	<b>1,84</b>	<b>1,78</b>	<b>1,73</b>	<b>1,70</b>
1,90	1,85	1,78	1,74	1,69	1,63	1,60	1,55	1,52	1,48	1,46	1,44
<b>2,46</b>	<b>2,39</b>	<b>2,26</b>	<b>2,18</b>	<b>2,10</b>	<b>2,00</b>	<b>1,94</b>	<b>1,86</b>	<b>1,82</b>	<b>1,76</b>	<b>1,71</b>	<b>1,68</b>
1,88	1,83	1,76	1,72	1,67	1,61	1,58	1,52	1,50	1,46	1,43	1,41
<b>2,43</b>	<b>2,35</b>	<b>2,23</b>	<b>2,15</b>	<b>2,06</b>	<b>1,96</b>	<b>1,90</b>	<b>1,82</b>	<b>1,78</b>	<b>1,71</b>	<b>1,66</b>	<b>1,64</b>
1,86	1,81	1,75	1,70	1,65	1,59	1,56	1,50	1,48	1,44	1,41	1,39
<b>2,40</b>	<b>2,32</b>	<b>2,20</b>	<b>2,12</b>	<b>2,03</b>	<b>1,93</b>	<b>1,87</b>	<b>1,79</b>	<b>1,74</b>	<b>1,68</b>	<b>1,63</b>	<b>1,60</b>
1,85	1,80	1,73	1,68	1,63	1,57	1,54	1,49	1,46	1,42	1,39	1,37
<b>2,37</b>	<b>2,30</b>	<b>2,18</b>	<b>2,09</b>	<b>2,00</b>	<b>1,90</b>	<b>1,84</b>	<b>1,76</b>	<b>1,71</b>	<b>1,64</b>	<b>1,60</b>	<b>1,56</b>
1,84	1,79	1,72	1,67	1,62	1,56	1,53	1,47	1,45	1,40	1,37	1,35
<b>2,35</b>	<b>2,28</b>	<b>2,15</b>	<b>2,07</b>	<b>1,98</b>	<b>1,88</b>	<b>1,82</b>	<b>1,74</b>	<b>1,69</b>	<b>1,62</b>	<b>1,56</b>	<b>1,53</b>
1,82	1,77	1,70	1,65	1,60	1,54	1,51	1,45	1,42	1,38	1,35	1,32
<b>2,32</b>	<b>2,24</b>	<b>2,11</b>	<b>2,03</b>	<b>1,94</b>	<b>1,84</b>	<b>1,78</b>	<b>1,70</b>	<b>1,65</b>	<b>1,57</b>	<b>1,52</b>	<b>1,49</b>

