МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«УФИМСКИЙ ГОСУДАРСТВЕННЫЙ НЕФТЯНОЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебное пособие

2-е издание, исправленное и дополненное

Уфа
Издательство УГНТУ
2016

УДК 519.2(07)
ББК 22.172Я7
М34

Утверждено Редакционно-издательским советом УГНТУ
в качестве учебного пособия

Рецензенты:
Профессор кафедры программирования и вычислительной математики
Башкирского государственного педагогического университета им. М. Акмуллы,
доктор физико-математических наук Р.М. Асадуллин

Профессор кафедры информационных технологий и компьютерной математики
Башкирского государственного университета,
доктор физико-математических наук Н.Д. Морозкин

Авторы:

Бахтизин Р.Н., Галиуллин М.М., Галиакбарова Э.В., Гимаев Р.Г., Зарипов Р.М., Исламгулова Г.Ф., Ковалева Э.А., Лазарев В.А., Майский Р.А., Мухаметзянов И.З., Нагаева З.М., Сахарова Л.А., Сокова И.А., Сулейманов И.Н., Умергалина Т.В., Фаткуллин Н.Ю., Хайбуллин Р.Я., Хакимова З.Р., Чернятьева М.Р., Юлдыбаев Л.Х., Шамшович В.Ф., Шварева Е.Н., Якупов В.М., Янчушка А.П., Абзалимов Р.Р., Акмадиева Т.Р., Аносова Е.П., Ахтямов Н.Т., Байрамгулова Р.С.

М34 Математическая статистика: учеб. пособие / Р.Н.Бахтизин и др., под ред. Т.В.Умергалиной. – 2-е изд., испр. и доп. -Уфа: Изд-во УГНТУ, 2016.-92 с.

ISBN 978-5-7831-1402-1

Учебно-методическое пособие содержит разделы «Теоретические основы», «Методические указания для студентов», «Материалы для самостоятельной работы студентов».

В разделе «Теоретические основы» содержится теоретический материал, необходимый для приобретения студентами навыков обработки статистических данных. Материал изложен в объеме, предусмотренном ФГОС3+. В разделе «Материалы для самостоятельной работы студентов» предложены разработки для выполнения четырех лабораторных работ. Раздел «Материалы для самостоятельной работы студентов» включает в себя: 67 вариантов выборок для выполнения первых трех лабораторных работ и 30 вариантов выборок для четвертой лабораторной работы.

Пособие разработано для студентов, обучающихся по всем направлениям подготовки и специальностям технического вуза.

Первое издание вышло в 2008 году.

УДК 519.2(07)
ББК 22.172Я7

ISBN 978-5-7831-1402-1

СОДЕРЖАНИЕ

1. Теоретические основы

1.1. ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ

Математическая статистика возникла в XVIII в. и создавалась параллельно с теорией вероятностей. Дальнейшее развитие этой дисциплины (начало XX в.) обязано, в первую очередь, П.Л. Чебышеву, А.А. Маркову, А.М.Ляпунову. Основные результаты, ставшие в настоящее время классическими, были получены учеными англо-американской школы – К. Пирсоном, Р.Фишером, Ю.Нейманом, А.Вальдом, В.Феллером и другими и российскими математиками – В.И.Романовским, Е.Е.Слуцким, А.Н.Колмогоровым, Н.В. Смирновым. Годом рождения современной математической статистики следует считать 1933 г. – год опубликования работы академика А.Н.Колмогорова «Основные понятия теории вероятностей». Именно в это время математическую статистику выделили из теории вероятностей в отдельную дисциплину.

1.2. ПРЕДМЕТ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ И ЕЕ ОСНОВНЫЕ ЗАДАЧИ. ВЫБОРКА. СТАТИСТИЧЕСКИЙ РЯД. ЭМПИРИЧЕСКИЙ ЗАКОН РАСПРЕДЕЛЕНИЯ. ПОЛИГОН И ГИСТОГРАММА

В теории вероятностей, если мы изучаем случайную величину X, ее закон распределения считается заданным, и мы можем достоверно ответить на любой вопрос, касающийся данной случайной величины. В математической статистике ситуация прямо противоположная – мы ничего не знаем о законе распределения изучаемой случайной величины X. У нас имеются только некоторые ее наблюдения или измерения. Понятно, что по конечному числу наблюдений невозможно достоверно сделать какие-либо выводы об изучаемой случайной величине. Ясно также, что чем больше таких наблюдений, тем более надежными будут наши приближенные выводы. В этом состоит основная особенность математической статистики – она не определяет достоверно закономерности поведения изучаемых случайных явлений, а оценивает их с той или иной степенью достоверности. Но при неограниченном увеличении числа наблюдений выводы математической статистики становятся практически достоверными. Поэтому содержание этой дисциплины – как и сколько сделать наблюдений и как их обработать, чтобы ответить на интересующий нас вопрос о случайном явлении с требуемой степенью достоверности.

Итак, установление закономерностей, которым подчинены массовые случайные явления, основано на изучении статистических данных - результатов наблюдений. Математическая статистика решает две главные задачи: указать способы сбора и группировки (если данных очень много) статистических сведений (результатов наблюдений) и разработать методы анализа собранных статистических данных в зависимости от целей исследования.

Математическая статистика – это раздел математики, занимающийся разработкой методов сбора, обработки и анализа статистических данных для получения научных и практических выводов. Теория вероятностей является теоретической основой для математической статистики.

Пусть требуется изучить совокупность однородных объектов относительно качественного или количественного признака, характеризующего эти объекты, например некоторое предприятие выпускает партию одинаковых деталей. Если контролируют детали по размеру – это количественный признак.

Можно производить этот контроль сплошным обследованием, то есть измерять каждый из объектов совокупности. Но на практике сплошное обследование применяется редко:

а) из-за очень большого числа объектов;

б) из-за того, что иногда обследование заключается в физическом уничтожении, например, проверяем взрываемость гранат или проверяем на крепость произведенную посуду и т.д.

В таких случаях производится случайный отбор ограниченного (небольшого) числа объектов, которые и подвергают изучению.

Выборочной совокупностью (выборкой) называется совокупность случайно отобранных однородных объектов.

Генеральной совокупностью называется совокупность всех однородных объектов, из которых производится выборка.

Объемом совокупности (выборочной или генеральной) называется **число** объектов этой совокупности.

При наборе выборки можно поступать двояко: после того как объект отобран и над ним произведено наблюдение, он может быть возвращен либо не возвращен в генеральную совокупность. В связи с этим выборки подразделяются на повторные и бесповторные.

Для того чтобы по данным выборки можно было достаточно уверенно судить об интересующем нас признаке генеральной совокупности, необходимо, чтобы объекты выборки правильно его представляли. Это требование коротко формулируется так: выборка должна быть репрезентативной (представительной).

Способы отбора выборки:

1. Отбор, не требующий расчленения генеральной совокупности на части:

а) простой случайный бесповторный;

б) простой случайный повторный.

2. Отбор, при котором генеральная совокупность разбивается на части (если объем генеральной совокупности слишком большой):

а) типический отбор. Объекты отбираются не из всей генеральной совокупности, а из ее «типичных» частей. Например, цех из тридцати станков производит одну и ту же деталь. Тогда отбор делается по одной или по две детали с каждого станка в случайные моменты времени;

б) механический отбор. Например, если нужно выбрать 5% деталей, то выбирают не случайно, а каждую двадцатую деталь;

в) серийный отбор. Объекты выбирают не по одному, а сериями.

Итак, пусть в соответствии с теорией отбора из генеральной совокупности значений некоторого количественного признака произведена выборка объема n:

$$X = \{x_1, x_2, x_3, ..., x_n\},$$

где $x_i (i = 1,..., n)$ - числовые значения, которые могут быть положительными, отрицательными, целыми или дробными.

Таблица 1.1

№	1	2	3	…	n
x	x_1	x_2	x_3	…	x_n

Таблица 1.1. называется <u>простым статистическим рядом</u>, являющимся первичной формой представления статистического материала. Необходимо выяснить (с той или иной степенно достоверности), каким закономерностям подчиняются значения нашей выборки. С этой целью производятся следующие преобразования в выборке.

Из данных табл. 1.1 находят x_{min} и x_{max}, соответственно наименьшее и наибольшее значения выборки. Затем данные табл. 1.1 располагают в порядке возрастания. Тогда выборка $X = \{x_1, x_2, x_3, ..., x_n\}$, записанная в порядке возрастания, называется <u>вариационным рядом</u>, а её значения –<u>вариантами.</u>

Размах выборки – это длина основного интервала $[x_{min}; x_{max}]$, в который попадают все значения выборки. Вычисляется ***размах*** выборки следующим образом: $d = x_{max} - x_{min}$. Затем по формуле

$$k = 1 + 4 \cdot [\lg n], \tag{1.1}$$

где $[\lg n]$ - целая часть числа $\lg n$, определяется число k. Данное число задает количество подынтервалов (классов), на которые разбиваем основной интервал. Длина h подынтервалов и их границы $a_j \left(j = \overline{0, k}\right)$ вычисляются по формулам

$$h = d / k, \tag{1.2}$$

$$a_0 = x_{min}; \quad a_1 = a_0 + h; \dots; \ a_i = a_{i-1} + h; \dots; \ a_k = a_{max}. \tag{1.3}$$

Далее находятся частоты $m_j\left(j=\overline{1,\,k}\right)$ и относительные частоты $\mu_j=\dfrac{m_j}{n}\ \left(j=\overline{1,\,k}\right)$ попадания значений выборки X в j-й подынтервал. Причем для частот должно выполняться равенство $\sum\limits_{j=1}^{k} m_j=n$, а для относительных частот соответственно $\sum\limits_{j=1}^{k}\mu_j=1$.

Результаты проведенных расчетов сводятся в табл. 1.2 и 1.3.

Таблица 1.2

x	$[a_0;\,a_1)$	$[a_1;\,a_2)$	$\ldots$	$[a_{k-1};\,a_k]$
m	m_1	m_2	$\ldots$	m_k

Таблица 1.3

x	$[a_0;\,a_1)$	$[a_1;\,a_2)$	$\ldots$	$[a_{k-1};\,a_k]$
μ	μ_1	μ_2	$\ldots$	μ_k

Далее находятся середины подынтервалов:

$$b_1=\frac{a_1+a_0}{2};\quad b_2=\frac{a_2+a_1}{2};\ \ldots\ ;\quad b_k=\frac{a_k+a_{k-1}}{2}$$

и после этого составляется еще одна таблица (табл. 1.4), которая называется статистическим рядом распределения.

Таблица 1.4

x	b_1	b_2	$\ldots$	b_k
μ	μ_1	μ_2	$\ldots$	μ_k

Статистический ряд распределения является оценкой теоретического ряда распределения и сходится к нему по вероятности. В дальнейшем будем его называть **эмпирическим** (на греческом языке empeiria- опыт) законом распределения исследуемой выборки.

Сгруппированные данные табл. 1.4 несут в себе меньше информации, чем выборочные, так как в них теряется информация о порядке следования выборочных значений. При группировке также фактически происходит округление наблюдаемых значений выборки внутри j-го класса (подынтервала) до значения b_j, что приводит к потере информации о распределении исследуемой случайной величины внутри каждого класса. Преимуществом же сгруппированных данных является их компактность и большая наглядность.

В целях визуального изучения полученных в табл. 1.2, 1.3, 1.4 данных пользуются различными способами их графического изображения. К ним относятся гистограмма и полигон.

Для построения гистограммы относительных частот используются данные табл. 1.3. В декартовой системе координат на оси OX откладываются границы подынтервалов. По оси OY откладываются величины $\mu_J \left(j = \overline{1, k}\right)$.

Гистограммой относительных частот называется ступенчатая фигура, состоящая из прямоугольников, основаниями которых служат частичные подынтервалы длины h, а высоты равны числам $\mu_J \left(j = \overline{1, k}\right)$. Аналогичным образом по данным табл. 1.2 строится гистограмма частот.

Для построения полигона относительных частот используются данные табл. 1.4. В декартовой системе координат на оси OX находятся x_{min} и x_{max}, то есть изображаются границы основного интервала. Затем наносятся значения середин подынтервалов $b_j \left(j = \overline{1, k}\right)$. По оси OY откладываются значения, соответствующие относительным частотам $\mu_j \left(j = \overline{1, k}\right)$.

Полигоном относительных частот называется ломаная, отрезки которой соединяют точки $(b_1; \mu_1)$; $(b_2; \mu_2)$; … ; $(b_k; \mu_k)$. Полигон относительных частот есть визуальное представление эмпирического закона распределения выборки.

<u>Статистической функцией распределения</u> $F^*(x)$ называется относительная частота наступления события $(X < x)$:

$$F^*(x) = \mu(X < x) = \frac{m_x}{n},$$

где $m_x = \sum\limits_{x_i < x} m_i$ – накопительная частота.

Рассмотренные формы статистических (эмпирических) законов распределения тем лучше приближают неизвестные генеральные (теоретические) законы, чем больше объем выборки n, что следует из предельных теорем теории вероятностей.

Любая функция выборки $\varphi = \varphi(x_1, x_2, ..., x_n)$ называется ***статистикой***. Статистика является случайной величиной, так как на различных реализациях выборки она получает различные наблюдаемые значения. Статистиками являются: частоты m_j, границы классов a_j и их середины b_j, размах выборки. Статистический ряд распределения также является статистикой. Из определения статистики следует, что любая функция от статистик также является статистикой, поэтому статистикой является любая функция от сгруппированных данных (см. табл.1.4).

Статистики служат для оценки любых характеристик изучаемой случайной величины: вероятностей случайных событий, связанных с изучаемой величиной, ее числовых характеристик, параметров закона распределения и так далее. Изучение статистик на основе теории вероятностей есть теоретическое ядро математической статистики.

1.3. СТАТИСТИЧЕСКИЕ ОЦЕНКИ ГЕНЕРАЛЬНЫХ ПАРАМЕТРОВ. ТОЧЕЧНЫЕ И ИНТЕРВАЛЬНЫЕ ОЦЕНКИ ДЛЯ МАТЕМАТИЧЕСКОГО ОЖИДАНИЯ И ДИСПЕРСИИ

Пусть изучается случайная величина (СВ) X, необходимо определить ее числовые характеристики. По выборке (результатам n испытаний) $x_1, x_2, ..., x_n$ требуется найти оценку $\tilde{\theta}$ неизвестного параметра $\theta_г$.

Обозначим через X_i случайную величину, представляющую собой результат $i-$го испытания. Статистической оценкой неизвестного параметра $\theta_г$ назовем функцию $\tilde{\theta} = \varphi(X_1, X_2, ..., X_n)$. Тем самым $\tilde{\theta}$ является случайной величиной. Значение этой функции $\tilde{\theta} = \varphi(X_1, X_2, ..., X_n)$ при полученных результатах испытаний $X_i = x_i$, $i = \overline{1, n}$ будем рассматривать как приближенное значение неизвестного параметра $\theta_г$. Приведенное определение оценки отражает только самое общее требование, что оценка должна определяться по значениям выборки.

Оценки параметров подразделяются на точечные и интервальные.

Точечной называется статистическая оценка генерального параметра $\theta_г$, которая определяется одним числом $\tilde{\theta}$.

Интервальной называется оценка генерального параметра $\theta_г$, которая определяется двумя числами $\tilde{\theta}$ и $\tilde{\tilde{\theta}}$ - концами интервала, покрывающего оцениваемый генеральный параметр $\theta_г$.

Для того чтобы точечная оценка давала «хорошие» приближения оцениваемого параметра, она должна быть: несмещенной, эффективной, состоятельной.

Несмещенной называют такую точечную оценку $\tilde{\theta}$, математическое ожидание которой равно оцениваемому генеральному параметру при любом объеме выборки, то есть

$$M\left[\tilde{\theta}\right] = \theta_г. \qquad (1.4)$$

Если равенство (1.4) нарушается, то в этом случае оценка $\tilde{\theta}$ называется ***смещенной***.

Эффективной называется точечная оценка $\tilde{\theta}$, которая (при заданном объеме выборки) имеет наименьшую возможную дисперсию, то есть

$$D\left(\tilde{\theta}\right) = \min. \tag{1.4а}$$

Состоятельной называется точечная оценка $\tilde{\theta}$, которая (с увеличением объема выборки) стремится по вероятности к оцениваемому параметру $\theta_г$, то есть для любого достаточно малого $\delta > 0$

$$\lim_{n\to\infty} P\left(\left|\theta_г - \tilde{\theta}\right| < \delta\right) = 1. \tag{1.4б}$$

Математическое ожидание $M_г(x)$ случайной величины X

В качестве оценки математического ожидания случайной величины X по выборке $\{x_1, x_2, x_3, ..., x_n\}$ используют выборочное среднее $\bar{x}_в$ – выборочное математическое ожидание: если выборка сгруппирована и для этой выборки составлена табл.1.4, то

$$\bar{x}_в = \sum_{j=1}^{k} b_j \mu_j, \tag{1.5}$$

или за $\bar{x}_в$ принимают среднее арифметическое значений выборки:

$$\bar{x}_в = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{1.5а}$$

Эта оценка является несмещенной, состоятельной (см. [1], [4] из основной литературы). Эффективность данной оценки зависит от вида закона распределения нашей исследуемой совокупности.

Дисперсия $D_г(x)$ случайной величины X

В качестве оценки дисперсии $D_г$ случайной величины X по выборке $\{x_1, x_2, x_3, ..., x_n\}$ возьмем выборочную дисперсию $D_в$. Выборочной дисперсией $D_в$ называют среднее арифметическое квадратов отклонений наблюдаемых значений x_i от среднего значения $\bar{x}_в$:

$$D_в = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}_в\right)^2 \tag{1.6}$$

или по данным табл. 1.4

$$D_в = \sum_{j=1}^{k} \left(b_j - \bar{x}_в \right)^2 \mu_j, \tag{1.6a}$$

где $\mu_j = \dfrac{m_j}{n}, \quad j = \overline{1,k}$.