$k_2$ — степени свободы для меньшей дисперсии	$k_1$ — степени свободы											
	1	2	3	4	5	6	7	8	9	10	11	12
100	3,94	3,09	2,70	2,46	2,30	2,19	2,10	2,03	1,97	1,92	1,88	1,85
	<b>6,90</b>	<b>4,82</b>	<b>3,98</b>	<b>3,51</b>	<b>3,20</b>	<b>2,99</b>	<b>2,82</b>	<b>2,69</b>	<b>2,59</b>	<b>2,51</b>	<b>2,43</b>	<b>2,36</b>
125	3,92	3,07	2,68	2,44	2,29	2,17	2,08	2,01	1,95	1,90	1,86	1,83
	<b>6,84</b>	<b>4,78</b>	<b>3,94</b>	<b>3,47</b>	<b>3,17</b>	<b>2,95</b>	<b>2,79</b>	<b>2,65</b>	<b>2,56</b>	<b>2,47</b>	<b>2,40</b>	<b>2,33</b>
150	3,91	3,06	2,67	2,43	2,27	2,16	2,07	2,00	1,94	1,89	1,85	1,82
	<b>6,81</b>	<b>4,75</b>	<b>3,91</b>	<b>3,44</b>	<b>3,14</b>	<b>2,92</b>	<b>2,76</b>	<b>2,62</b>	<b>2,53</b>	<b>2,44</b>	<b>2,37</b>	<b>2,30</b>
200	3,89	3,04	2,65	2,41	2,26	2,14	2,05	1,98	1,92	1,87	1,83	1,80
	<b>6,76</b>	<b>4,71</b>	<b>3,88</b>	<b>3,41</b>	<b>3,11</b>	<b>2,90</b>	<b>2,73</b>	<b>2,60</b>	<b>2,50</b>	<b>2,41</b>	<b>2,34</b>	<b>2,28</b>
400	3,86	3,02	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	1,81	1,78
	<b>6,70</b>	<b>4,66</b>	<b>3,83</b>	<b>3,36</b>	<b>3,06</b>	<b>2,85</b>	<b>2,69</b>	<b>2,55</b>	<b>2,46</b>	<b>2,37</b>	<b>2,29</b>	<b>2,23</b>
1000	3,85	3,00	2,61	2,38	2,22	2,10	2,02	1,95	1,89	1,84	1,80	1,76
	<b>6,66</b>	<b>4,62</b>	<b>3,80</b>	<b>3,34</b>	<b>3,04</b>	<b>2,82</b>	<b>2,66</b>	<b>2,53</b>	<b>2,43</b>	<b>2,34</b>	<b>2,26</b>	<b>2,20</b>
$\infty$	3,84	2,99	2,60	2,37	2,21	2,09	2,01	1,94	1,88	1,83	1,79	1,75
	<b>6,64</b>	<b>4,60</b>	<b>3,78</b>	<b>3,32</b>	<b>3,02</b>	<b>2,80</b>	<b>2,64</b>	<b>2,51</b>	<b>2,41</b>	<b>2,32</b>	<b>2,24</b>	<b>2,18</b>

\* Пятипроцентные пределы  $F_5(k_1, k_2)$  набраны обычным шрифтом, а

для большей дисперсии											
14	16	20	24	30	40	50	75	100	200	500	$\infty$
1,79	1,75	1,68	1,63	1,57	1,51	1,48	1,42	1,39	1,34	1,30	1,28
<b>2,26</b>	<b>2,19</b>	<b>2,06</b>	<b>1,98</b>	<b>1,89</b>	<b>1,79</b>	<b>1,73</b>	<b>1,64</b>	<b>1,59</b>	<b>1,51</b>	<b>1,46</b>	<b>1,43</b>
1,77	1,72	1,65	1,60	1,55	1,49	1,45	1,39	1,36	1,31	1,27	1,25
<b>2,23</b>	<b>2,15</b>	<b>2,03</b>	<b>1,94</b>	<b>1,85</b>	<b>1,75</b>	<b>1,68</b>	<b>1,59</b>	<b>1,54</b>	<b>1,46</b>	<b>1,40</b>	<b>1,37</b>
1,76	1,71	1,64	1,59	1,54	1,47	1,44	1,37	1,34	1,29	1,25	1,22
<b>2,20</b>	<b>2,12</b>	<b>2,00</b>	<b>1,91</b>	<b>1,83</b>	<b>1,72</b>	<b>1,66</b>	<b>1,56</b>	<b>1,51</b>	<b>1,43</b>	<b>1,37</b>	<b>1,33</b>
1,74	1,69	1,62	1,57	1,52	1,45	1,42	1,35	1,32	1,26	1,22	1,19
<b>2,17</b>	<b>2,09</b>	<b>1,97</b>	<b>1,88</b>	<b>1,79</b>	<b>1,69</b>	<b>1,62</b>	<b>1,53</b>	<b>1,48</b>	<b>1,39</b>	<b>1,33</b>	<b>1,28</b>
1,72	1,67	1,60	1,54	1,49	1,42	1,38	1,32	1,28	1,22	1,16	1,13
<b>2,12</b>	<b>2,04</b>	<b>1,92</b>	<b>1,84</b>	<b>1,74</b>	<b>1,64</b>	<b>1,57</b>	<b>1,47</b>	<b>1,42</b>	<b>1,32</b>	<b>1,24</b>	<b>1,19</b>
1,70	1,65	1,58	1,53	1,47	1,41	1,36	1,30	1,26	1,19	1,13	1,08
<b>2,09</b>	<b>2,01</b>	<b>1,89</b>	<b>1,81</b>	<b>1,71</b>	<b>1,61</b>	<b>1,54</b>	<b>1,44</b>	<b>1,38</b>	<b>1,28</b>	<b>1,19</b>	<b>1,11</b>
1,69	1,64	1,57	1,52	1,46	1,40	1,35	1,28	1,24	1,17	1,11	1,00
<b>2,07</b>	<b>1,99</b>	<b>1,87</b>	<b>1,79</b>	<b>1,69</b>	<b>1,59</b>	<b>1,52</b>	<b>1,41</b>	<b>1,36</b>	<b>1,25</b>	<b>1,15</b>	<b>1,09</b>