Формулы для вычисления дисперсии можно упростить. Покажем это на примере формулы (1.6 а):

$$D_в = \frac{1}{n} \sum_{j=1}^{k} m_j \left(b_j - \bar{x}_в \right)^2 = \frac{1}{n} \sum_{j=1}^{k} \left(b_j^2 - 2 b_j \ \bar{x}_в + \bar{x}_в^2 \right) m_j =$$

$$= \frac{1}{n} \sum_{j=1}^{k} b_j^2 \, m_j - 2 \bar{x}_в \cdot \frac{1}{n} \sum_{j=1}^{k} b_j \, m_j + \bar{x}_в^2 \sum_{j=1}^{k} m_j =$$

$$= M\left[\bar{x}_в^2 \right] - 2 \bar{x}_в \ \bar{x}_в + \left(\bar{x}_в \right)^2 = M\left[\bar{x}_в^2 \right] - \bar{x}_в^2 .$$

Итак, мы доказали теорему:

Теорема. Дисперсия $D_в$ равна среднему квадратов значений вариант минус квадрат общей средней:

$$D_в = \overline{x_в^2} - \bar{x}_в^2, \tag{1.7}$$

где $\overline{x_в^2} = \dfrac{1}{n} \sum\limits_{j=1}^{k} b_j^2 \, m_j; \quad \bar{x}_в = \dfrac{1}{n} \sum\limits_{j=1}^{k} b_j \, m_j$. Здесь b_j - данные из табл. 1.4;

m_J - данные из табл. 1.2.

Данная оценка является состоятельной, но $D_В$ является смещенной оценкой (см. [1] в дополнительной литературе).

Выборочную дисперсию $D_В$ можно легко исправить так, чтобы получить несмещенную оценку. В качестве несмещенной оценки дисперсии $D_г$ используют исправленную дисперсию s^2:

$$s^2 = \frac{n}{n-1} D_в . \tag{1.8}$$

Исправленная дисперсия s^2 остается состоятельной, так как $\lim\limits_{n \to \infty} \dfrac{n}{n-1} = 1$ и множитель $\dfrac{n}{n-1}$ не влияет на состоятельность оценки. При больших объемах выборки n различие между $D_г$ и s^2 становится незначительным, а при малых

n в качестве характеристики рассеивания надо использовать исправленную дисперсию.

Выборочным средним квадратическим отклонением называют число $\sigma_в$, равное $\sigma_в = \sqrt{D_в}$.

Стандартным отклонением s называют корень квадратный из исправленной дисперсии.

При выборке малого объема $(n < 30)$ точечная оценка может значительно отличаться от оцениваемого генерального параметра, то есть приводить к грубым ошибкам. Поэтому при небольшом объеме выборки следует пользоваться **интервальными оценками.**

Пусть найденная (по данным выборки) статистическая оценка $\tilde{\theta}$ является оценкой неизвестного генерального параметра $\theta_г$. Ясно, что $\tilde{\theta}$ тем точнее определяет $\theta_г$, чем меньше значение разности $\left|\theta_г - \tilde{\theta}\right|$. То есть при $\left|\theta_г - \tilde{\theta}\right| < \delta$ $(\delta > 0)$ чем меньше δ, тем оценка $\theta_г$ точнее. Значит, положительное число δ характеризует ***точность*** оценки.

Надежностью (доверительной вероятностью) оценки $\tilde{\theta}$ называется вероятность γ, с которой осуществляется событие $\left|\theta_г - \tilde{\theta}\right| < \delta$, то есть

$$\gamma = P\left(\left|\theta_г - \tilde{\theta}\right| < \delta\right). \tag{1.9}$$

Обычно надежность оценки (доверительная вероятность γ) задается. Причем в качестве γ берут число, близкое к единице (0,95; 0,99; 0,999).

Доверительным называется интервал, который с заданной надежностью γ покрывает оцениваемый генеральный параметр. В соотношении (1.9), если раскрыть модуль, получается $P\left(-\delta < \theta_г - \tilde{\theta} < \delta\right) = \gamma$ или $P\left(\tilde{\theta} - \delta < \theta_г < \tilde{\theta} + \delta\right) = \gamma$. Тогда интервал $\left(\tilde{\theta} - \delta;\ \tilde{\theta} + \delta\right)$ и есть доверительный интервал. Из общих соображений ясно, что длина доверительного интервала будет зависеть от объема выборки n и доверительной вероятности γ.

Построение ***доверительного интервала*** для оценки математического ожидания ***при известном*** σ

Пусть случайная величина X распределена нормально, где $\sigma -$ известно. Требуется по выборке $\{x_1, x_2, x_3, ..., x_n\}$ оценить неизвестное математическое ожидание $M_г(x)$. Наилучшей оценкой генерального математического ожидания является выборочное среднее $\overline{x}_в$ (см. формулу (1.5) или (1.5а)). Так как случайная величина X имеет нормальное распределение, то $\overline{x}_{в}$ является случайной величиной также с нормальным распределением. Причем $M(\overline{x}_{в}) = M_{г}(x)$; $D(\overline{x}_{в}) = \frac{\sigma^2}{n}$ и $\sigma(\overline{x}_{в}) = \frac{\sigma}{\sqrt{n}}$.

Потребуем, чтобы выполнялось равенство (1.9): γ=P(|$M_{г}(x)$-$\overline{x}_{в}$|<δ), где γ- заданное число.

Из теории вероятностей (см.[1] из основной литературы) известна формула $P(\ |a - X| < \alpha) = 2\Phi\,(\frac{\alpha}{\sigma})$, где $\Phi\,(\frac{\alpha}{\sigma})$ - функция Лапласа. Заменим параметры последнего соотношения нашими обозначениями: a на $M_{г}(x)$,

X на $\overline{x}_{в}$, α на δ, σ на $\sigma(\overline{x}_{в}) = \frac{\sigma}{\sqrt{n}}$:

(*) $\gamma = P(\ |M_{г}(x) - \overline{x}_{в}| < \delta) = 2\Phi\,(\frac{\delta \cdot \sqrt{n}}{\sigma})$. Введем обозначение

$t = \frac{\delta \cdot \sqrt{n}}{\sigma}$, тогда $\delta = \frac{t \cdot \sigma}{\sqrt{n}}$ называется **точностью** оценки. Теперь запишем

соотношение (*) в следующем виде: $P(\ |M_{г}(x) - \overline{x}_{в}| < \frac{t \cdot \sigma}{\sqrt{n}}) = \gamma = 2\Phi\,(t)$

или $$P\ (\ \overline{x}_{в} - \frac{t \cdot \sigma}{\sqrt{n}}\ <\ M_{г}(x)\ <\ \overline{x}_{в} + \frac{t \cdot \sigma}{\sqrt{n}}\ )\ =\ \gamma. \qquad (1.10)$$

Так как γ нам заранее задано, тогда из соотношения γ=2Φ(t) имеем Φ(t)= = γ/2 и по таблице значений функции Лапласа (прил. 1) находим её аргумент, то есть t, чтобы подставить в (1.10).

Смысл соотношения (1.10) таков: с надежностью γ можно утверждать, что доверительный интервал ($\overline{x}_{в} - \frac{t \cdot \sigma}{\sqrt{n}}$; $\overline{x}_{в} + \frac{t \cdot \sigma}{\sqrt{n}}$) покрывает неизвестный параметр $M_{г}(x)$. Точность оценки равна $\delta = \frac{t \cdot \sigma}{\sqrt{n}}$.

Аналогично могут быть построены интервальные оценки для дисперсии, коэффициента корреляции и других параметров.

Построение ***доверительного интервала*** для оценки математического ожидания ***при неизвестном*** σ

Пусть случайная величина X –распределена по нормальному закону, где m и σ -генеральные параметры неизвестны.

Требуется по выборке x_1, x_2, ... , x_n построить интервальную оценку неизвестного математического ожидания. Найдем точечные несмещенные оценки для m и σ: $\overline{x}_{в} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $s = \sqrt{\frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}_{в}^2\right)}$. По данным выборки можно построить случайную величину $T=\frac{\overline{x}-m}{s}\sqrt{n}$, которая имеет распределение Стьюдента $S(t,n)$ с $k=n-1$ степенями свободы. Распределение Стьюдента зависит только от объема выборки n и при $n\to\infty$ переходит в нормированное нормальное распределение $N(0,1)$. На практике при $n>30$ t - распределение можно заменять распределением $N(0,1)$. На рис.1.1 представлен график t - распределения.

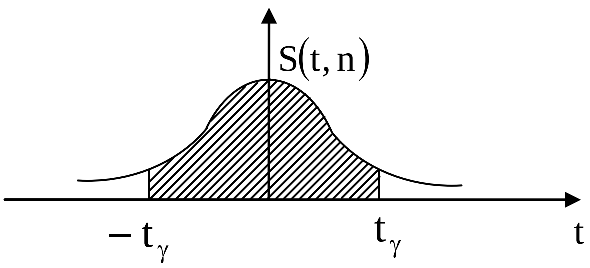

Рис. 1.1

Доверительный интервал для СВ T определяется равенством

$$P\left(|T|<t_\gamma\right)=2\int_0^{t_\gamma}S(t,n)dt=\gamma .$$

По таблице распределения Стьюдента по γ и $k=n-1$ можно найти t_γ. Преобразуем неравенство:

$$|T|<t_\gamma \iff \left|\frac{\overline{x}-m}{s}\sqrt{n}\right|<t_\gamma \iff |\overline{x}-m|<\frac{t_\gamma\cdot s}{\sqrt{n}} \iff$$

$$\overline{x}-\frac{t_\gamma\cdot s}{\sqrt{n}}<m<\overline{x}+\frac{t_\gamma\cdot s}{\sqrt{n}}$$

Откуда

$$P\left(\overline{x}-\frac{t_\gamma\cdot s}{\sqrt{n}}<m<\overline{x}+\frac{t_\gamma\cdot s}{\sqrt{n}}\right)=\gamma=1-\alpha .$$

Тем самым по выборке, пользуясь распределением Стьюдента, можно построить доверительный интервал

$$\left(\overline{x}_в-\frac{t_\gamma\cdot s}{\sqrt{n}};\ \overline{x}_в+\frac{t_\gamma\cdot s}{\sqrt{n}}\right),$$

покрывающий неизвестный параметр m с надежностью γ.

Метод моментов

Снова возвратимся к точечным оценкам. Были сформулированы основные свойства оценок, но ничего не было сказано о способах их получения.

Одним из первых методов оценивания параметров был метод моментов, разработанный К.Пирсоном. Достоинство метода – его простота. Метод моментов точечной оценки неизвестных параметров распределения состоит в приравнивании теоретических моментов рассматриваемого распределения соответствующим выборочным моментам того же порядка.

Пусть СВ X – задана плотностью $f(x,\theta)$, где θ - неизвестный параметр. Надо найти его точечную оценку. Для оценки одного параметра достаточно иметь одно уравнение, содержащее этот параметр. Приравняем первые начальные моменты: $\alpha_1 = \tilde{\alpha}_1 \Leftrightarrow M[X] = \overline{x}_в$, где $M[X] = \int\limits_{-\infty}^{\infty} xf(x,\theta)dx = \varphi(\theta)$.

Имеем уравнение: $\varphi(\theta) = \overline{x}_в$, решив которое, можно найти оценку θ неизвестного параметра: $\tilde{\theta} = \psi(x_1, x_2, ..., x_n)$.

ПРИМЕР 1.1. По выборке x_1, x_2, ..., x_n методом моментов найти точечную оценку неизвестного параметра λ показательного распределения $f(x) = \lambda \cdot e^{-\lambda x}$, $x \geq 0$.

<u>Решение</u>. Приравняем $M[X] = \overline{x}_в = \frac{1}{n}\sum\limits_{i=1}^{n} x_i$. Для показательного распределения:

$$M[X] = \frac{1}{\lambda} = \overline{x}_в \Rightarrow \tilde{\lambda} = \frac{1}{\overline{x}_в}.$$

Ответ: $\tilde{\lambda} = \frac{1}{\overline{x}_в}$.

Пусть СВ X имеет плотность распределения $f(x, \theta_1, \theta_2)$ с двумя неизвестными параметрами. Для отыскания оценок двух параметров надо иметь два уравнения относительно этих параметров. Приравняем $\alpha_1 = \tilde{\alpha}_1$ и $\mu_1 = \tilde{\mu}_2$, где $\alpha_1 = M[X]$, $\tilde{\alpha}_1 = \overline{x}_в$, $\mu_2 = D$, $\tilde{\mu}_2 = D_в$. Имеем систему:

$$\begin{cases} M[X] = \overline{x}_в \\ D[X] = D_в \end{cases} \Leftrightarrow \begin{matrix} \int\limits_{-\infty}^{\infty} xf(x,\theta_1,\theta_2)dx = \overline{x}_в \\ \int\limits_{-\infty}^{\infty} (x - m_x)^2 f(x,\theta_1,\theta_2)dx = D_в \end{matrix}$$

Решив систему, найдем $\tilde{\theta}_1 = \psi_1(x_1, x_2, ..., x_n)$, $\tilde{\theta}_2 = \psi_2(x_1, x_2, ..., x_n)$. Метод моментов дает состоятельные оценки.

ПРИМЕР 1.2. По выборке x_1, x_2, ..., x_n методом моментов найти оценки неизвестных параметров m и σ нормального распределения.

<u>Решение</u>. Имеем $\begin{cases} M[X] = \overline{x}_в \\ D[X] = D_в \end{cases} \Leftrightarrow \begin{cases} m = \overline{x}_в \\ \sigma^2 = D_в \end{cases}$.

Ответ: $\begin{cases} m = \overline{x}_в \\ \sigma = \sqrt{D_в} \end{cases}.$

Метод наибольшего правдоподобия

Основным методом получения точечных оценок неизвестных параметров распределения является метод наибольшего правдоподобия. Рассмотрим основную идею метода.

Пусть по результатам выборки требуется оценить неизвестный параметр θ случайная величина X с законом распределения $f(x,\theta)$.

Функция $L = f(x_1,\theta)\cdot f(x_2,\theta)\cdot \ldots \cdot f(x_n,\theta)$ (1.11)
называется функцией правдоподобия.

Для дискретной случайной величины функция правдоподобия имеет вид $L = P(x_1,\theta)\cdot P(x_2,\theta)\cdot \ldots \cdot P(x_n,\theta)$ где $P(X_k = x_k) = P(x_к,\theta)$.

При фиксированных $x_1,\ x_2,\ \ldots,\ x_n$ функцию L будем рассматривать как функцию от параметра θ. По методу наибольшего правдоподобия за оценку параметра θ принимают значение аргумента $\tilde{\theta} = \tilde{\theta}(x_1,\ x_2,\ \ldots,\ x_n)$, при котором L имеет максимальное значение. Поскольку $\ln L$ при фиксированных $(x_1,\ x_2,\ \ldots,\ x_n)$ достигает максимума при том же значении параметра θ, что и L, то для нахождения оценки решают уравнение правдоподобия:

$$\frac{\partial \ln L}{\partial \theta} = 0. \qquad (1.12)$$

Если закон распределения случайной величины X зависит от двух параметров $f(x,\theta_1,\theta_2)$, то функция правдоподобия $L = f(X_1,\theta_1,\theta_2)\cdot f(X_2,\theta_1,\theta_2)\cdot\ldots\cdot f(X_n,\theta_1,\theta_2)$, а уравнения правдоподобия имеют вид: $\frac{\partial \ln L}{\partial Q_1} = 0,\ \frac{\partial \ln L}{\partial Q_2} = 0$.

ПРИМЕР 1.3. Методом наибольшего правдоподобия найти оценку параметра λ распределения Пуассона.

<u>Решение</u>. Распределение Пуассона: $P_n(m) = \frac{\lambda^m}{m!}e^{-\lambda}$. Составим функцию правдоподобия:

$$L = P(x_1,\lambda)\cdot P(x_2,\lambda)\cdot \ldots \cdot P(x_n,\lambda) = \frac{\lambda^{x_1}}{x_1!}e^{-\lambda}\cdot\frac{\lambda^{x_2}}{x_2!}e^{-\lambda}\cdot\ \ldots\ \cdot\frac{\lambda^{x_n}}{x_n!}e^{-\lambda};$$

$$\ln L = \sum_{i=1}^{n} x_i \cdot \ln\lambda - n\lambda - \ln(x_1!\cdot x_2!\cdot\ldots\cdot x_n!).$$

Уравнение правдоподобия $\frac{d\ln L}{d\lambda} = 0 \Rightarrow \frac{\sum_{i=1}^{n} x_i}{\lambda} - n = 0 \Rightarrow \lambda = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}_в,$

т.к. $\dfrac{d^2 \ln L}{d\lambda^2} = -\dfrac{1}{\lambda^2}\sum_{i=1}^{n} x_i < 0$, то $\hat{\lambda} = \overline{x}_в$ - точка максимума.

Ответ: оценка наибольшего правдоподобия $\hat{\lambda} = \overline{x}_в$.

ПРИМЕР 1.4. Методом наибольшего правдоподобия найти оценки параметров $m = M_г(x)$ и $\sigma = \sqrt{D_г(x)}$ нормального распределения: $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$ по выборке $x_1,\ x_2,\ \ldots,\ x_n$.

<u>Решение</u>. Функция правдоподобия

$$L = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_1-m)^2}{2\sigma^2}} \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_2-m)^2}{2\sigma^2}} \cdot \ldots \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_n-m)^2}{2\sigma^2}} \Rightarrow$$

$$L = \frac{1}{\left(\sigma\sqrt{2\pi}\right)^n} e^{-\sum_{i=1}^{n}(x_i-m)^2/2\sigma^2};$$

$$\ln L = -n\ln\sigma + \ln\frac{1}{\left(\sqrt{2\pi}\right)^n} - \frac{\sum_{i=1}^{n}(x_i - m)^2}{2\sigma^2};$$

$$\frac{\partial \ln L}{\partial m} = \frac{\sum_{i=1}^{n} x_i - nm}{\sigma^2}; \quad \frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{\sum_{i=1}^{n}(x_i - m)^2}{\sigma^3}.$$

Уравнения правдоподобия:

$$\frac{\sum_{i=1}^{n} x_i - nm}{\sigma^2} = 0; \quad -\frac{n}{\sigma} + \frac{\sum_{i=1}^{n}(x_i - m)^2}{\sigma^3} = 0 \quad \Rightarrow$$

$$\tilde{m} = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}_в; \quad \tilde{\sigma}^2 = \frac{\sum_{i=1}^{n}(x_i - m)^2}{n} = D_в.$$

Оценка для дисперсии получилась смещенной.

Метод наименьших квадратов

Метод наибольшего правдоподобия всегда приводит к состоятельным оценкам, хотя иногда и смещенным, и наилучшим образом использует всю информацию о неизвестном параметре, содержащуюся в выборке. Однако часто его применение связано с необходимостью решения сложных систем уравнений.

Другим способом, имеющим большое практическое применение в задачах оценивания неизвестных параметров генеральной совокупности по выборке и часто приводящим к более простым выкладкам, является метод наименьших квадратов (МНК). Подробнее о МНК смотри [10] – учебные пособия кафедры.

Идея МНК применительно к оцениванию параметров сводится к тому, чтобы в качестве оценки неизвестного параметра принимать значение, которое минимизирует сумму квадратов отклонений между оценкой и параметром для всех значений выборки:

$$\sum_{i=1}^{n}\left(\theta-\hat{\theta}(x_i)\right)^2=\min. \qquad (1.13)$$

Если X – нормально распределенная случайная величина с параметрами $m=M_г(x)$ и $\sigma=\sqrt{D_г(x)}$, то по методу наибольшего правдоподобия необходимо максимизировать

$$L=\frac{1}{\left(\sigma\sqrt{2\pi}\right)^n}\; e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-m)^2},$$

что приводит к минимизации $\sum_{i=1}^{n}(x_i-m)^2$. Условие $(x_i-m)^2=\min$ является требованием МНК.

Итак, если случайная величина имеет нормальное распределение, то метод максимального правдоподобия и метод наименьших квадратов дают одинаковые результаты.

Особенно часто МНК применяют в задачах выравнивания или сглаживания статистических рядов.

Пусть в результате испытаний получен ряд точек $(x_1,\ y_1)$, $(x_2,\ y_2)$, ... $(x_n,\ y_n)$, изображенных на рис. 1.2.

Пусть заранее известно, что зависимость между X и Y линейная: $Y=A+BX$ и разброс обусловлен случайными ошибками. Требуется найти оценки параметров A, B, которые наилучшим образом в смысле МНК описывали бы искомую зависимость по результатам испытаний. В качестве оценок A и B выбираются значения $\hat{A}$ и $\hat{B}$, минимизирующие сумму квадратов отклонений измеренных значений y_i от вычисленных $y_{iт}=A+Bx_i$:

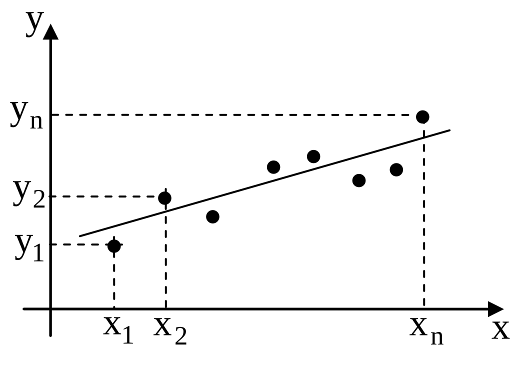

Рис. 1.2

$$Q(A,B)=\sum_{i=1}^{n}\left(y_i-y_{iт}\right)^2=\min \quad\Leftrightarrow\quad Q(A,B)=\sum_{i=1}^{n}\left(y_i-A-Bx_i\right)^2=\min \quad .$$

Условия экстремума функции двух переменных $Q(A,B)$:

$$\frac{\partial Q}{\partial A}=0;\quad \frac{\partial Q}{\partial B}=0.$$

Имеем:

$$\begin{cases} 2\sum_i (y_i - A - Bx_i)(-1) = 0 \\ 2\sum_i (y_i - A - Bx_i)(-x_i) = 0 \end{cases} \Leftrightarrow \begin{cases} nA + B\sum_i x_i = \sum_i y_i \\ A\sum_i x_i + B\sum_i x_i^2 = \sum_i x_i y_i \end{cases} \qquad (1.14)$$

Решая систему (1.14), найдем $\hat{A}$ и $\hat{B}$. Можно показать, что при найденных $\hat{A}$ и $\hat{B}$ $Q(A,B) = \min$. Тогда $y = \hat{A} + \hat{B}x$ - искомая зависимость.

Рассмотренные в данном разделе методы статистического оценивания параметров распределения используются на практике для обработки результатов измерений. Навыки применения этих методов необходимы для проведения исследовательских работ.

1.4. ПРОВЕРКА СТАТИСТИЧЕСКОЙ ГИПОТЕЗЫ О ЗАКОНЕ РАСПРЕДЕЛЕНИЯ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ

Статистической ***гипотезой*** $\boldsymbol{H}_0$ называется любое предположение относительно закона распределения исследуемой случайной величины X.

Гипотезы бывают **простые и сложные**. Простая гипотеза полностью определяет закон распределения величины X в отличии от сложной.

Гипотезы бывают **параметрическими и непараметрическими**. В первом случае мы имеем предположение о параметрах распределения при известном законе, а во втором – о самом виде закона распределения.

ПРИМЕР 1.5. Гипотеза H_0 о том, что математическое ожидание нормальной случайной величины равно $\bar{x}_в$ при условии, что дисперсия $D_г$ известна, является простой параметрической. Если же дисперсия $D_г$ неизвестна, то гипотеза будет сложной параметрической.

ПРИМЕР 1.6. Гипотеза H_0 о том, что случайная величина распределена по нормальному (или по какому-то другому) закону, является сложной непараметрической.

Наряду с выдвинутой гипотезой H_0 рассматривают противоречащую ей гипотезу H_1. Если выдвинутая гипотеза H_0 будет отвергнута, то имеет место противоречащая ей гипотеза H_1.

Критерием проверки статистической гипотезы называется некоторое правило, позволяющее принять ее или отвергнуть. Причем критерии строятся с помощью случайной величины K (часто именно ее называют критерием), для которой известно распределение. Наблюдаемым значением критерия $K_{набл.}$ называют значение критерия, вычисленное по данным выборки.

В случае проверки гипотез возможны ошибки:

Ошибка 1-го рода состоит в том, что будет отвергнута правильная гипотеза. Вероятность ошибки первого рода α называется уровнем значимости критерия, по которому производится проверка.

Ошибка 2-го рода состоит в том, что будет принята неправильная гипотеза. Если β - вероятность ошибки второго рода, то величина $1-\beta$ называется мощностью критерия.

Параметрические гипотезы проверяются с помощью критериев ***значимости***, а непараметрические – с помощью критериев ***согласия***.

Критической областью называется совокупность значений критерия, при которых нулевую гипотезу отвергают. Если уровень значимости α уже выбран и задан объем выборки, то критическую область следует строить так, чтобы мощность критерия была максимальной. Выполнение этого требования должно обеспечить минимальную ошибку второго рода, что более желательно.

Основной принцип проверки статистических гипотез: если $K_{\text{набл.}}$ принадлежит критической области – гипотезу H_0 отвергают, если же $K_{\text{набл.}}$ принадлежит области принятия гипотезы, то гипотезу H_0 принимают.

Остановимся только на гипотезе о законе распределения генеральной совокупности.

В разделе 1.2 при группировке данных выборочной совокупности получена табл. 1.4 – эмпирический закон распределения выборки X и по данным этой таблицы можно построить полигон относительных частот. Относительные частоты иногда называют эмпирическими вероятностями. Из визуального наблюдения полигона делается вывод (выдвигается гипотеза H_0) о законе распределения, например H_0: генеральная совокупность распределена по нормальному закону.

И выдвигается гипотеза, противоречащая гипотезе H_0 или ее отвергающая.

Проверка гипотезы о предполагаемом законе неизвестного распределения производится при помощи специально подобранной случайной величины – ***критерия согласия***. Разработано несколько таких критериев: χ^2 - Пирсона, Колмогорова, Смирнова и др. Рассмотрим критерий χ^2 - Пирсона, как классический пример применительно к проверке гипотезы о нормальном законе распределения генеральной совокупности.

Будем сравнивать эмпирические и теоретические вероятности. Обычно они различаются. Случайно ли это расхождение? Возможно, что расхождение случайно (незначимо) и объясняется либо малым числом исходных данных, либо способом их группировки или другими причинами. Возможно, что данное расхождение неслучайно (значимо) и объясняется тем, что теоретические вероятности вычислены исходя из неверной гипотезы о нормальном распределении генеральной совокупности.

Пусть нам задан уровень значимости α (γ - доверительная вероятность, то есть вероятность принять верную гипотезу; α - это вероятность отвергнуть верную гипотезу, причем $\alpha+\gamma=1$). Для того чтобы при заданном α проверить гипотезу о нормальном распределении генеральной совокупности, надо вычислить теоретические вероятности. Плотность распределения для нормального закона есть функция:

$$f(x)=\frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot e^{-(x-M[x])^2/(2\cdot D[x])}. \tag{1.15}$$

Тогда, пользуясь формулой нахождения вероятности попадания случайной величины в интервал:

$$p(a<x<b)=\int_a^b f(x)dx,$$

имеем для всех $j=\overline{1,k}$:

$$p_j=p(a_{j-1}<x<a_j)=\frac{1}{\sqrt{2\pi}\cdot\sigma_в}\cdot\int_{a_{j-1}}^{a_j} e^{-(x-\overline{x}_в)^2/(2\cdot D_в)}dx=$$

$$=\frac{1}{\sqrt{2\pi}\cdot\sigma_в}\cdot e^{(b_j-\overline{x}_в)^2/(2\cdot D_в)}\cdot\int_{a_{j-1}}^{a_j}dx=\frac{1}{\sqrt{2\pi}\cdot\sigma_в}\cdot e^{-(b_j-\overline{x}_в)^2/(2\cdot D_в)}\cdot h, \tag{1.16}$$

где a_j $(j=\overline{0,k})$ - границы частичных подынтервалов (см.табл. 1.2);

b_j - середина $j-$го частичного подынтервала;

h - длина частичного подынтервала (см. формулу (1.2)).

Составляется сводная таблица на основе данных табл.1.4 и рассчитанных теоретических вероятностей:

Таблица 1.5

x	b_1	b_2	$\dots$	b_j	$\dots$	b_k	
μ	μ_1	μ_2	$\dots$	μ_j	$\dots$	μ_k	эмпирические вероят.
p	p_1	p_2	$\dots$	p_j	$\dots$	p_k	теоретические вероят.

Оценка отклонения эмпирических вероятностей μ_j $(j=\overline{1,k})$ от теоретических вероятностей p_j $(j=\overline{1,k})$ производится с помощью критерия Пирсона χ^2:

$$\chi^2 = \sum_{j=1}^{k} \frac{(\mu_j - p_j)^2 \cdot n}{p_j}. \qquad (1.17)$$

Эта величина случайная, так как в различных опытах она принимает различные заранее неизвестные значения. Ясно, что чем меньше различаются эмпирические и теоретические частоты, тем меньше величина критерия и, следовательно, он характеризует близость эмпирического и теоретического распределений.

Доказано, что при $n \to \infty$ закон распределения случайной величины (1.17) независимо от того, какому закону распределения подчинена генеральная совокупность, стремится к закону распределения χ^2 с r степенями свободы. Поэтому случайная величина (1.17) обозначена через χ^2, а сам критерий называют ***критерием согласия*** «хи квадрат».

По таблице критических точек распределения χ^2 (прил. 4) по заданному уровню значимости α и числу степеней свободы $r = k - 1 - s$ (k - количество подынтервалов, s - число параметров предполагаемого распределения) находим критическое значение $\chi^2_{кр}(\alpha, r)$ правосторонней критической области.

Правило 1.1. Надо вычислить наблюдаемое значение критерия $\chi^2_{набл.} = \sum_{j=1}^{k} \frac{(\mu_j - p_j)^2 \cdot n}{p_j}$, если $\chi^2_{набл.} < \chi^2_{кр.}$, тогда нет оснований отвергать гипотезу H_0 о нормальном законе распределения генеральной совокупности (то есть эмпирические и теоретические частоты различаются незначимо (случайно)).

Правило 1.2. Если $\chi^2_{набл.} > \chi^2_{кр.}$, тогда гипотеза H_0 отвергается.

1.5. СТАТИСТИЧЕСКАЯ И КОРРЕЛЯЦИОННАЯ ЗАВИСИМОСТИ. ЭМПИРИЧЕСКАЯ И ТЕОРЕТИЧЕСКАЯ ЛИНИИ РЕГРЕССИИ

Две случайные величины могут быть связаны между собой функциональной зависимостью, либо зависимостью другого рода, называемой статистической, либо быть независимыми. Строгая функциональная зависимость для случайных величин реализуется редко, так как обе величины (или одна из них) подвержены различным случайным факторам.

ПРИМЕР 1.7. Рассмотрим две таблицы значений, которые принимают случайные величины X и Y.

Таблица 1.6

X	1	7	13	19	25	31	37
Y	15	3	25	2	10	16	8

Таблица 1.7

X	1	7	13	19	25	31	37
Y	20	12	15	9	9	3	0

Изобразим эти данные в декартовой системе координат, откладывая значения случайной величины X по оси OX, а значения случайной величины Y по оси OY.

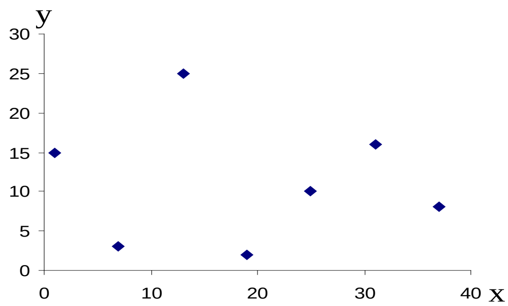

Рис. 1.3

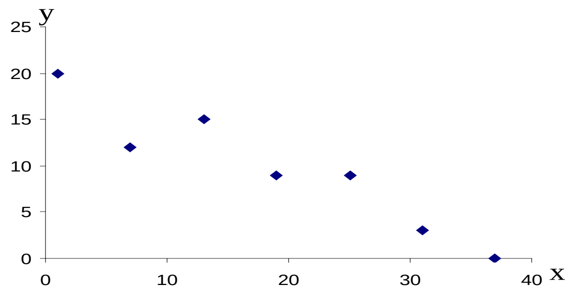

Рис. 1.4

Данные табл.1.6 представлены на рис.1.3, данные табл. 1.7 – на рис.1.4. Из рис.1.3 видно, что данные табл.1.6 не связаны между собой. А вот из рис.1.4 просматривается какая-то зависимость между X и Y, причем выражена обратная зависимость: с увеличением значений случайной величины X, значения случайной величины Y уменьшаются.

Статистической называют зависимость, при которой изменение одной из величин влечет изменение распределения другой. Если при изменении одной из величин изменяется среднее значение другой, то в этом случае статистическую зависимость называют ***корреляционной.*** Значит, корреляционная зависимость есть частный случай статистической зависимости.

Чтобы установить наличие и характер связи между двумя случайными величинами X и Y, нужно привести к удобному виду исходный цифровой материал. Наглядной (удобной) формой представления данных является корреляционная таблица.

Таблица 1.8

X \ Y	y_1	y_2	$\cdots$	y_j	$\cdots$	y_ℓ	m_x
x_1	m_{11}	m_{12}	$\cdots$	m_{1j}	$\cdots$	$m_{1\ell}$	m_{x_1}
x_2	m_{21}	m_{22}	$\cdots$	m_{2j}	$\cdots$	$m_{2\ell}$	m_{x_2}
…	…	…	…	…	…	…	…
x_i	m_{i1}	m_{i2}	$\cdots$	m_{ij}	$\cdots$	$m_{i\ell}$	m_{x_i}
…	…	…	…	…	…	…	…
x_k	m_{k1}	m_{k2}	$\cdots$	m_{kj}	$\cdots$	$m_{k\ell}$	m_{x_k}
m_y	m_{y_1}	m_{y_2}	$\cdots$	m_{y_j}	$\cdots$	m_{y_ℓ}	m

Здесь $x_1, x_2, \ldots, x_k$; $y_1, y_2, \ldots, y_\ell$ - середины подынтервалов сгруппированных выборок X и Y (см. 1.2); m_{ij} - частота, с которой встречается пара $(x_i; y_j)$. В последнем столбце и в последней строке таблицы помещены суммарные частоты, соответствующие значению $X = x_i$, и соответственно $Y = y_i$, то есть

$$m_{x_i} = m_{i1} + m_{i2} + \ldots + m_{i\ell};$$

$$m_{y_j} = m_{1j} + m_{2j} + \ldots + m_{kj},$$ тогда должно быть

$$m_{x_1} + m_{x_2} + \ldots + m_{x_k} = \sum_{i=1}^{k} m_{x_i} = m$$ и

$$m_{y_1} + m_{y_2} + \ldots + m_{y_\ell} = \sum_{i=1}^{\ell} m_{y_j} = m.$$

m - общее количество пар значений $(x_i; y_j)$.

Каждая $i-$я строка табл. 1.8 представляет собой (совместно с первой строкой) некоторое распределение случайной величины Y, соответствующее данному значению случайной величины $X = x_i$. Такое распределение называется ***условным распределением***. Последняя строка табл.1.8 совместно с первой строкой образует безусловное распределение случайной величины Y (ее эмпирический закон распределения):

Таблица 1.9

Y	y_1	y_2	$\cdots$	y_j	$\cdots$	y_ℓ
m_y	m_{y_1}	m_{y_2}	$\cdots$	m_{y_j}	$\cdots$	m_{y_ℓ}

Каждый $j-$й столбец табл.1.8 представляет собой совместно с первым столбцом некоторое распределение случайной величины X, соответствующее данному значению случайной величины $Y = y_j$ (то есть условное распределение). Последний столбец табл. 1.8 совместно с первым столбцом образует безусловное распределение случайной величины X (ее эмпирический закон распределения):

Таблица 1.10

X	x_1	x_2	$\cdots$	x_i	$\cdots$	x_k
m_x	m_{x_1}	m_{x_2}	$\cdots$	m_{x_i}	$\cdots$	m_{x_k}

По данным табл. 1.9 и 1.10 вычисляем средние значения:

$$\overline{Y} = \left(\sum_{j=1}^{\ell} y_j \cdot m_{y_j} \right) \Big/ m; \qquad \overline{X} = \left(\sum_{i=1}^{k} x_i \cdot m_{x_i} \right) \Big/ m \tag{1.18}$$

и средние квадратические отклонения:

$$\sigma_Y^2 = \left(\sum_{j=1}^{\ell} \left(y_j - \overline{Y}\right)^2 \cdot m_{y_j} \right) \Big/ m \ \ \sigma_X^2 = \left(\sum_{i=1}^{k} \left(x_i - \overline{X}\right)^2 \cdot m_{x_i} \right) \Big/ m. \tag{1.19}$$

Замечание 1. Рекомендуется сделать два рисунка – это графические изображения эмпирических законов распределения случайных величин X и Y в виде распределения частот. На рисунках нанести средние значения $\overline{X}$ и $\overline{Y}$.