однопроцентные  $F_1(k_1, k_2)$  — полужирным.

Пяти- и однопроцентные пределы для отношений  $G_i$  к сумме эмпирических дисперсий, полученных

$r-1$	1	2	3	4	5	6	
Сумма $k$ независимых выборок объема $r$	2	0,9985	0,9750	0,9392	0,9057	0,8772	0,8534
		<b>0,999</b>	<b>0,9950</b>	<b>0,9794</b>	<b>0,9586</b>	<b>0,9373</b>	<b>0,9172</b>
	3	0,9669	0,8709	0,7977	0,7457	0,7071	0,6771
		<b>0,9933</b>	<b>0,9423</b>	<b>0,8831</b>	<b>0,8335</b>	<b>0,7933</b>	<b>0,7606</b>
	4	0,9065	0,7679	0,6841	0,6287	0,5895	0,5598
		<b>0,9676</b>	<b>0,8643</b>	<b>0,7814</b>	<b>0,7212</b>	<b>0,6761</b>	<b>0,6410</b>
	5	0,8412	0,6838	0,5981	0,5441	0,5065	0,4783
		<b>0,9279</b>	<b>0,7885</b>	<b>0,6957</b>	<b>0,6329</b>	<b>0,5875</b>	<b>0,5531</b>
	6	0,7808	0,6161	0,5321	0,4803	0,4447	0,4184
		<b>0,8828</b>	<b>0,7218</b>	<b>0,6258</b>	<b>0,5635</b>	<b>0,5195</b>	<b>0,4866</b>
	7	0,7271	0,5612	0,4800	0,4307	0,3974	0,3726
		<b>0,8376</b>	<b>0,6644</b>	<b>0,5685</b>	<b>0,5080</b>	<b>0,4659</b>	<b>0,4347</b>
	8	0,6798	0,5157	0,4377	0,3910	0,3595	0,3362
		<b>0,7945</b>	<b>0,6152</b>	<b>0,5209</b>	<b>0,4627</b>	<b>0,4226</b>	<b>0,3932</b>
	9	0,6385	0,4775	0,4027	0,3584	0,3286	0,3067
		<b>0,7544</b>	<b>0,5727</b>	<b>0,4810</b>	<b>0,4251</b>	<b>0,3870</b>	<b>0,3592</b>
	10	0,6020	0,4450	0,3733	0,3311	0,3029	0,2823
		<b>0,7175</b>	<b>0,5358</b>	<b>0,4469</b>	<b>0,3934</b>	<b>0,3572</b>	<b>0,3308</b>
	12	0,5410	0,3924	0,3264	0,2880	0,2624	0,2439
		<b>0,6528</b>	<b>0,4751</b>	<b>0,3919</b>	<b>0,3328</b>	<b>0,3099</b>	<b>0,2861</b>
15	0,4709	0,3346	0,2758	0,2419	0,2195	0,2034	
	<b>0,5747</b>	<b>0,4069</b>	<b>0,3317</b>	<b>0,2882</b>	<b>0,2593</b>	<b>0,2386</b>	

(распределение  $k$  Кохрана) наибольшей эмпирической дисперсии из  $k$  независимых выборок объема  $r^*$

7	8	9	10	16	36	144	$\infty$
0,8332	0,8159	0,8010	0,7880	0,7341	0,6602	0,5813	0,5000
<b>0,8988</b>	<b>0,8823</b>	<b>0,8674</b>	<b>0,8539</b>	<b>0,7949</b>	<b>0,7067</b>	<b>0,6062</b>	<b>0,5000</b>
0,6530	0,6333	0,6167	0,6025	0,5466	0,4748	0,4031	0,3333
<b>0,7335</b>	<b>0,7107</b>	<b>0,6912</b>	<b>0,6743</b>	<b>0,6059</b>	<b>0,5153</b>	<b>0,4230</b>	<b>0,3333</b>
0,5365	0,5175	0,5017	0,4884	0,4366	0,3720	0,3093	0,2500
<b>0,6129</b>	<b>0,5897</b>	<b>0,5702</b>	<b>0,5536</b>	<b>0,4884</b>	<b>0,4057</b>	<b>0,3251</b>	<b>0,2500</b>
0,4564	0,4387	0,4241	0,4118	0,3645	0,3066	0,2513	0,2000
<b>0,5259</b>	<b>0,5037</b>	<b>0,4854</b>	<b>0,4697</b>	<b>0,4094</b>	<b>0,3351</b>	<b>0,2644</b>	<b>0,2000</b>
0,3980	0,3817	0,3682	0,3568	0,3135	0,2612	0,2119	0,1667
<b>0,4608</b>	<b>0,4401</b>	<b>0,4229</b>	<b>0,4084</b>	<b>0,3529</b>	<b>0,2858</b>	<b>0,2229</b>	<b>0,1667</b>
0,3535	0,3384	0,3259	0,3154	0,2756	0,2278	0,1833	0,1429
<b>0,4105</b>	<b>0,3911</b>	<b>0,3751</b>	<b>0,3616</b>	<b>0,3105</b>	<b>0,2494</b>	<b>0,1929</b>	<b>0,1429</b>
0,3185	0,3043	0,2926	0,2829	0,2462	0,2022	0,1616	0,1250
<b>0,3704</b>	<b>0,3522</b>	<b>0,3373</b>	<b>0,3248</b>	<b>0,2779</b>	<b>0,2214</b>	<b>0,1700</b>	<b>0,1250</b>
0,2901	0,2768	0,2659	0,2568	0,2226	0,1820	0,1446	0,1111
<b>0,3378</b>	<b>0,3207</b>	<b>0,3067</b>	<b>0,2950</b>	<b>0,2514</b>	<b>0,1992</b>	<b>0,1521</b>	<b>0,1111</b>
0,2666	0,2541	0,2439	0,2353	0,2032	0,1655	0,1308	0,1000
<b>0,3106</b>	<b>0,2945</b>	<b>0,2813</b>	<b>0,2704</b>	<b>0,2297</b>	<b>0,1811</b>	<b>0,1376</b>	<b>0,1000</b>
0,2299	0,2187	0,2098	0,2020	0,1737	0,1403	0,1100	0,8333
<b>0,2680</b>	<b>0,2535</b>	<b>0,2419</b>	<b>0,2320</b>	<b>0,1961</b>	<b>0,1535</b>	<b>0,1157</b>	<b>0,0833</b>
0,1911	0,1815	0,1736	0,1671	0,1429	0,1144	0,0889	0,0667
<b>0,2228</b>	<b>0,2104</b>	<b>0,2002</b>	<b>0,1918</b>	<b>0,1612</b>	<b>0,1251</b>	<b>0,0934</b>	<b>0,0667</b>

$r-1$	1	2	3	4	5	6
20	0,3894	0,2705	0,2205	0,1921	0,1735	0,1602
	<b>0,4799</b>	<b>0,3297</b>	<b>0,2654</b>	<b>0,2288</b>	<b>0,2048</b>	<b>0,1877</b>
24	0,3434	0,2354	0,1907	0,1656	0,1493	0,1374
	<b>0,4247</b>	<b>0,2871</b>	<b>0,2295</b>	<b>0,1970</b>	<b>0,1759</b>	<b>0,1608</b>
30	0,2929	0,1980	0,1593	0,1377	0,1237	0,1137
	<b>0,3632</b>	<b>0,2412</b>	<b>0,1913</b>	<b>0,1635</b>	<b>0,1454</b>	<b>0,1327</b>
40	0,2370	0,1576	0,1259	0,1082	0,0968	0,0887
	<b>0,2940</b>	<b>0,1915</b>	<b>0,1508</b>	<b>0,1281</b>	<b>0,1135</b>	<b>0,1033</b>
60	0,1737	0,1131	0,0895	0,0765	0,0682	0,0623
	<b>0,2151</b>	<b>0,1371</b>	<b>0,1069</b>	<b>0,0902</b>	<b>0,0796</b>	<b>0,0722</b>
120	0,0998	0,0632	0,0495	0,0419	0,0371	0,0337
	<b>0,1225</b>	<b>0,0759</b>	<b>0,0585</b>	<b>0,0489</b>	<b>0,0429</b>	<b>0,0387</b>
$\infty$	0	0	0	0	0	0
	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>