Уточним определение корреляционной зависимости. Для этого введем понятие **условной средней**. Для каждой i-й строки табл.1.8 (совместно с первой строкой) можно вычислить среднее значение случайной величины Y (по формуле 1.18), которое называется ***условным средним***:

$$\overline{Y}_{x_i} = \overline{Y}\left(X = x_i\right) \quad \left(i = 1, 2, \ldots, k\right).$$

Так как каждому значению x_i соответствует одно значение условной средней, то, очевидно, условная средняя $\overline{Y}_{x_i}$ есть функция от X. В этом случае говорят, что случайная величина Y зависит от X корреляционно.

Корреляционной зависимостью Y от X называют функциональную зависимость условной средней $\overline{Y}_x$ от X:

$$\overline{Y}_x = f\left(X\right). \tag{1.20}$$

Уравнение (1.20) называется уравнением ***регрессии*** Y на X; функция $f\left(X\right)$ называется ***регрессией*** Y на X; график функции $f\left(X\right)$ - ***линией регрессии*** Y на X.

Аналогично для каждого $j-$го столбца табл.1.8 (совместно с первым столбцом) можно вычислить среднее значение случайной величины X по формуле (1.18), которое называется ***условным средним***:

$$\overline{X}_{y_j} = \overline{X}\left(Y = y_j\right) \quad \left(j = 1, 2, \ldots, \ell\right).$$

Тогда корреляционной зависимостью X от Y называется функциональная зависимость $\overline{X}_y$ от Y:

$$\overline{X}_y = \varphi\left(Y\right). \tag{1.21}$$

Уравнение (1.21) называется ***уравнением регрессии*** X на Y; функция $\varphi(Y)$ называется ***регрессией*** X на Y; график функции $\varphi(Y)$ - ***линией регрессии*** X на Y.

Замечание 2. Рассматриваемые два уравнения регрессии существенно различны и не могут быть получены одно из другого.

Изучение корреляционной связи будем проводить при решении двух основных задач:

- определение ***формы*** корреляционной связи, то есть вида теоретической функции регрессии (она может быть линейной и нелинейной);
- определение ***тесноты*** (силы) корреляционной связи.

Наиболее простой и важный случай корреляционной зависимости - линейная регрессия. В этом случае теоретическое уравнение линейной регрессии Y на X (формула 1.20) имеет вид

$$\overline{Y}_x = aX + b. \tag{1.22}$$

Коэффициент a в уравнении (1.22) называют ***коэффициентом регрессии*** Y на X и обозначают ρ_{YX} $(a = \rho_{YX})$. Оценки неизвестных параметров ρ_{YX} и b рассчитаем, применяя данные табл.1.8:

$$a = \rho_{YX} = \frac{\left(\sum_{i=1}^{k}\sum_{j=1}^{\ell} m_{ij}\cdot x_i \cdot y_j\right)\Big/ m - \overline{X}\cdot\overline{Y}}{\left(\sum_{i=1}^{k} x_i^2 \cdot m_{x_i}\right)\Big/ m - \left(\overline{X}\right)^2}. \tag{1.23}$$

$$b = \overline{Y} - \rho_{YX}\cdot\overline{X}, \tag{1.24}$$

где $\overline{X}$ и $\overline{Y}$ - средние значения случайных величин $\overline{X}$ и $\overline{Y}$, вычисленные по формулам (1.18).

Сделаем графическое изображение так называемой эмпирической линии регрессии Y на X и теоретической линии регрессии Y на X. Для этого в декартовой системе координат по оси OX откладываем значения $x_1, x_2, ..., x_k$ из табл. 1.8, по оси OY откладываем значения условных средних $\overline{Y}_{x_i}$. Тогда ломаная, соединяющая точки $\left(x_1, \overline{Y}_{x_1}\right)$; $\left(x_2, \overline{Y}_{x_2}\right)$; … ; $\left(x_k, \overline{Y}_{x_k}\right)$, и будет эмпирической линией регрессии Y на X. Здесь же на данном графике строим теоретическую линию регрессии, то есть прямую $\overline{Y}_x = \rho_{YX} X + b$ с вычисленными коэффициентами.

Замечание 3. Поскольку формулы (1.23) и (1.24) получены по методу наименьших квадратов, то по сути этого метода, теоретическая линия регрессии должна на графике быть в «середине» ломаной.

Аналогично можно поставить вопрос о нахождении теоретического уравнения линейной регрессии X на Y (формула 1.21), которое в этом случае имеет вид

$$\overline{X}_y = a_1 Y + b_1. \tag{1.25}$$

Коэффициент a_1 в уравнении (1.25) называют ***коэффициентом регрессии*** X на Y и обозначают ρ_{XY} $(a_1 = \rho_{XY})$. Оценки неизвестных параметров ρ_{XY} и b_1 рассчитываются по данным табл.1.8:

$$a_1 = \rho_{XY} = \frac{\left(\sum_{i=1}^{k}\sum_{j=1}^{\ell} m_{ij} \cdot x_i \cdot y_j\right) \Big/ m - \overline{X} \cdot \overline{Y}}{\left(\sum_{j=1}^{\ell} y_j^2 \cdot m_{y_j}\right) \Big/ m - \left(\overline{Y}\right)^2}. \tag{1.26}$$

$$b_1 = \overline{X} - \rho_{XY} \cdot \overline{Y}, \tag{1.27}$$

где $\overline{X}$ и $\overline{Y}$ средние значения случайных величин X и Y, вычисленные по формулам (1.18).

Далее целесообразно сделать графическое изображение эмпирической и теоретической линий регрессии X на Y аналогично вышеизложенному.

В случае линейной регрессии задача определения тесноты связи сводится к вычислению эмпирического (выборочного) коэффициента корреляции, который можно вычислить по одной из формул:

$$r_B = \rho_{YX} \cdot \frac{\sigma_X}{\sigma_Y} \quad \text{или} \quad r_B = \rho_{XY} \cdot \frac{\sigma_Y}{\sigma_X}, \tag{1.28}$$

где σ_X, σ_Y - значения средних квадратических отклонений, вычисленных по формуле (1.19).

Приведем свойства выборочного коэффициента корреляции:

1. $|r_B| \le 1$ или $-1 \le r_B \le 1$.

2. Если $r_B = 0$, тогда случайные величины X и Y не связаны линейной корреляционной зависимостью (но могут быть связаны нелинейной корреляционной или даже функциональной зависимостью).

3. С возрастанием абсолютной величины выборочного коэффициента корреляции линейная корреляционная зависимость становится более тесной и при $|r_B| = 1$ переходит в линейную функциональную зависимость.

4. Если $r_B = 1$ $(r_B = -1)$, тогда X и Y связаны прямой (обратной) линейной функциональной зависимостью.

Замечание 4. Однако эмпирический коэффициент корреляции является весьма условным показателем даже линейной связи, так как он является средней пропорциональной величиной между коэффициентами регрессии. В теории корреляции существует понятие корреляционного отношения, которое является более естественным и общим показателем степени тесноты связи, так как не связано с формой зависимости.

2. Методические указания для студентов

2.1. ЛАБОРАТОРНАЯ РАБОТА № 1
«Первичная обработка статистических данных»

Из данных, входящих в выборку $X=\{x_1, x_2, \ldots, x_n\}$ (см. табл.1.1), находим x_{min} и x_{max}, соответственно наименьшее и наибольшее значения выборки, и вычисляем число $d = x_{max} - x_{min}$, называемое ***размахом выборки***. Размах выборки – это длина основного интервала, в который попадают все значения выборки. Далее значения x_i $\left(i=\overline{1,n}\right)$ можно упорядочить, то есть расположить в порядке возрастания. Тогда выборка $X=\{x_1, x_2, \ldots, x_n\}$, записанная по возрастанию, называется ***вариационным рядом*** и её значения x_i - вариантами. По формуле $k = 1 + 4\cdot[\lg n]$, где n- объем выборки X, $[\lg n]$– целая часть числа $\lg n$, определим число k. Данное число k задает количество подынтервалов, на которые разбиваем основной интервал $[x_{min}; x_{max}]$. Вычисляем длину подынтервалов по формуле

$$h = \frac{d}{k} \tag{2.1}$$

и затем – границы подынтервалов:

$$a_0 = x_{min};\ \ a_1 = a_0 + h; \ldots; a_j = a_{j-1} + h; \ldots; a_k = x_{max}. \tag{2.2}$$

Находим m_j $\left(j=\overline{1,k}\right)$– частоты и $\mu_j = \dfrac{m_j}{n}\left(j=\overline{1,k}\right)$ относительные частоты попадания значений выборки X в $j-$й подынтервал. Причем должно быть $\sum\limits_{j=1}^{k} m_j = m_1 + m_2 + \ldots + m_k = n$; для относительных частот: $\sum\limits_{j=1}^{k} \mu_j = 1$.

В результате проведенных расчетов получаем две таблицы:

Таблица 2.1

X	$[a_0; a_1)$	$[a_1; a_2)$	$\ldots$	$[a_{k-1}; a_k]$
m	m_1	m_2	$\ldots$	m_k

Таблица 2.2

X	$[a_0; a_1)$	$[a_1; a_2)$	$\ldots$	$[a_{k-1}; a_k]$
μ	μ_1	μ_2	$\ldots$	μ_k

Далее, если найти середины подынтервалов:

$$b_1 = \frac{a_1 + a_0}{2};\ b_2 = \frac{a_2 + a_1}{2};\ldots;b_k = \frac{a_k + a_{k-1}}{2}, \qquad (2.3)$$

Таблица 2.3

X	b_1	b_2	$\ldots$	b_k
μ	μ_1	μ_2	$\ldots$	μ_k

В целях наглядности полученных в табл. 2.1, 2.2, 2.3 данных пользуются различными способами их графического изображения. К ним относятся гистограмма и полигон.

Для построения гистограммы относительных частот используем данные табл.2.2. В декартовой системе координат на оси $0X$ находим значения x_{min} и x_{max} и тем самым находим границы основного интервала, в который попадают все значения выборки. Затем на этом интервале откладываем границы подынтервалов. По оси $0Y$ откладываем значения относительных частот μ_j $\left(j = \overline{1, k}\right)$. Тогда гистограммой относительных частот назовем ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные подынтервалы длины h, а высоты равны числам μ_j $\left(j = \overline{1, k}\right)$. Аналогично, по данным табл. 2.1, строится гистограмма частот.

Для построения полигона относительных частот используем данные табл. 2.3. В декартовой системе координат на оси $0X$ находим x_{min} и x_{max}, то есть изображаем границы основного интервала. Затем наносим значения середин подынтервалов b_j. По оси $0Y$ откладываем значения, соответствующие относительным частотам μ_j.

Полигоном относительных частот называют ломаную, отрезки которой соединяют точки $(b_1, \mu_1); (b_2, \mu_2);\ldots;(b_k, \mu_k)$.

Данные табл. 2.3 представляют эмпирический закон распределения выборки, а полигон относительных частот есть его визуальное представление.

Эмпирической функцией распределения (функцией распределения выборки) называют функцию $F^*(x)$, определяющую для каждого значения x относительную частоту события $X < x$.

Таким образом, $F^*(x) = \frac{n_x}{n}$, где n_x – число вариант, меньших x, n – объем выборки.

Для каждой реализации выборки эмпирическая функция распределения однозначно определена и обладает всеми свойствами теоретической функции распределения:

1) значения эмпирической функции принадлежат отрезку $[0;1]$;

2) $F^*(x)$ – не убывающая функция;

3) если $x_1 -$ наименьшая варианта, то $F^*(x)=0$ при $x \le x_1$; если $x_k -$наибольшая варианта, то $F^*(x)=1$ при $x > x_k$.

Эмпирическая функция распределения выборки является оценкой теоретической функции распределения генеральной совокупности.

ПРИМЕР 2.1

Дана выборка X из генеральной совокупности объема n=100.

Таблица 2.4

254	1158	522	524	972	736	401	347	208	368
1485	812	1032	226	428	368	676	671	587	701
701	1171	443	683	786	895	267	597	51	941
659	400	484	876	570	241	678	127	728	903
424	245	531	986	1017	429	732	1021	430	153
513	520	221	1074	826	65	389	1180	504	325
294	447	1459	589	307	461	1434	559	837	743
382	387	967	446	763	767	349	853	578	652
285	628	688	517	380	375	878	409	109	621
712	476	432	721	1300	577	580	909	690	757

1. Находим из выборки x_{min} и x_{max}, рассчитываем размах выборки d:

$$x_{max} = 1485; \quad x_{min} = 51; \quad d = x_{max} - x_{min} = 1485 - 51 = 1434.$$

2. Составим вариационный ряд, для чего всю последовательность выборки расположим в порядке возрастания

51	65	109	127	153	208	221	226	241	245
254	267	285	294	307	325	347	349	368	368
375	380	382	387	389	400	401	409	424	428
429	430	432	443	446	447	461	476	484	504
513	517	520	522	524	531	559	570	577	578
580	587	589	597	621	628	652	659	671	676
678	683	688	690	701	701	712	721	728	732
736	743	757	763	767	786	812	826	837	853
876	878	895	903	909	941	967	972	986	1017
1021	1032	1074	1158	1171	1180	1300	1434	1459	1485

3. Вычисляем число $k -$ количество частичных подынтервалов, на которое разбиваем нашу выборку X: $k = 1 + 4 \cdot \lg 100 = 9$. Исходя из этого вычисляем длину подынтервалов $h = \dfrac{1434}{9} \approx 159{,}333$ и границы подынтервалов a_j, $j = \overline{1,9}$: $a_0 = x_{min} = 51;$

$$a_1 = a_0 + h = 51 + 159{,}333 = 210{,}333;$$

$$a_2 = a_1 + h = 210{,}333 + 159{,}333 = 369{,}666;$$

$a_3 = a_2 + h = 369{,}666 + 159{,}333 = 528{,}999$;

$a_4 = a_3 + h = 528{,}999 + 159{,}333 = 688{,}332$;

$a_5 = a_4 + h = 688{,}332 + 159{,}333 = 847{,}665$;

$a_6 = a_5 + h = 847{,}665 + 159{,}333 = 1006{,}998$;

$a_7 = a_6 + h = 1006{,}998 + 159{,}333 = 1166{,}331$;

$a_8 = a_7 + h = 1166{,}331 + 159{,}333 = 1325{,}664$;

$a_9 = a_8 + h = x_{max} = 1485$.

4. Рассчитываем частоты – число попаданий в подынтервалы значений из выборки, то есть $m_1, m_2, \ldots, m_9$.

$[51;\ 210{,}333)$ $m_1 = 5$

$[210{,}333;\ 369{,}666)$ $m_2 = 14$

$[369{,}666;\ 528{,}999)$ $m_3 = 25$

$[528{,}999;\ 688{,}332)$ $m_4 = 18$

$[688{,}332;\ 847{,}665)$ $m_5 = 16$

$[847{,}665;\ 1006{,}998)$ $m_6 = 10$

$[1006{,}998;\ 1166{,}331)$ $m_7 = 5$

$[1166{,}331;\ 1325{,}664)$ $m_8 = 3$

$[1325{,}664; 1485]$ $m_9 = 3$.

Контроль: $\sum_{j=1}^{9} m_j = n$ $\quad 5 + 14 + 25 + 18 + 16 + 10 + 5 + 3 = 100$.

На основе полученных данных заполняем таблицу:

X	[51; 210,333)	[210,333; 369,666)	[369,666; 528,999)	[528,999; 688,332)	[688,332; 847,665)
m	6	14	25	18	16

X	[847,665; 1006,998)	[1006,998; 1166,331)	[1166,331; 1325,664)	[1325,664; 1485]
m	10	5	3	3

5. Считаем середины подынтервалов $b_1, b_2, \cdots, b_9$ и относительные частоты $\mu_1, \mu_2, \ldots, \mu_9$.

$$b_j = \frac{a_{j-1} + a_j}{2}; \quad \mu_j = \frac{m_j}{N}, \quad j = \overline{1,9};$$

$$b_1 = \frac{a_0 + a_1}{2} = \frac{51 + 210{,}333}{2} = 130{,}666;$$

$$b_2 = \frac{a_1 + a_2}{2} = \frac{210{,}333 + 369{,}666}{2} = 290;$$

$$b_3 = \frac{a_2 + a_3}{2} = \frac{369{,}666 + 528{,}999}{2} = 449{,}3;$$

$$b_4 = \frac{a_3 + a_4}{2} = \frac{528{,}999 + 688{,}332}{2} = 608{,}6;$$

$$b_5 = \frac{a_4 + a_5}{2} = \frac{688{,}332 + 847{,}665}{2} = 768;$$

$$b_6 = \frac{a_5 + a_6}{2} = \frac{847{,}665 + 1006{,}998}{2} = 927{,}3;$$

$$b_7 = \frac{a_6 + a_7}{2} = \frac{1006{,}998 + 1166{,}331}{2} = 1086{,}66;$$

$$b_8 = \frac{a_7 + a_8}{2} = \frac{1166{,}331 + 1325{,}664}{2} = 1246;$$

$$b_9 = \frac{a_8 + a_9}{2} = \frac{1325{,}664 + 1485}{2} = 1405{,}3.$$

Относительные частоты:

$$\mu_1 = \frac{m_1}{n} = \frac{6}{100} = 0{,}06; \qquad \mu_6 = \frac{m_6}{n} = \frac{10}{100} = 0{,}1;$$

$$\mu_2 = \frac{m_2}{n} = \frac{14}{100} = 0{,}14; \qquad \mu_7 = \frac{m_7}{n} = \frac{5}{100} = 0{,}05;$$

$$\mu_3 = \frac{m_3}{n} = \frac{25}{100} = 0{,}25; \qquad \mu_8 = \frac{m_8}{n} = \frac{3}{100} = 0{,}03;$$

$$\mu_4 = \frac{m_4}{n} = \frac{18}{100} = 0{,}18; \qquad \mu_9 = \frac{m_9}{n} = \frac{3}{100} = 0{,}03.$$

$$\mu_5 = \frac{m_5}{n} = \frac{16}{100} = 0{,}16;$$

Контроль:

$$\sum_{j=1}^{9} \mu_j = 1 \qquad 0{,}06 + 0{,}14 + 0{,}25 + 0{,}18 + 0{,}16 + 0{,}1 + 0{,}05 + 0{,}03 + 0{,}03 = 1.$$

В результате имеем таблицу:

Таблица 2.5

X	130,6	290	449,3	608,6	768	927,3	1086,6	1246	1405
μ_i	0,06	0,14	0,25	0,18	0,16	0,1	0,05	0,03	0,03

6. Из данных табл. 2.5 получим эмпирический закон распределения относительных частот и визуальное его представление, то есть строим гистограмму (рис. 2.1) и полигон распределения относительных частот (рис. 2.2).