\* Пятипроцентные пределы набраны обычным шрифтом, а однопроцентные —

Окончание таблицы

7	8	9	10	16	36	144	$\infty$
0,1501	0,1422	0,1357	0,1303	0,1108	0,0879	0,0675	0,0500
<b>0,1748</b>	<b>0,1646</b>	<b>0,1567</b>	<b>0,1501</b>	<b>0,1248</b>	<b>0,0960</b>	<b>0,0709</b>	<b>0,0500</b>
0,1286	0,1216	0,1160	0,1113	0,0942	0,0743	0,0567	0,0417
<b>0,1495</b>	<b>0,1406</b>	<b>0,1338</b>	<b>0,1283</b>	<b>0,1060</b>	<b>0,0810</b>	<b>0,0595</b>	<b>0,0417</b>
0,1061	0,1002	0,0958	0,0921	0,0771	0,0604	0,0457	0,0333
<b>0,1232</b>	<b>0,1157</b>	<b>0,1100</b>	<b>0,1054</b>	<b>0,0867</b>	<b>0,0658</b>	<b>0,0480</b>	<b>0,0333</b>
0,0827	0,0780	0,0745	0,0713	0,0595	0,0462	0,0347	0,0250
<b>0,0957</b>	<b>0,0898</b>	<b>0,0853</b>	<b>0,0816</b>	<b>0,0668</b>	<b>0,0503</b>	<b>0,0363</b>	<b>0,0250</b>
0,0583	0,0552	0,0520	0,0497	0,0411	0,0316	0,0234	0,0167
<b>0,0668</b>	<b>0,0625</b>	<b>0,0594</b>	<b>0,0567</b>	<b>0,0461</b>	<b>0,0344</b>	<b>0,0245</b>	<b>0,0167</b>
0,0312	0,0292	0,0279	0,0266	0,0218	0,0165	0,0120	0,0083
<b>0,0357</b>	<b>0,0334</b>	<b>0,0316</b>	<b>0,0302</b>	<b>0,0242</b>	<b>0,0178</b>	<b>0,0125</b>	<b>0,0083</b>
0	0	0	0	0	0	0	0
<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>

полужирным.



**Характеристики компонент факторов, определяющих величину потерь линейного времени по техническим причинам**

Таблица П6.1

**Характеристики компонент для первого набора факторов**

Факторы	Компоненты						
	1	2	3	4	5	6	7
$X_2$	0,55	0,03	-0,07	0,08	-0,08	0,19	0,04
$X_3$	-0,50	0,60	-0,28	-0,04	-0,23	0,22	0,46
$X_4$	0,05	0,48	0,78	-0,17	0,02	0,29	-0,24
$X_5$	0,01	0,36	-0,08	0,86	0,06	-0,16	-0,32
$X_6$	0,42	0,39	0,07	-0,16	0,14	-0,69	0,38
$X_7$	0,27	0,39	-0,55	-0,39	0,22	0,16	-0,52
$X_8$	0,44	-0,01	-0,04	0,23	0,48	0,55	0,48
$\lambda_j$	2,84	0,02	1,42	1,24	0,15	0,78	0,53

Таблица П6.2

**Характеристики компонент для второго набора факторов**

Факторы	Компоненты			
	1	2	3	4
$X_{12}$	-0,68	-0,71	-0,15	-0,06
$X_{13}$	0,58	0,62	-0,10	-0,51
$X_{14}$	0,34	0,11	0,87	0,33
$X_{15}$	0,29	0,30	-0,46	0,79
$\lambda_j$	1,76	0,31	0,89	1,03

Таблица П6.3

**Характеристики компонент для третьего набора факторов**

Факторы	Компоненты					
	1	2	3	4	5	6
$X_{18}$	0,09	0,82	0,13	-0,09	-0,22	0,49

Факторы	Компоненты					
	1	2	3	4	5	6
$X_{19}$	-0,52	0,06	-0,20	0,80	-0,19	0,10
$X_{20}$	-0,48	-0,06	0,82	-0,13	-0,24	-0,16
$X_{21}$	0,50	0,03	0,51	0,56	0,48	0,02
$X_{23}$	-0,32	0,52	-0,10	-0,07	0,57	-0,54
$X_{24}$	-0,38	-0,22	0,04	-0,16	0,59	0,65
$\lambda_j$	3,24	1,13	0,32	0,12	0,44	0,74

Таблица П6.4

**Характеристики компонент для четвертого набора факторов**

Факторы	Компоненты				
	1	2	3	4	5
$X_1$	0,14	0,68	-0,61	-0,35	0,12
$X_9$	0,14	0,67	0,64	0,31	0,18
$X_{10}$	0,56	-0,04	0,34	-0,59	-0,47
$X_{11}$	0,56	-0,003	-0,33	0,65	-0,39
$X_{27}$	0,58	-0,29	-0,01	-0,05	0,76
$\lambda_j$	2,48	1,20	0,95	0,25	0,12

Таблица П6.5

**Характеристики компонент для пятого набора факторов**

Факторы	Компоненты		
	1	2	3
$X_{26}$	0,50	0,80	0,33
$X_{30}$	-0,57	0,59	-0,57
$X_{34}$	-0,65	0,09	0,75
$\lambda_j$	1,92	0,76	0,32

## Характеристики компонент для шестого набора факторов

Факторы	Компоненты	
	1	2
$X_{31}$	0,71	-0,71
$X_{33}$	0,71	0,71
$\lambda_y$	1,54	0,46

## Темы практических занятий по курсу основы научных исследований

1. Сбор и предварительная обработка статистических данных.
2. Определение вида закона распределения случайных величин.
3. Построение однофакторных моделей с использованием программного обеспечения Microsoft Office Excel.
4. Выбор вида математической модели с использованием программного обеспечения Microsoft Office Excel.
5. Методика корреляционного анализа и определение значимых факторов.
6. Методика регрессионного анализа с использованием программного обеспечения STATISTICA.
7. Методика компонентного анализа с использованием программного обеспечения STATISTICA.

1. *Адлер Ю. П.* Планирование эксперимента при поиске оптимальных условий / Ю. П. Адлер, Е. В. Маркова, Ю. В. Грановский. — М. : Наука, 1976. — 280 с.
2. *Болдин А. П.* Основы научных исследований и УНИРС : учеб. пособие / А. П. Болдин, В. А. Максимов. — М. : Изд-во МАДИ (ГТУ), 2002. — 276 с.
3. *Болдин А. П.* Основы научных исследований и УНИРС : В 2 кн. Кн. 2. Специальные методы и методологические подходы : учеб. пособие / А. П. Болдин, В. А. Максимов. — М. : Изд-во МАДИ (ГТУ), 2004. — 181 с.
4. *Боровиков В.* STATISTICA : искусство анализа данных на компьютере / В. Боровиков. — СПб. : Питер, 2001. — 656 с.
5. *Дубров А. М.* Многомерные статистические методы : учебник / А. М. Дубров, В. С. Мхитарян, Л. И. Трошин. — М. : Финансы и статистика, 1998. — 398 с.
6. *Мирошников Л. В.* Диагностирование технического состояния автомобилей на автотранспортных предприятиях / Л. В. Мирошников, А. П. Болдин, В. И. Пал. — М. : Транспорт, 1977. — 263 с.
7. Таблицы планов экспериментов для факторных полиномиальных моделей. — М. : Metallurgia, 1982. — 150 с.
8. Техническая эксплуатация автомобилей : учебник / [Е. С. Кузнецов, А. П. Болдин, В. М. Власов и др.] ; под ред. Е. С. Кузнецова. — 4-е изд., перераб. и доп. — М. : Наука, 2001. — 535 с.
9. Факторный, дискриминантный и кластерный анализ : пер. с англ. / [Дж.-О. Ким, Ч. У. Мьюллер, У. Р. Клекка и др.] ; под ред. А. М. Хотинского, С. Б. Королева. — М. : Финансы и статистика, 1989. — 215 с.
10. РД 50-690-89. Методические указания. Надежность в технике. Методы оценки показателей надежности по экспериментальным данным. — М. : Изд-во стандартов, 1990. — 132 с.