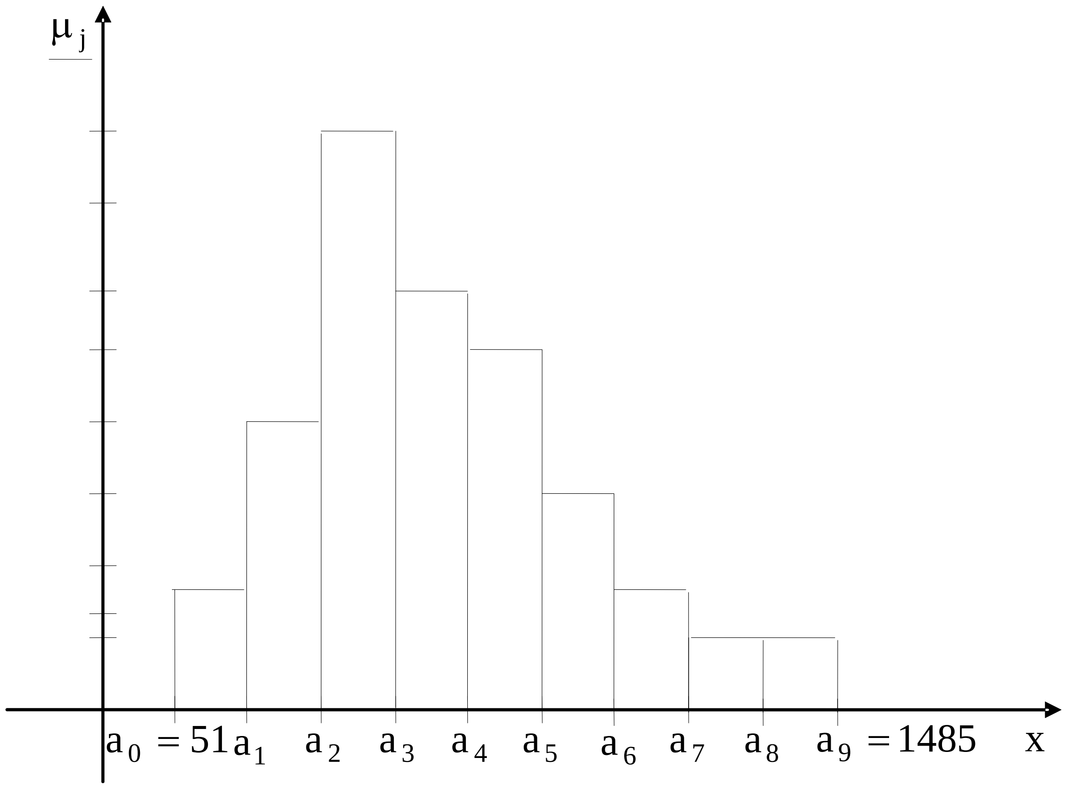

Рис. 2.1

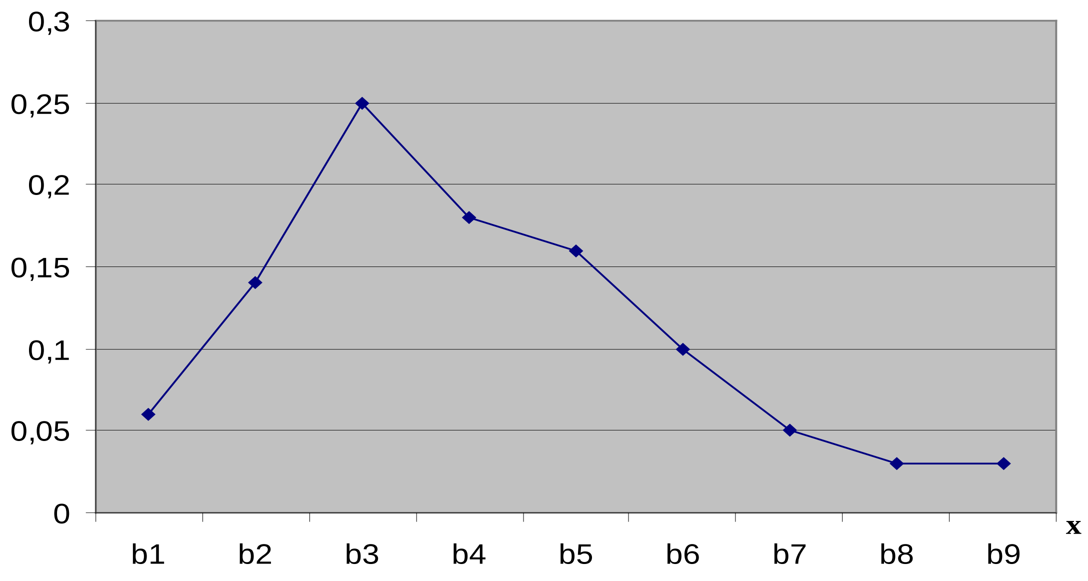

Рис. 2.2

7. Составим эмпирическую функцию распределения. По определению, для группированного статистического ряда $F^*(x)$ имеет вид

$$F^*(x)=\begin{cases} 0 & x \le 51 \\ 0{,}06 & \text{при } 51 < x \le 210{,}333 \\ 0{,}2 & \text{при } 210{,}333 < x \le 369{,}666 \\ 0{,}45 & \text{при } 369{,}666 < x \le 528{,}999 \\ 0{,}63 & \text{при } 528{,}999 < x \le 688{,}332 \\ 0{,}79 & \text{при } 688{,}332 < x \le 847{,}665 \\ 0{,}89 & \text{при } 847{,}665 < x \le 1006{,}998 \\ 0{,}94 & \text{при } 1006{,}998 < x \le 1166{,}331 \\ 0{,}97 & \text{при } 1166{,}331 < x \le 1325{,}664 \\ 0{,}99 & \text{при } 1325{,}664 < x \le 1485 \\ 1 & \text{при } x > 1485 \end{cases}$$

Построим график $F^*(x)$.

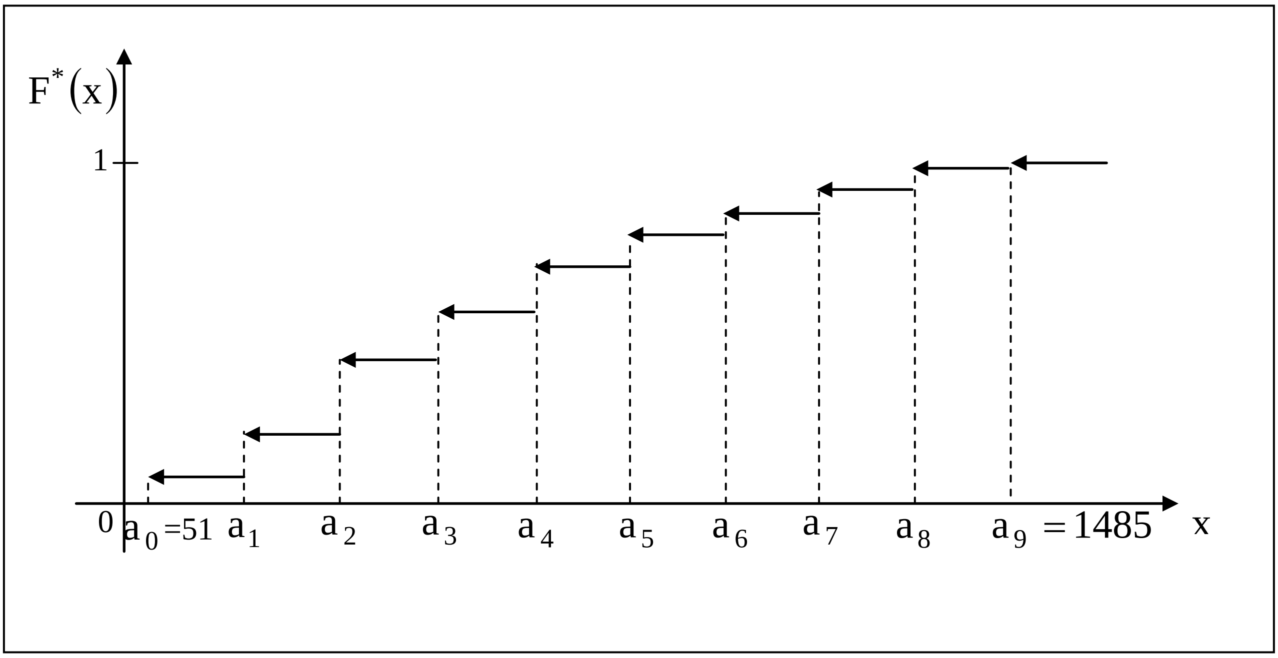

Рис.2.3

2.2. ЛАБОРАТОРНАЯ РАБОТА № 2

«Расчет точечных и интервальных оценок генерального математического ожидания и дисперсии»

Пусть $\theta_г$ – некоторый параметр генеральной совокупности, который невозможно вычислить. Но знать его значение (хотя бы приближенное, оценочное) надо! Поэтому по выборочным данным производят расчет статистических оценок данного генерального параметра.

Точечной называют статистическую оценку генерального параметра $\theta_г$, которая определяется одним числом $\tilde{\theta}$. Точечная оценка $\tilde{\theta}$ может быть несмещенной и смещенной.

Несмещенной называют такую точечную оценку $\tilde{\theta}$, математическое ожидание которой равно оцениваемому генеральному параметру при любом объеме выборки, то есть

$$M\left[\tilde{\theta}\right] = \theta_г. \tag{2.4}$$

Если равенство (2.4) нарушается, то в этом случае точечная оценка $\theta_в$ называется ***смещенной.***

Несмещенной оценкой генеральной средней (математического ожидания генеральной совокупности) служит выборочная средняя:

$$\overline{x}_в = \sum_{j=1}^{k} b_j \cdot \frac{m_j}{n} = \sum_{j=1}^{k} b_j \cdot \mu_j, \tag{2.5}$$

которую считаем по данным табл. 2.3.

Смещенной оценкой генеральной дисперсии $D_г$ служит выборочная дисперсия:

$$D_в = \sum_{j=1}^{k} \left(b_j - \overline{x}_в\right)^2 \cdot \frac{m_j}{n} = \sum_{j=1}^{k} \left(b_j - \overline{x}_в\right)^2 \cdot \mu_j, \tag{2.6}$$

где b_j; μ_j – из табл.2.3. Иногда более удобно пользоваться другой формулой для вычисления выборочной дисперсии:

$$D_в = \sum_{j=1}^{k} b_j^2 \cdot \mu_j - \left(\overline{x}_в\right)^2. \tag{2.6a}$$

<u>Замечание.</u> Поскольку $D_в$ является смещенной оценкой, то ее «исправляют» следующим образом:

$$s^2 = \frac{n \cdot D_в}{n-1}. \tag{2.7}$$

Полученная оценка s^2 – это несмещенная дисперсия, а s – выборочное среднее квадратическое отклонение.

При выборке малого объема точечная оценка может значительно отличаться от оцениваемого генерального параметра, то есть приводит к грубым ошибкам, поэтому при небольшом объеме выборки следует пользоваться интервальными оценками.

Интервальной называют оценку, которая определяется двумя числами – концами интервала, покрывающего оцениваемый генеральный параметр $\theta_г$.

Доверительным называют интервал, который с заданной надежностью (доверительной вероятностью) γ покрывает оцениваемый генеральный параметр, то есть с которой осуществляется неравенство $\left|\theta_г - \tilde{\theta}\right| < \delta$.

Обычно надежность оценки (доверительная вероятность γ) задается. Причем в качестве γ берут число, близкое к единице (0,95; 0,99; 0,999).

Итак, пусть вероятность того, что $\left|\theta_г - \tilde{\theta}\right| < \delta$ равна γ, то есть

$$P\left(\left|\theta_г - \tilde{\theta}\right| < \delta\right) = \gamma, \tag{2.8}$$

или

$$P\left(\tilde{\theta} - \delta < \theta_г < \tilde{\theta} + \delta\right) = \gamma, \tag{2.8a}$$

тогда интервал $\left(\tilde{\theta} - \delta; \tilde{\theta} + \delta\right)$ и есть доверительный интервал.

Для оценки математического ожидания $M_г(x)$ нормально распределенной генеральной совокупности X по выборочной средней $\overline{x}_в$ при известном среднем квадратическом отклонении $\sigma_г = \sqrt{D_г}$ служит доверительный интервал.

$$\overline{x}_в - t\frac{\sigma_г}{\sqrt{n}} < M_г(x) < \overline{x}_в + t\cdot\frac{\sigma_г}{\sqrt{n}}, \tag{2.9}$$

где $t\frac{\sigma_г}{\sqrt{n}} = \delta -$ точность оценки; $n -$ объем выборки; $t -$ это такое значение аргумента функции Лапласа $\Phi(t)$ (прил. 1), при котором $\Phi(t) = \frac{\gamma}{2}$.

Для оценки математического ожидания нормально распределенной генеральной совокупности X по выборочной средней $\overline{x}_в$ при неизвестном среднем квадратическом отклонении $\sigma_г = \sqrt{D_г}$ (при объеме выборки $N > 30$) служит доверительный интервал

$$\overline{x}_в - t_\gamma\cdot\frac{s}{\sqrt{n}} < M_г(x) < \overline{x}_в + t_\gamma\cdot\frac{s}{\sqrt{n}}, \tag{2.10}$$

где t_γ находим по таблице (приложение 2) по заданным n и γ.

Для оценки среднего квадратического отклонения $\sigma_г$ нормально распределенной генеральной совокупности X с доверительной вероятностью служат доверительные интервалы:

$$\sigma_в(1-q) < \sigma_г < \sigma_в(1+q) \quad (\text{при} \quad q<1)$$
$$0 < \sigma_г < \sigma_в(1+q) \quad (\text{при} \quad q>1), \qquad (2.11)$$

где $q -$ находим по таблице (приложение 3) при заданных n и γ.

Замечание. Для $M_г(x)$ предлагается построить доверительные интервалы для двух значений вероятности $\gamma = 0{,}95$ и $0{,}99$. Провести анализ, как меняются границы интервалов с увеличением доверительной вероятности.

ПРИМЕР 2.2. Найти точечные и интервальные оценки генерального математического ожидания и генеральной дисперсии, исходя из данных примера 2.1.

1. По данным табл. 2.5 рассчитываем выборочное математическое ожидание и выборочную дисперсию, среднее квадратическое отклонение.

$$\bar{x}_в = 0{,}06 \cdot 130{,}666 + 0{,}14 \cdot 290 + 0{,}25 \cdot 449{,}333 + 0{,}18 \cdot 608{,}666 +$$
$$+ 0{,}16 \cdot 768 + 0{,}1 \cdot 927{,}332 + 0{,}05 \cdot 1086{,}66 + 0{,}03 \cdot 1246 +$$
$$+ 0{,}03 \cdot 1405{,}332 = 619{,}81.$$

$$D_в = (130{,}666 - 619{,}81)^2 \cdot 0{,}06 + (290{,}0 - 619{,}81)^2 \cdot 0{,}14 +$$
$$+ (449{,}333 - 619{,}81)^2 \cdot 0{,}25 + (608{,}666 - 619{,}81)^2 \cdot 0{,}18 +$$
$$+ (768 - 619{,}81)^2 \cdot 0{,}16 + (927{,}332 - 619{,}81)^2 \cdot 0{,}1 +$$
$$+ (1086{,}66 - 619{,}81)^2 \cdot 0{,}05 + (1246 - 619{,}81)^2 \cdot 0{,}03 +$$
$$+ (1405{,}332 - 619{,}81)^2 \cdot 0{,}03 = 91015.$$

$$\sigma_в = \sqrt{D_в} = \sqrt{91015} \approx 301{,}69.$$

По данным табл. 2.4 вычисляем еще одну точечную характеристику - среднее арифметическое значение нашей выборки $\overline{X} = \sum\limits_{i=1}^{n} x_i / n$: $\overline{X} = 615{,}80$.

2. Делаем расчет интервальных оценок, то есть будем строить доверительные интервалы с доверительной вероятностью γ.

а) $\gamma = 0{,}95$; $\sigma_в = 301{,}69$; $\bar{x}_в = 619{,}81$.

Ищем соответствующее значение t_γ по таблице в прил. 1 $t_\gamma = 1{,}96$.

Точность оценки $\delta = \dfrac{1{,}96 \cdot 301{,}69}{\sqrt{100}} \approx 59{,}3$. Тогда

$$p(619{,}81 - 59{,}3 < M_г(x) < 619{,}81 + 59{,}3) = 0{,}95;$$
$$p(560{,}51 < M_г(x) < 679{,}11) = 0{,}95.$$

б) $\gamma = 0{,}99$; $t_\gamma = 2{,}57$.

$$\delta = \frac{2{,}57 \cdot 301{,}69}{\sqrt{100}} = 77{,}53.$$

$$p(619{,}81 - 77{,}53 < m < 619{,}81 + 77{,}53) = 0{,}99;$$

$$p(542{,}28 < M_г(x) < 697{,}34) = 0{,}99.$$

Строим полученные интервалы на полигоне распределения относительных частот.