Предисловие.....	3
Введение.....	4
<b>Глава 1. Основные понятия и определения учебного курса «Основы научных исследований» .....</b>	<b>7</b>
1.1. Понятия о науке .....	7
1.2. Характерные черты современной науки.....	9
1.3. Определение и классификация научных исследований .....	11
1.4. Методы научного исследования при технической эксплуатации автомобилей .....	12
1.5. Выбор темы научного исследования.....	14
1.6. Этапы научного исследования .....	17
1.7. Основные цели и подходы научного исследования, сущность пассивного и активного эксперимента.....	18
<b>Глава 2. Применение закономерностей рассеяния непрерывных случайных величин при проведении исследований эксплуатационной надежности автомобилей и других показателей их работы на автотранспортных предприятиях .....</b>	<b>22</b>
2.1. Случайные величины и возможности обработки экспериментальных данных на их основе компьютерными программами .....	22
2.2. Обработка случайных величин, связанных с рассеянием изучаемого показателя, на примере изучения долговечности автомобильных деталей, узлов и агрегатов .....	24
2.3. Графическая интерпретация случайных величин и построение гистограмм.....	27
2.4. Законы распределения случайных величин .....	35
2.5. Проверка соответствия закона распределения эмпирическим данным на основе критерия Пирсона .....	43
2.6. Понятие доверительного интервала и доверительной вероятности при статистической оценке характеристик рассеяния случайных величин .....	48
2.7. Определение объема выборки и организация наблюдений за автомобилями при изучении показателей их работы в эксплуатации .....	56

<b>Глава 3. Использование критериев Стьюдента, Фишера и дисперсионного анализа при выявлении расхождения сравнимых выборок случайных величин и обосновании возможности их объединения. Разделение смешанных выборок.....</b>	<b>63</b>
3.1. Простейший случай проверки «нулевой» гипотезы о принадлежности двух выборок одной генеральной совокупности.....	63
3.2. Однофакторный и многофакторный дисперсионные анализы как общие методы проверки расхождения между средними при большом количестве статистических выборок.....	69
3.3. Применение кластерного анализа и метода подбора закона распределения в ограниченном диапазоне данных для разделения смешанных выборок.....	76
3.4. Пример использования принципов разделения и объединения выборок для определения нормативов метода диагностирования экологической безопасности карбюраторных автомобилей при их испытаниях на ненагруженных беговых барабанах.....	80
<b>Глава 4. Сглаживание стохастических зависимостей. Корреляционный и регрессионный анализы.....</b>	<b>91</b>
4.1. Сглаживание стохастических экспериментальных зависимостей по методу наименьших квадратов для случая однофакторной линейной регрессии.....	91
4.2. Коэффициент детерминации и его использование для оценки точности и адекватности однофакторной модели линейной регрессии.....	96
4.3. Матричные способы определения коэффициентов уравнений многофакторных регрессий, представляемых полиномами $n$ -й степени.....	101
4.4. Оценка точности и адекватности многофакторной регрессионной модели линейного и нелинейного (степенного) видов.....	107
4.5. Осуществление прогноза по разработанным регрессионным моделям и выявление аномальных исходных данных.....	111
<b>Глава 5. Применение активных многофакторных экспериментов при решении задач технической эксплуатации автомобилей.....</b>	<b>119</b>
5.1. Простейший случай статистического планирования активного однофакторного эксперимента.....	119
5.2. Планирование активного двухфакторного эксперимента.....	132
5.3. Ортогональное планирование активного эксперимента для линейной модели с количеством факторов больше двух и возможность сокращения числа основных опытов за счет использования реплик различной дробности.....	142
5.4. Планирование эксперимента при поиске оптимальных условий.....	154

5.5. Нелинейное планирование активного эксперимента для получения моделей многофакторных зависимостей второго порядка и поиска экстремальных значений функции отклика.....	165
<b>Глава 6. Особенности компонентного анализа и основные предпосылки для его применения при управлении процессами технической эксплуатацией автомобилей.....</b>	<b>181</b>
6.1. Основные принципиальные подходы при оценке влияющих факторов с использованием многошагового регрессионного и компонентного анализов.....	181
6.2. Метод главных компонент.....	185
6.2.1. Общая характеристика метода главных компонент.....	185
6.2.2. Вычисление главных компонент.....	188
6.2.3. Основные числовые характеристики главных компонент.....	189
6.2.4. Выбор главных компонент и переход к обобщенным факторам.....	191
6.3. Примеры использования компонентного анализа при решении задач управления процессами технической эксплуатации автомобилей.....	198
<b>Глава 7. Имитационное моделирование как метод получения количественных оценок перспективных организационных и технологических систем поддержания работоспособности автомобилей.....</b>	<b>237</b>
7.1. Возможности имитационного моделирования в исследовании вариантов применения внешнего и встроенного диагностирования на автомобильном транспорте.....	237
7.2. Основные стратегии поддержания исправного технического состояния для отдельного элемента (детали, узла, агрегата) автомобиля.....	243
7.3. Основные организационно-технологические варианты обслуживания и ремонта автомобилей на АТП общего пользования, подлежащие модельному исследованию.....	254
7.4. Результаты моделирования основных вариантов организации ТО и ремонта на основе использования стационарного и встроенного диагностирования на автотранспортных предприятиях общего пользования.....	268
<b>Глава 8. Приборное и метрологическое обеспечение научных исследований на автотранспортных предприятиях.....</b>	<b>285</b>
8.1. Основные понятия и определения в области метрологии.....	285
8.2. Метрологическая служба.....	288
8.3. Метрологическое обеспечение научных исследований.....	290
8.4. Нормирование метрологических характеристик.....	294
8.5. Измерение физических величин, источники ошибок.....	295
8.6. Виды ошибок.....	297

Заключение.....	302
<b>Приложения</b> .....	<b>303</b>
Приложение 1 .....	303
Приложение 2 .....	305
Приложение 3 .....	309
Приложение 4 .....	312
Приложение 5 .....	322
Приложение 6 .....	326
Приложение 7 .....	329
Список литературы.....	330

*Учебное издание*

**Болдин Адольф Петрович,  
Максимов Виктор Александрович**  
**Основы научных исследований**

**Учебник**

Редактор *И. В. Мочалова*  
Технический редактор *Е. Ф. Коржуева*  
Компьютерная верстка: *Г. Ю. Никитина*  
Корректор *А. Ю. Гончарова*

Изд. № 10114728. Подписано в печать 19.12.2011. Формат 60 × 90/16.  
Гарнитура «Ньютон». Бумага офс. № 1. Печать офсетная. Усл. печ. л. 21,0.  
Тираж 1 000 экз. Заказ № 5295.

ООО Издательский центр «Академия». [www.academia-moscow.ru](http://www.academia-moscow.ru)

125252, Москва, ул. Зорге, д. 15, корп. 1, пом. 266.

Адрес для корреспонденции 129085, Москва, пр-т Мира, 101В, стр. 1, а/я 48.  
Тел./факс: (495) 648-0507, 616-0029.

Санитарно-эпидемиологическое заключение № РОСС RU АЕ51. Н 14964  
от 21.12.2010.

Отпечатано с электронных носителей издательства.

ОАО «Тверской полиграфический комбинат», 170024, г. Тверь, пр-т Ленина, 5.  
Телефон: (4822) 44-52-03, 44-50-34. Телефон/факс: (4822) 44-42-15.

Home page — [www.tverpk.ru](http://www.tverpk.ru) Электронная почта (E-mail) — [sales@tverpk.ru](mailto:sales@tverpk.ru)