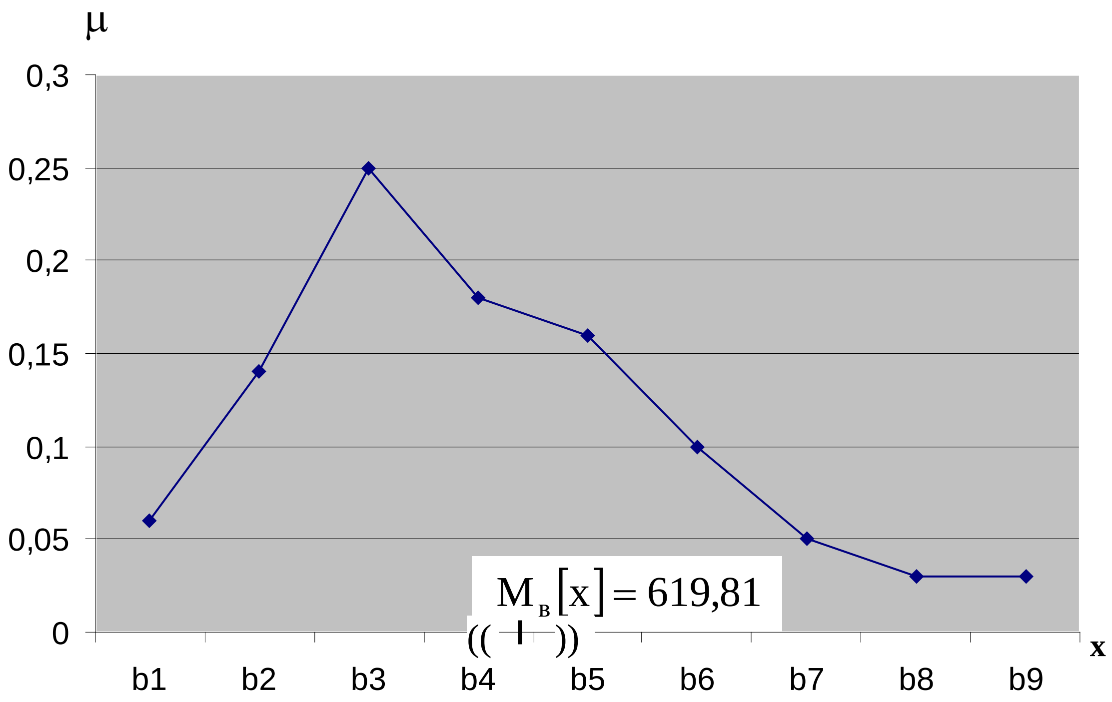

Рис.2.4

2.3. ЛАБОРАТОРНАЯ РАБОТА № 3

«Проверка гипотезы о законе распределения генеральной совокупности»

В лабораторной работе №1 в результате первичной обработки исходных данных получено эмпирическое распределение (табл. 2.3) и по данным этой таблицы построен полигон относительных частот. Относительные частоты иногда называют эмпирическими вероятностями. Из визуального наблюдения полигона можно сделать один из следующих выводов:

Гипотеза A_o**:** Генеральная совокупность распределена по нормальному закону;

Гипотеза B_o**:** Генеральная совокупность распределена по показательному закону;

Гипотеза C_o**:** Генеральная совокупность распределена по равномерному закону.

Гипотеза A_o

Для того чтобы при заданном уровне значимости α ($\gamma -$ доверительная вероятность, то есть вероятность принять верную гипотезу; $\alpha -$ это вероятность отвергнуть верную гипотезу $(\alpha + \gamma = 1)$) проверить гипотезу о нормальном распределении генеральной совокупности, надо:

1. Вычислить $\overline{x}_в$, $\sigma_в$ (лабораторная работа № 2).

2. Вычислить теоретические вероятности p_j $\left(j = \overline{1,\, k}\right)$. Поскольку плотность распределения для нормального закона есть

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-(x - M[x])^2 / (2D[x])}. \tag{2.12}$$

Тогда

$$p_j = p\left(a_{j-1} < x < a_j\right) = \frac{1}{\sqrt{2\pi}\,\sigma_в} \cdot \int\limits_{a_{j-1}}^{a_j} e^{-(x - \overline{x}_в)^2 / (2D_в)} dx =$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_в} \cdot e^{-\left(b_j - \overline{x}_в\right)^2 / (2D_в)} \cdot \int\limits_{a_{j-1}}^{a_j} dx = \frac{1}{\sqrt{2\pi}\,\sigma_в} \cdot e^{-\left(b_j - \overline{x}_в\right)^2 / (2D_в)} \cdot h, \tag{2.13}$$

$$j = \overline{1, k}$$

где a_j – границы частичных интервалов;

b_j – середина $j -$го частичного интервала;

$h -$длина частичного интервала (см. формулу (2.2)).

3. Составить сводную таблицу на основе данных табл. 2.3 и рассчитанных теоретических вероятностей:

Таблица 2.6

X	b_1	b_2	…	b_j	…	b_k	
μ	μ_1	μ_2	…	μ_j	…	μ_k	← эмпирические вероятности
p	p_1	p_2	…	p_j	…	p_k	← теоретические вероятности

4. Оценку отклонения эмпирических вероятностей μ_j $\left(j = \overline{1, k}\right)$ от теоретических вероятностей p_j $\left(j = \overline{1, k}\right)$ произвести с помощью критерия Пирсона χ^2:

$$\chi^2_{набл.} = \sum_{j=1}^{k} \frac{\left(\mu_j - p_j\right)^2 \cdot n}{p_j}. \tag{2.14}$$

5. По таблице критических точек распределения χ^2 (приложение 4) по заданному уровню значимости α и числу степеней свободы $r = k - s - 1$ ($k -$

количество подынтервалов, s - число параметров распределения) найти критическое значение $\chi^2_{кр.}\left(\alpha, r\right)$ правосторонней критической области.

Правило 2.1. Если $\chi^2_{набл.} < \chi^2_{кр.}$, тогда нет оснований отвергать гипотезу A_0 о нормальном распределении генеральной совокупности (то есть эмпирические и теоретические частоты различаются незначимо (случайно).

Правило 2.2. Если $\chi^2_{набл.} > \chi^2_{кр.}$, тогда гипотезу A_0 отвергаем.

Гипотеза B_0

Для того чтобы при заданном уровне значимости α проверить гипотезу о показательном распределении генеральной совокупности, надо:

1. Вычислить $\overline{x}_в$ (лабораторная работа № 1). Принять в качестве оценки параметра λ показательного распределения величину, обратную выборочной средней:

$$\lambda = \frac{1}{\overline{x}_в}. \tag{2.15}$$

2. Вычислить теоретические вероятности $p_j \left(j = \overline{1,k}\right)$. Поскольку плотность распределения для показательного (экспоненциального) закона есть

$$f(x) = \begin{cases} 0, & \text{при } x < 0, \\ \lambda \cdot e^{-\lambda x}, & \text{при } x \geq 0, \end{cases} \tag{2.16}$$

тогда

$$p_j = p\left(a_{j-1} < x < a_j\right) = \lambda \int\limits_{a_{j-1}}^{a_j} e^{-\lambda \cdot x}\, dx =$$

$$= -e^{-\lambda \cdot x}\Big|_{a_{j-1}}^{a_j} = -\left(e^{-\lambda \cdot a_j} - e^{-\lambda \cdot a_{j-1}} \right) = e^{-\lambda \cdot a_{j-1}} - e^{-\lambda \cdot a_j},$$

где $a_j \left(j = 1, \overline{k}\right)$ – границы частичных интервалов;

λ – вычисляем по формуле (2.15).

3. Составить сводную таблицу на основе данных табл. 2.3 и рассчитанных теоретических вероятностей (см. табл. 2.6).

4. Оценку отклонения эмпирических вероятностей $\mu_j \left(j = \overline{1,k}\right)$ от теоретических вероятностей $p_j \left(j = \overline{1,k}\right)$ произвести с помощью критерия Пирсона χ^2 (формула (2.14)).

5. По таблице критических точек распределения χ^2 по заданному уровню значимости α и по числу степеней свободы $r = k - s - 1$ (k – количество подынтервалов, s - число параметров распределения) найти критическое значение $\chi^2_{кр.}\left(\alpha, r\right)$ правосторонней критической области (см. прил.4).

Далее необходимо проанализировать в соответствии с правилами 2.1 и 2.2 (для предыдущей гипотезы).

Гипотеза C_0

Для того чтобы при заданном уровне значимости α проверить гипотезу о равномерном распределении генеральной совокупности, надо:

1. Оценить параметры a и c – концы интервала, в котором наблюдались возможные значения x, по формулам (через a^* и c^* обозначены оценки параметров):

$$a^* = \bar{x}_в - \sqrt{3}\cdot\sigma_в, \quad c^* = \bar{x}_в + \sqrt{3}\cdot\sigma_в. \tag{2.17}$$

2. Вычислить теоретические частоты $p_j \left(j = \overline{1,k}\right)$. Поскольку плотность распределения для равномерного закона есть

$$f(x) = \frac{1}{c^* - a^*}, \tag{2.18}$$

тогда

$$\begin{gathered} p_j = p\left(a_{j-1} < x < a_j\right) = \int\limits_{a_{j-1}}^{a_j} \frac{1}{c^* - a^*}dx = \\ = \frac{1}{c^* - a^*}\cdot\int\limits_{a_{j-1}}^{a_j} dx = \frac{1}{c^* - a^*}\cdot h, \end{gathered} \tag{2.19}$$

где $a_j \left(j = \overline{1,k}\right)$ – границы частичных интервалов;

h – длина частичных интервалов.

Получаем, что все $p_j \left(j = \overline{1,k}\right)$ равны одному числу $\dfrac{h}{c^* - a^*}$.

3. Составить сводную таблицу на основе эмпирических вероятностей и рассчитанных теоретических вероятностей (см. табл. 2.6).

4. Оценку отклонения эмпирических вероятностей $\mu_j \left(j = \overline{1,k}\right)$ от теоретических вероятностей $p_j \left(j = \overline{1,k}\right)$ произвести с помощью критерия Пирсона χ^2 (формула (2.14)).

5. По таблице критических точек распределения χ^2 по заданному уровню значимости α и по числу степеней свободы $r = k - s - 1$ (k – количество подынтервалов, s - число параметров распределений) найти критические значения $\chi^2_{кр.}(\alpha, r)$ правосторонней критической области.

Далее анализируем в соответствии с правилами 2.1 и 2.2 (см. гипотезу А).

<u>*Замечание 1.*</u> После составления табл.5 необходимо сделать на одном рисунке два графика: ломаную эмпирических вероятностей и кривую теоретических вероятностей.

Замечание 2. Здесь же на этом рисунке рекомендуется нанести:

а) $\overline{x}_в$; б) доверительный интервал, построенный для одной из доверительных вероятностей γ (например, для $\lambda = 0{,}95$); в) интервал, построенный по правилу «3-х сигм».

Замечание 3. Данный рисунок является наглядным результатом работы, проделанной в лабораторных работах 1,2,3.

ПРИМЕР 2.3. Используя критерий согласия Пирсона при уровне значимости $\alpha = 0{,}01$, проверить гипотезу о нормальном распределении генеральной совокупности.

1. Переносим из лабораторной работы № 1 полигон (рис.2.2) распределения относительных частот, табл. 2.5, из лабораторной работы №2 (пример 2.2) $\overline{x}_в$ =619,81, $\sigma_в$ =301,69.

X	130,6	290	449,3	608,6	768	927,3	1086,6	1246	1405
μ	0,06	0,14	0,25	0,18	0,16	0,1	0,05	0,03	0,03

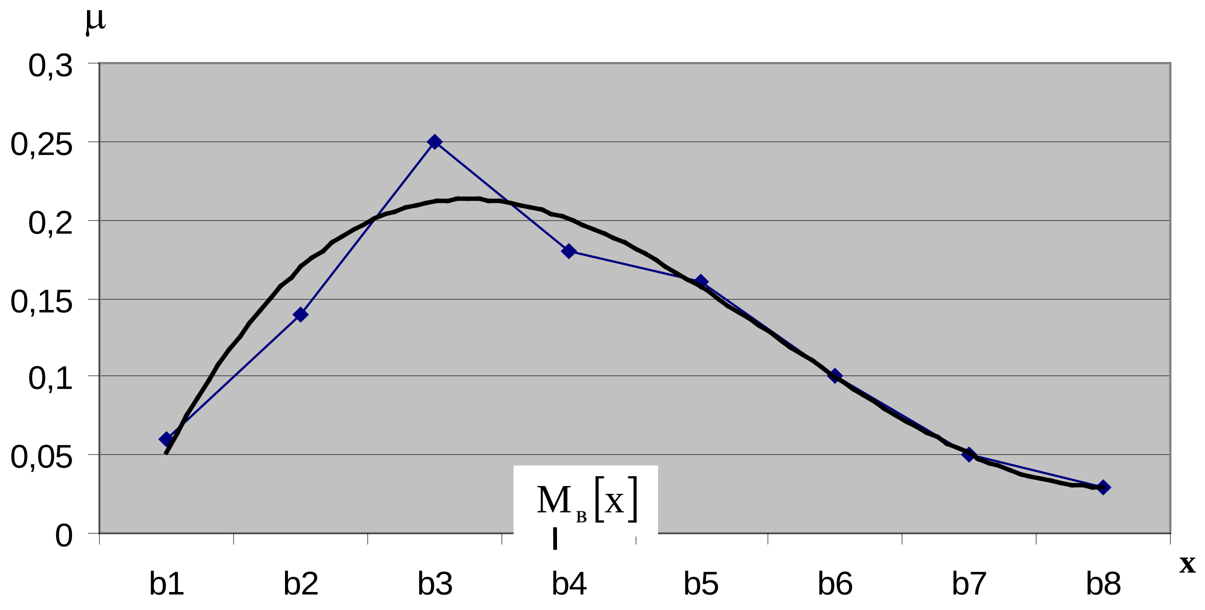

Рис. 2.5

2. Из визуального наблюдения ломаной делаем предположение (ставим гипотезу) о законе распределения генеральной совокупности, то есть ставим гипотезу H_0: выборка распределена по нормальному закону.

3. Вычислим теоретические вероятности p_j. Для этого записываем функцию плотности $f(x)$ для нормального закона:

$$f(x)=\frac{1}{\sqrt{2\pi}\cdot\sigma_в}\cdot e^{-\frac{(x-\bar{x}_в)^2}{2\sigma_в^2}}.$$

Соответственно $f(x)=\dfrac{1}{\sqrt{2\cdot 3{,}14}\cdot 301{,}69}\cdot e^{-\frac{(x-619{,}81)^2}{2\cdot(301{,}69)^2}}$ или

$$f(x)=0{,}001\cdot e^{-\frac{(x-619{,}81)^2}{2\cdot(301{,}69)^2}}.$$

Тогда $p_j=\dfrac{1}{\sqrt{2\pi}\cdot 301{,}69}\cdot e^{-\frac{(b_j-619{,}81)^2}{2\cdot(301{,}69)^2}}\cdot 159{,}333.$

$$p_1=0{,}003\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(130{,}666-619{,}81)^2}{2\cdot(301{,}69)^2}}\cdot 159{,}333=$$

$$=0{,}478\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(1{,}62)^2}{2}}=0{,}478\cdot 0{,}1109=0{,}05;$$

$$p_2=0{,}478\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(290-619{,}81)^2}{2\cdot(301{,}69)^2}}=0{,}478\cdot 0{,}2203=0{,}105;$$

$$p_3=0{,}478\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(449{,}333-619{,}819)^2}{2\cdot 301{,}69^2}}=0{,}478\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{(0{,}565)^2}{2}}=$$

$$=0{,}478\cdot 0{,}3391=0{,}162;$$

$$p_4=0{,}478\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(608{,}666-619{,}819)^2}{2\cdot 301{,}69^2}}=0{,}478\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{(0{,}037)^2}{2}}=$$

$$=0{,}478\cdot 0{,}3986=0{,}19;$$

$$p_5=0{,}478\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(768-619{,}819)^2}{2\cdot 301{,}69^2}}=0{,}169;$$

$$p_6=0{,}478\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(927{,}332-619{,}819)^2}{2\cdot 301{,}69^2}}=0{,}11;$$

$$p_7=0{,}478\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(1086{,}66-619{,}819)^2}{2\cdot 301{,}69^2}}=0{,}06;$$

$$p_8 = 0{,}478 \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(1246-619{,}819)^2}{2 \cdot 301{,}69^2}} = 0{,}02;$$

$$p_9 = 0{,}478 \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(1405{,}332-619{,}819)^2}{2 \cdot 301{,}69^2}} = 0{,}007.$$

4. Составляем табл. 2.7 распределения теоретических вероятностей.

Таблица 2.7

X	130,6	290	449,3	608,6	768	927,3	1086,6	1246	1405
μ	0,06	0,14	0,25	0,18	0,16	0,1	0,05	0,03	0,03
p	0,05	0,105	0,162	0,19	0,17	0,11	0,06	0,02	0,007

Отметим теоретические вероятности на полигоне относительных частот.

5. Рассчитаем значение критерия $\chi^2_{\text{набл.}}$ (см. формулу (2.14)):

$$\chi^2_{\text{набл.}} = \frac{(0{,}06-0{,}05)^2}{0{,}05}\cdot 100 + \frac{(0{,}14-0{,}11)^2}{0{,}11}\cdot 100 + \frac{(0{,}25-0{,}16)^2}{0{,}16}\cdot 100 +$$

$$+\frac{(0{,}18-0{,}19)^2}{0{,}19}\cdot 100 + \frac{(0{,}16-0{,}17)^2}{0{,}17}\cdot 100 + \frac{(0{,}1-0{,}11)^2}{0{,}11}\cdot 100 +$$

$$+\frac{(0{,}03-0{,}03)^2}{0{,}02}\cdot 100 + \frac{(0{,}03-0{,}007)^2}{0{,}007}\cdot 100 = 0{,}2 + 0{,}8 + 5{,}06 + 0{,}05 +$$

$$+0{,}06 + 0{,}09 + 0{,}17 + 0{,}5 + 7{,}56 = 14{,}48.$$

6. Из таблицы «Критические точки распределения χ^2» (прил. 4) находим соответствующее нашим значениям $\chi^2_{\text{кр.}}(\alpha, r)$.

$\alpha = 0{,}01;\ r = k - 3$

$k = 9 \quad \Rightarrow\ r = 6.\ \chi^2_{\text{кр.}}(0{,}01; 6) = 16{,}8.$

Сравниваем $\chi^2_{\text{набл.}}$ и $\chi^2_{\text{кр.}}$. Так как $\chi^2_{\text{набл.}} < \chi^2_{\text{кр.}}$ $(14{,}48 < 16{,}8)$, то гипотеза A_0 (выборка распределена по нормальному закону) принимается по правилу 2.1.

2.4 ЛАБОРАТОРНАЯ РАБОТА № 4
«Расчет параметров корреляционной зависимости. Вывод линейной зависимости».

Две случайные величины могут быть связаны функциональной зависимостью, либо зависимостью другого рода, называемой статистической, либо быть независимыми. Строгая функциональная зависимость для случайных величин реализуется редко, так как обе величины (или одна из них) подвержены различным случайным факторам (даже если среди этих факторов есть общие).

Статистической называют зависимость, при которой изменение одной из величин влечет изменение распределения другой. Если при изменении одной из величин изменяется среднее значение другой, то в этом случае статистическую зависимость называют ***корреляционной.*** Значит, корреляционная зависимость есть частный случай статистической зависимости.

Чтобы установить наличие и характер статистической связи между двумя случайными величинами X и Y, нужно привести к удобному виду исходный цифровой материал. Наглядной (удобной) формой представления данных является корреляционная таблица:

Таблица 2.8

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$	y_ℓ	m_x
x_1	m_{11}	m_{12}	$\cdots$	m_{1j}	$\cdots$	$m_{1\ell}$	m_{x1}
x_2	m_{21}	m_{22}	$\cdots$	m_{2j}	$\cdots$	$m_{2\ell}$	m_{x2}
…	…	…	…	…	…	…	…
x_i	m_{i1}	m_{i2}	$\cdots$	m_{ij}	$\cdots$	$m_{i\ell}$	m_{xi}
…	…	…	…	…	…	…	…
x_k	m_{k1}	m_{k2}	$\cdots$	m_{kj}	$\cdots$	$m_{k\ell}$	m_{xk}
m_y	m_{y1}	m_{y2}	$\cdots$	m_{yj}	$\cdots$	$m_{y\ell}$	m

Здесь $x_1, x_2, \ldots, x_k; y_1, y_2, \ldots, y_\ell -$ середины подинтервалов сгруппированных выборок X и Y (см. лабораторную работу №1); m_{ij} $-$частота, с которой встречается пара $(x_i; y_j)$. В последнем столбце и в последней строке таблицы помещены суммарные частоты, соответствующие значению $X = x_i$ и, соответственно, $Y = y_j$, то есть

$m_{xi} = m_{i1} + m_{i2} + \ldots + m_{i\ell}$;

$m_{yj} = m_{1j} + m_{2j} + \ldots + m_{kj}$, тогда должно быть

$$m_{x_1} + m_{x_2} + \ldots + m_{x_k} = \sum_{i=1}^{k} m_{x_i} = m \text{ и}$$

$$m_{y_1} + m_{y_2} + \ldots + m_{y_j} + \ldots + m_{y_\ell} = \sum_{j=1}^{\ell} m_{y_j} = m;$$

m – общее количество пар значений $(x_i; y_j)$.

Каждая i – я строка табл. 2.8 представляет собой совместно с первой строкой некоторое распределение случайной величины Y, соответствующей данному значению случайной величины $X = x_i$. Такое распределение называется ***условным распределением***. Последняя строка табл. 2.8 совместно с первой строкой образует безусловное распределение случайной величины Y (ее эмпирический закон распределения):

Таблица 2.9

Y	y_1	y_2	$\cdots$	y_j	$\cdots$	y_ℓ
m_y	m_{y_1}	m_{y_2}	$\cdots$	m_{y_j}	$\cdots$	m_{y_ℓ}

Каждый j – й столбец табл. 2.8 представляет собой совместно с первым столбцом некоторое распределение случайной величины X, соответствующее данному значению случайной величины $Y = y_j$ (то есть условное распределение). Последний столбец табл. 2.8 совместно с первым столбцом образует безусловное распределение случайной величины X (ее эмпирический закон распределения):

Таблица 2.10

X	x_1	x_2	$\cdots$	x_i	$\cdots$	x_k
m_x	m_{x_1}	m_{x_2}	$\cdots$	m_{x_i}	$\cdots$	m_{x_k}

По данным табл. 2.9 и. 2.10 вычисляем средние значения:

$$\overline{Y} = \left(\sum_{j=1}^{\ell} y_j \cdot m_{yj} \right) \Big/ m; \quad \overline{X} = \left(\sum_{i=1}^{k} x_i \cdot m_{xi} \right) \Big/ m \tag{2.20}$$

и средние квадратические отклонения:

$$\sigma_Y^2 = \left(\sum_{j=1}^{\ell} \left(y_j - \overline{Y}\right)^2 \cdot m_{yj} \right) \Big/ m; \; \sigma_X^2 = \left(\sum_{i=1}^{k} \left(x_i - \overline{X}\right)^2 \cdot m_{xi} \right) \Big/ m. \tag{2.21}$$

<u>*Замечание.*</u> Рекомендуется сделать два рисунка – это графические изображения эмпирических законов распределения случайных величин X и Y в

виде полигонов распределения частот. На рисунках нанести средние значения $\overline{X}$ и $\overline{Y}$.

Уточним определение корреляционной зависимости. Для этого введем понятие условной средней. Для каждой $i-$й строки табл. 2.8 (совместно с первой строкой) можно вычислить среднее значение случайной величины Y (по формуле 2.20), которое называется ***условным средним***

$$\overline{Y}_{x_i} = \overline{Y}(X = x_i) \quad (i = 1,2,\ldots,k).$$

Так как каждому значению x_i соответствует одно значение условной средней, то очевидно, условная средняя $\overline{Y}_{x_i}$ есть функция от X. В этом случае говорят, что случайная величина Y зависит от X корреляционно.

Корреляционной зависимостью Y от X называют функциональную зависимость условной средней $\overline{Y}_x$ от X:

$$\overline{Y}_x = f(X). \tag{2.22}$$

Уравнение (2.22) называется уравнением ***регрессии*** Y на X; функция $f(x)$ называется ***регрессией*** Y на X; график функции $f(x)$ - ***линией регрессии*** Y на X.

Аналогично, для каждого $j-$го столбца табл. 2.8 (совместно с первым столбцом) можно вычислить среднее значение случайной величины X по формуле (2.20), которое называется ***условным средним***

$$\overline{X}_{y_j} = \overline{X}(Y = y_j) \quad (j = 1,2,\ldots,\ell).$$

Тогда корреляционной зависимостью X от Y называется функциональная зависимость средней $\overline{X}_{y_i}$ от Y:

$$\overline{X}_y = \varphi(Y), \tag{2.23}$$

уравнение (2.23) называется ***уравнением регрессии*** X на Y; функция $\varphi(Y)-$ ***регрессией*** X на Y; график функции $\varphi(Y)-$***линией регрессии*** X на Y.

Замечание. Рассматриваемые два уравнения регрессии существенно различны и не могут быть получены одно из другого.

Изучение корреляционной связи будем проводить при решении двух основных задач:

- определение формы корреляционной связи, то есть вида теоретической функции регрессии (она может быть линейной и нелинейной);
- определение тесноты (силы) корреляционной связи.

Наиболее простой и важный случай корреляционной зависимости – линейная регрессия. В этом случае теоретическое уравнение линейной регрессии Y на X (формула 2.22) имеет вид

$$\overline{Y}_x = a\,X + b. \tag{2.24}$$

Коэффициент a в уравнении (2.24) называют ***коэффициентом регрессии*** X на Y и обозначают $\rho_{YX}\left(a = \rho_{YX}\right)$. Оценки неизвестных параметров ρ_{YX} и b рассчитаем применяя данные табл. 2.8:

$$a = \rho_{YX} = \frac{\left(\sum\limits_{i=1}^{k}\sum\limits_{j=1}^{\ell} m_{ij}\cdot x_i \cdot y_j\right)\Big/ m - \overline{X}\cdot\overline{Y}}{\left(\sum\limits_{i=1}^{k} x_i^2 \cdot m_{x_i}\right)\Big/ m - \left(\overline{X}\right)^2}. \qquad (2.25)$$

$$b = \overline{Y} - \rho_{YX}\cdot\overline{X}, \qquad (2.26)$$

где $\overline{X}$ и $\overline{Y}$ - средние значения случайных величин $\overline{X}$ и $\overline{Y}$, вычисленные по формулам (2.20).

Сделаем графическое изображение так называемой эмпирической линии регрессии Y на X и теоретической линии регрессии Y на X. Для этого в декартовой системе координат по оси $0X$ откладываем значения $x_1, x_2, \ldots, x_k$ из табл. 2.8, по оси $0Y$ откладываем значения условных средних $\overline{Y_{x_i}}$. Тогда ломаная, соединяющая точки $\left(x_1, \overline{Y}_{x_1}\right);\left(x_2, \overline{Y}_{x_2}\right);\ldots$ $\left(x_k, \overline{Y}_{x_k}\right)$, и будет эмпирической линией регрессии Y на X. Здесь же на данном графике строим теоретическую линию регрессии, то есть прямую $\overline{Y}_x = \rho_{YX}\cdot X + b$ с вычисленными коэффициентами.

<u>Замечание .</u> Поскольку формулы (2.25) и (2.26) получены по методу наименьших квадратов, то по сути этого метода теоретическая линия регрессии должна на графике быть в «середине» ломаной линии.

Аналогично можно поставить вопрос о нахождении теоретического уравнения линейной регрессии X на Y (формула 2.23), которое имеет вид

$$\overline{X}_y = a_1 Y + b_1. \qquad (2.27)$$

Коэффициент a_1 в уравнении (2.27) называется ***коэффициентом регрессии*** X на Y и обозначается $\rho_{XY}\left(a_1 = \rho_{XY}\right)$. Оценки неизвестных параметров ρ_{XY} и b_1 рассчитываются по данным табл. 2.7:

$$a_1 = \rho_{XY} = \frac{\left(\sum\limits_{i=1}^{k}\sum\limits_{j=1}^{\ell} m_{ij}\, x_i \cdot y_j\right)\Big/ m - \overline{X}\,\overline{Y}}{\left(\sum\limits_{j=1}^{\ell} y_j^2 \cdot m_{y_j}\right)\Big/ m - \left(\overline{Y}\right)^2}; \qquad (2.28)$$

$$b_1 = \overline{X} - \rho_{YX}\cdot\overline{Y}, \qquad (2.29)$$

где $\overline{X}$ и $\overline{Y}$ – средние значения случайных величин X и Y, вычисленные по формулам (2.20).

Далее целесообразно сделать графическое изображение эмпирической и теоретической линий регрессии X на Y аналогично вышеизложенному.

В случае линейной регрессии задача определения тесноты связи сводится к вычислению эмпирического (выборочного) коэффициента корреляции, который можно вычислить по одной из формул:

$$r_B = \rho_{YX} \cdot \frac{\sigma_x}{\sigma_Y} \quad \text{или} \quad r_B = \rho_{XY} \cdot \frac{\sigma_Y}{\sigma_x}, \qquad (2.30)$$

где σ_X, σ_Y – значения средних квадратических отклонений, вычисленных по формуле (2.21).

Приведем свойства выборочного коэффициента корреляции:

1. $|r_B| \le 1$ или $-1 \le r_B \le 1$.

2. Если $r_B = 0$, тогда X и Y не связаны линейной корреляционной зависимостью (но могут быть связаны нелинейной корреляционной или даже функциональной зависимостью).

3. С возрастанием абсолютной величины выборочного коэффициента корреляции линейная корреляционная зависимость становится более тесной и при $|r_B| = 1$ переходит в функциональную зависимость.

4. Если $r_B = +1$ $(r_B = -1)$, тогда X и Y связаны прямой (обратной) линейной зависимостью.

Замечание. Однако эмпирический коэффициент корреляции является весьма условным показателем даже линейной связи, так как он является средней пропорциональной величиной между коэффициентами регрессии. В теории корреляции существует понятие корреляционного отношения, которое является более естественным и общим показателем степени тесноты связи, так как не связано с формой зависимости. Но в тему данного практикума корреляционные отношения не входят.

ПРИМЕР 2.4. Дана таблица распределения заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т).

X \ Y	6	12	18	24	30	m_x
10	4	-	-	-	-	4
15	2	6	-	-	-	8
20	-	2	5	2	-	9
25	-	-	40	8	4	52
30	-	-	5	7	7	19
35	-	-	-	-	8	8
m_y	6	8	50	17	19	$m = 100$

Находим эмпирические распределения каждой из компонент, их графические изображения, средние значения и средние квадратические отклонения (2.20) – (2.21).

X	10	15	20	25	30	35
$\mu_X = \frac{m_x}{m}$	$\frac{4}{100}$	$\frac{8}{100}$	$\frac{9}{100}$	$\frac{52}{100}$	$\frac{19}{100}$	$\frac{8}{100}$

Y	6	12	18	24	30
$\mu_y = \frac{m_y}{m}$	$\frac{6}{100}$	$\frac{8}{100}$	$\frac{50}{100}$	$\frac{17}{100}$	$\frac{19}{100}$

$$\overline{X} = 10 \cdot \frac{4}{100} + 15 \cdot \frac{8}{100} + 20 \cdot \frac{9}{100} + 25 \cdot \frac{52}{100} + 30 \cdot \frac{19}{100} + 35 \cdot \frac{8}{100} =$$
$$= (10 \cdot 4 + 15 \cdot 8 + 20 \cdot 9 + 25 \cdot 52 + 30 \cdot 19 + 35 \cdot 8)/100 = 24{,}9;$$

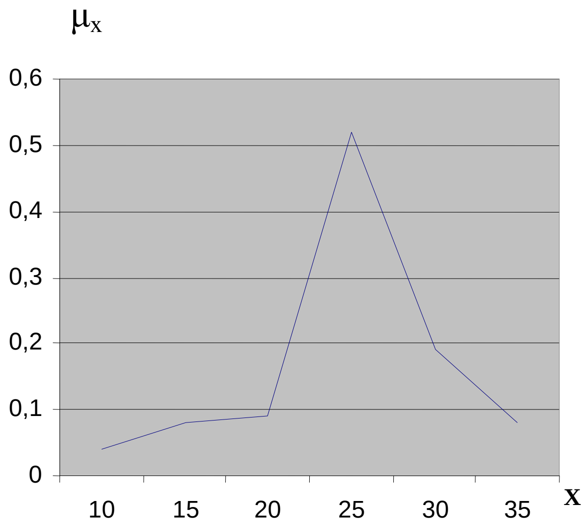

Рис. 2.6

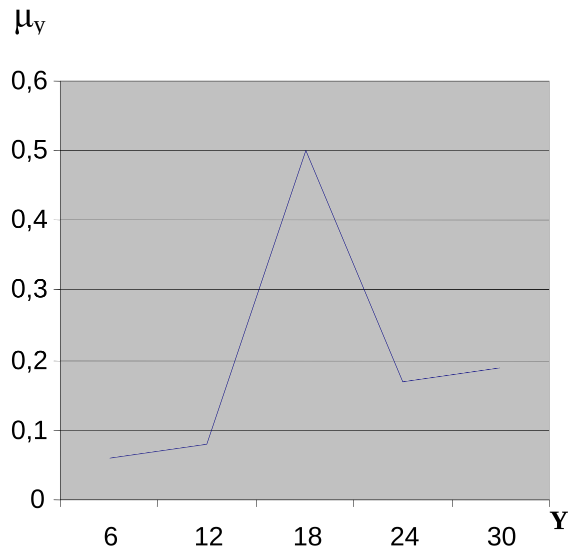

Рис. 2.7

$$\overline{Y} = 6 \cdot \frac{6}{100} + 12 \cdot \frac{8}{100} + 18 \cdot \frac{50}{100} + 24 \cdot \frac{17}{100} + 30 \cdot \frac{19}{100} =$$
$$= (6 \cdot 6 + 12 \cdot 8 + 18 \cdot 50 + 24 \cdot 17 + 30 \cdot 19)/100 = 20{,}1.$$
$$\sigma_X^2 = \left((10 - 24{,}9)^2 \cdot 4 + (15 - 24{,}9)^2 \cdot 8 + (20 - 24{,}9)^2 \cdot 9 + \right.$$
$$\left. + (25 - 24{,}9)^2 \cdot 52 + (30 - 24{,}9)^2 \cdot 19 + (35 - 24{,}9)^2 \cdot 8\right)/100 = 31{,}99$$
$$\sigma_Y^2 = \left((6 - 20{,}1)^2 \cdot 6 + (12 - 20{,}1)^2 \cdot 8 + (18 - 20{,}1)^2 \cdot 5 + \right.$$
$$\left. + (24 - 20{,}1)^2 \cdot 17 + (30 - 20{,}1)^2 \cdot 19\right)/100 = 40{,}6;$$
$$\sigma_X = \sqrt{31{,}99} = 5{,}66; \qquad \sigma_Y = \sqrt{40{,}6} = 6{,}37.$$

Перейдем к расчету данных для построения эмпирических линий регрессии Y на X и X на Y, то есть к расчету условных средних $\overline{Y}(X = x_i) = \left(\sum_{j=1}^{\ell} y_i \, m_{ij}\right) \Big/ m_{x_i}$.

$\overline{Y}(X=10)=6\cdot 4/4=6; \qquad \overline{Y}(X=15)=(6\cdot 2+12\cdot 6)/8=10{,}5;$

$\overline{Y}(X=20)=(12\cdot 2+18\cdot 5+24\cdot 2)/9=18;$

$\overline{Y}(X=25)=(18\cdot 40+24\cdot 8+30\cdot 4)/52=19{,}846;$

$\overline{Y}(X=30)=(18\cdot 5+24\cdot 7+30\cdot 7)/19=24{,}632;$

$\overline{Y}(X=35)=30\cdot 8/8=30,$

и условных средних $\overline{X}(Y=y_j)=\left(\sum_{i=1}^{k} x_i\, m_{i\,j}\right)\Big/ m_{y_j}$

$\overline{X}(Y=6)=(10\cdot 4+15\cdot 2)/6=11{,}7;$

$\overline{X}(Y=12)=(15\cdot 6+20\cdot 2)/8=16{,}25;$

$\overline{X}(Y=18)=(20\cdot 5+25\cdot 40+30\cdot 5)/50=25;$

$\overline{X}(Y=24)=(20\cdot 2+25\cdot 8+30\cdot 7)/17=26{,}5;$

$\overline{X}(Y=30)=(25\cdot 4+30\cdot 7+35\cdot 8)/19=31{,}1.$

Сделаем графическое изображение эмпирических линий регрессии:

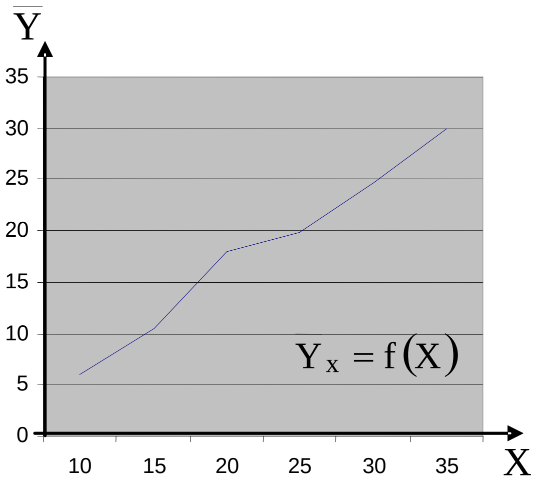

Рис. 2.8

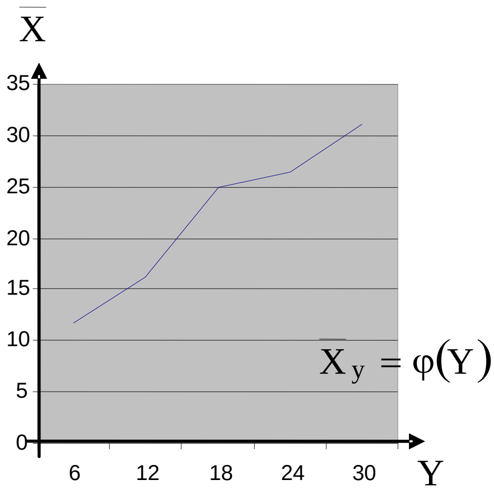

Рис. 2.9

Перейдем к расчету параметров для теоретических линий регрессии $\overline{Y}_x=a\,X+b$ и $\overline{X}_y=a_1 Y+b_1$. Прежде чем воспользоваться формулами (2.25) и (2.28), подготовим их общее значение выражения:

$$C=\sum_{i=1}^{k}\ \sum_{j=1}^{\ell}\left(x_i\, y_j\, m_{i\,j}\right)/m=$$

$$=(10\cdot 6\cdot 4+(15\cdot 6\cdot 2+15\cdot 12\cdot 6)+(20\cdot 12\cdot 2+20\cdot 18\cdot 5+20\cdot 24\cdot 2)+$$
$$+(25\cdot 18\cdot 40+25\cdot 24\cdot 8+25\cdot 30\cdot 4)+(30\cdot 18\cdot 5+30\cdot 24\cdot 7+30\cdot 30\cdot 7)+$$
$$+35\cdot 30\cdot 8)/100=(240+1260+3240+25800+14040+$$
$$+8400)/100=529{,}8.$$

Знаменатели формул (2.25) и (2.28) есть, соответственно, σ_X^2 и σ_Y^2, которые у нас уже рассчитаны, поэтому:

$$a=\rho_{YX}=\frac{529{,}8-24{,}9\cdot 20{,}1}{31{,}99}=\frac{529{,}8-500{,}49}{31{,}99}=\frac{29{,}31}{31{,}99}=0{,}916;$$

$$b = 20{,}1 - 0{,}916 \cdot 24{,}9 = 20{,}1 - 22{,}81 = -2{,}71;$$

$$a_1 = \rho_{XY} = \frac{529{,}8 - 24{,}9 \cdot 20{,}1}{40{,}6} = \frac{529{,}8 - 500{,}49}{40{,}6} = \frac{29{,}31}{40{,}6} = 0{,}722;$$

$$b_1 = 24{,}9 - 0{,}722 \cdot 20{,}1 = 24{,}9 - 14{,}51 = 10{,}39.$$

Тогда $\overline{Y} = 0{,}916\,X - 2{,}71$ и $\overline{X} = 0{,}722\,Y + 10{,}39$. Для графического изображения полученных прямых линий вновь вернемся к рисункам 2.3 и 2.4, чтобы совместить на одном графике эмпирическую и соответствующую ей теоретическую линии регрессии.

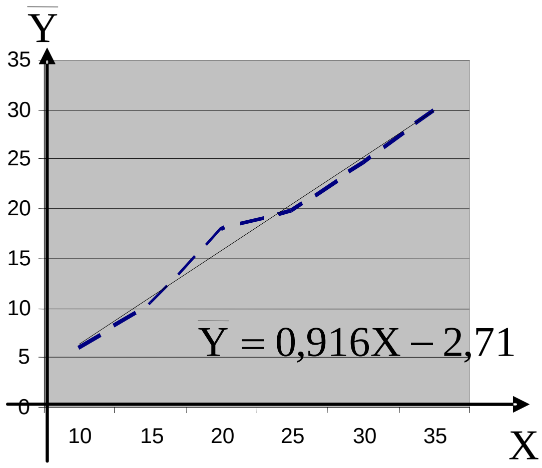

Рис. 2.10

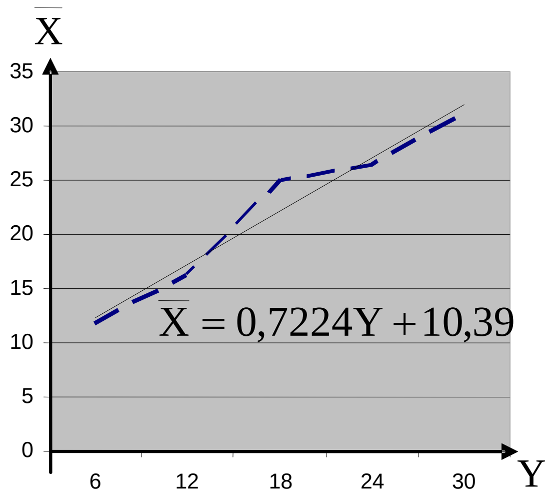

Рис. 2.11

Перейдем к вычислению эмпирического (выборочного) коэффициента корреляции по формуле (2.30):

$$r_B = \rho_{YX} \cdot \frac{\sigma_x}{\sigma_y} = \frac{0{,}916 \cdot 5{,}66}{6{,}37} \approx 0{,}81.$$

Итак: 1. Поскольку r_B достаточно близок к единице, то можно утверждать, что между случайными величинами X и Y существует линейная зависимость, причем достаточно тесная.

2. Поскольку $r_B > 0$, то эта линейная зависимость прямая.

Данные выводы можно сделать и из рис. 2.10, 2.11.

3. МАТЕРИАЛЫ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

3.1. ЛАБОРАТОРНЫЕ РАБОТЫ

ЗАДАЧА 1

Дана выборка из генеральной совокупности объема n. По выборке необходимо выполнить следующие три лабораторные работы.

Лабораторная работа №1. Первичная обработка статистических данных

1. Построить вариационный ряд.
2. Построить группированную выборку с числом интервалов $k = 5 \div 9$.
3. Построить гистограмму и полигон частот и (или) относительных частот.
4. Найти эмпирическую функцию распределения и построить ее график.

Лабораторная работа №2. Расчет точечных и интервальных оценок генерального математического ожидания и дисперсии.

1. По сгруппированной выборке вычислить точечные оценки математического ожидания и среднеквадратического отклонения.
2. Построить доверительные интервалы для генерального математического ожидания с доверительными вероятностями $\gamma_1 = 0{,}95$ и $\gamma_2 = 0{,}99$.

Лабораторная работа №3. Проверка гипотезы о законе распределения генеральной совокупности.

1. Перенести из Л.Р. №1 график полигона относительных частот.
2. Из визуального наблюдения полигона выбрать один из законов распределения (равномерный, нормальный, показательный) в качестве предполагаемого (теоретического) распределения.
3. Найти параметры теоретического распределения.
4. Построить на одном графике полигон относительных частот (выборочное распределение) и кривую теоретического распределения генеральной выборки.
5. Проверить гипотезу о том, что выборка имеет выбранное теоретическое распределение. Принять уровень значимости $\alpha = 0{,}01$.

Вариант 1. $n = 100; \ k = 9$

1.70	1.51	1.35	1.11	1.45	1.94	1.15	1.79	1.35	1.49
1.52	1.97	1.27	1.43	1.10	1.38	1.90	1.07	1.29	1.61
1.29	1.29	1.59	1.32	1.00	1.33	1.29	1.17	1.09	1.50
1.58	1.32	1.70	1.17	1.89	1.97	1.40	1.19	1.24	1.06
1.25	1.41	1.33	1.87	1.13	1.35	1.57	1.48	1.94	1.54
1.03	1.91	1.74	1.56	1.66	1.11	1.56	1.73	1.54	1.10
1.94	1.69	1.78	1.65	1.00	1.37	1.08	1.79	1.72	1.68
1.39	1.45	1.10	1.40	1.90	1.61	1.42	1.63	1.51	1.51
1.82	1.78	1.24	1.78	1.11	1.86	1.73	1.70	1.36	1.92
1.59	1.47	1.29	1.04	1.20	1.90	1.40	1.71	1.04	1.32

Вариант 2. $n = 100; \ k = 7$

0.64	0.39	0.00	0.56	0.05	1.07	0.02	0.01	0.00	0.31
0.06	0.02	0.35	0.13	0.14	0.04	0.42	0.22	0.44	0.08
0.54	1.13	0.47	0.20	0.09	0.25	0.14	0.20	0.01	0.60
0.31	0.36	0.28	0.36	0.91	0.17	0.49	0.08	0.46	0.29
0.63	0.12	0.45	0.07	0.43	0.51	0.84	0.83	0.28	0.37
0.17	0.07	0.04	0.31	0.40	0.12	0.20	0.21	0.60	0.09
0.06	0.24	0.43	0.32	0.16	0.31	0.29	0.60	0.77	0.76
0.22	0.26	0.09	0.18	0.24	0.17	0.10	0.33	0.24	0.44
0.03	0.43	0.08	0.96	0.40	0.22	0.01	0.03	0.94	0.14
0.30	0.27	0.05	0.24	0.03	0.15	0.52	0.16	0.48	0.38

Вариант 3. $n = 96; \ k = 5$

12.57	5.83	7.82	14.30	14.12	8.73	6.91	10.79
10.25	12.86	6.27	11.14	7.90	7.46	6.99	9.69
12.29	6.56	14.16	14.28	6.66	10.91	14.12	12.60
8.60	10.32	8.66	7.04	11.74	7.98	12.30	7.21
6.44	11.20	9.03	9.02	8.14	12.70	11.05	8.56
9.32	13.57	6.73	5.98	13.19	14.63	5.80	5.03
8.32	9.11	7.18	8.02	11.97	14.91	10.16	12.41
8.59	5.58	14.20	12.03	14.69	8.57	6.39	13.10
13.98	11.54	13.76	6.19	8.93	7.73	13.20	14.75
12.98	12.14	5.40	5.10	10.45	11.28	7.76	8.49
6.69	14.85	12.98	11.71	8.89	10.48	8.02	5.83
10.70	13.66	14.72	10.71	6.43	8.53	7.14	9.88

Вариант 4. $n = 96;\ \ k = 7$

1.37	0.42	-0.96	-0.88	1.47	3.14	-1.76	-1.35
1.61	0.02	-0.91	-2.59	1.48	-3.00	2.54	1.51
-1.71	1.60	0.00	-0.73	-2.06	-0.67	0.50	0.14
0.98	-0.14	1.93	-1.09	3.12	2.42	-2.09	-2.44
1.19	1.21	-0.09	1.50	2.71	0.09	-3.05	-0.12
2.30	-1.24	2.74	-3.07	-1.98	-0.86	3.12	-2.47
-0.12	-3.14	-2.91	-1.47	-0.74	-1.04	-1.48	-0.65
3.11	1.72	2.92	0.06	-0.05	-0.20	2.41	0.95
1.25	1.68	3.08	1.11	-1.46	2.29	2.65	-0.59
1.92	-1.60	-1.72	1.05	0.64	-2.51	-2.61	-1.14
1.11	-2.95	0.54	0.12	-2.29	2.61	-2.34	0.95
1.12	2.19	2.36	0.79	2.65	-0.63	1.06	3.12

Вариант 5. $n = 88;\ \ k = 5$

-3.55	-1.52	-6.23	5.94	5.26	-5.79	-4.96	3.87
-2.61	-5.33	-2.73	-5.90	-3.11	2.26	4.94	5.69
2.25	-5.40	1.10	4.60	-0.31	-4.69	-3.87	2.82
-4.75	-5.42	-6.19	5.85	3.78	-6.04	-4.86	3.77
-6.28	-6.02	-0.16	-4.45	-5.15	-5.86	5.44	1.71
-1.83	3.03	4.98	3.98	0.88	0.14	3.11	6.28
-1.65	-5.46	1.72	-1.57	-5.36	-5.93	6.07	0.12
4.96	-6.25	6.07	2.62	5.81	-0.63	4.37	4.64
-5.52	-5.85	1.62	2.89	-6.17	-0.38	-6.21	-4.71
-6.23	-6.23	-3.57	3.59	2.02	3.13	1.65	-6.16
5.22	4.23	-4.62	-4.62	6.04	1.80	-6.28	3.80

Вариант 6. $n = 99;\ \ k = 7$

2.17	2.39	2.80	2.54	1.79	1.77	2.09	1.65	2.91
2.53	1.92	2.82	2.10	2.90	1.64	2.93	2.58	1.72
2.23	2.12	1.70	2.61	2.38	1.88	2.21	2.13	2.09
1.97	3.10	2.54	2.61	1.60	2.61	1.88	3.15	2.42
2.72	3.41	2.47	2.53	3.31	2.74	1.67	2.97	2.46
2.36	3.45	2.66	2.56	3.43	1.80	3.48	2.45	2.65
2.53	2.09	2.45	2.11	1.98	2.19	2.59	3.20	2.33
2.58	2.75	2.77	2.41	2.08	2.98	2.08	3.26	2.64
1.93	1.59	1.91	2.84	3.12	2.29	1.65	2.31	3.32
3.23	3.32	2.17	3.35	1.88	3.30	3.06	2.07	2.16
2.87	2.18	3.25	2.01	1.93	3.35	1.78	3.37	2.32

Вариант 7. $n = 100; \; k = 9$

3.32	5.41	9.86	5.64	5.97	4.74	6.15	3.60	2.20	9.09
11.78	3.69	5.97	5.02	3.25	2.57	4.94	4.86	7.82	6.72
6.67	3.51	7.72	4.44	3.26	8.03	4.15	3.57	2.83	3.01
5.58	7.48	4.20	1.74	2.33	10.88	5.02	5.39	1.54	9.13
2.81	7.54	3.27	1.80	2.21	3.91	4.74	12.32	6.06	5.92
4.16	2.83	8.91	7.96	7.69	1.42	6.12	2.97	0.80	2.35
7.36	4.83	9.56	7.18	10.13	1.13	6.79	3.98	4.63	1.34
2.99	3.71	6.17	5.97	1.72	4.58	1.02	4.65	1.03	9.55
5.42	7.79	8.13	3.32	6.04	6.04	6.58	8.23	3.46	3.22
2.84	3.61	3.78	8.56	4.35	5.74	4.96	5.01	2.10	5.67

Вариант 8. $n = 96; \; k = 7$

-0.99	1.38	0.18	0.47	-0.72	-0.21	1.43	-1.62
-0.82	1.74	-1.41	1.65	0.84	1.28	0.32	-1.43
0.42	1.62	0.83	-0.25	-0.37	-0.54	0.76	1.55
1.02	-0.98	0.24	1.03	-0.14	0.21	2.16	-1.50
-0.96	-0.57	-0.66	-0.39	0.50	-1.14	-1.32	-1.44
1.66	-1.13	-1.10	1.21	0.05	-0.20	-2.02	1.47
1.02	0.31	-0.20	-0.31	0.83	-1.26	-0.15	-0.11
-1.34	0.20	-0.09	-1.16	1.50	2.95	-1.09	1.50
-1.42	-0.51	-0.48	-0.58	-1.10	-0.05	-0.29	-0.66
0.24	-0.83	-0.05	0.21	0.75	0.61	0.10	0.38
0.34	0.58	0.50	0.28	-0.38	0.92	-1.63	-0.09
-0.13	-0.27	-1.68	-0.46	1.01	-0.29	0.39	0.34

Вариант 9. $n = 100; \; k = 9$

0.79	0.03	0.18	2.51	2.88	0.59	0.40	0.08	0.02	0.37
0.54	1.06	0.12	0.17	1.42	3.27	0.29	1.57	1.89	0.70
0.28	0.54	0.74	1.28	0.23	1.01	0.41	0.93	0.66	0.05
0.26	1.02	0.43	0.37	0.90	0.16	2.21	0.69	1.06	3.13
0.52	0.86	2.53	0.59	0.76	0.01	1.36	3.20	2.59	1.12
1.84	1.12	1.32	1.65	0.34	1.14	0.61	0.18	0.36	1.88
1.37	0.35	1.11	1.17	0.04	0.15	0.83	2.75	0.20	0.32
0.05	1.03	0.51	1.36	0.23	0.05	0.40	0.87	0.31	0.40
2.02	0.95	0.55	0.24	2.16	0.20	0.12	1.21	1.49	0.97
0.09	1.49	0.52	1.63	3.33	0.52	0.12	0.14	3.19	0.42

Вариант 10. $n = 99$; $k = 7$

-1.00	0.92	0.06	-0.21	0.21	0.44	-0.14	-0.67	0.44
0.65	-0.33	0.19	-0.12	-0.77	0.15	0.67	-0.99	0.59
0.28	0.24	0.13	-0.37	0.14	-0.09	0.79	-0.64	0.30
-0.83	-0.17	-1.00	0.10	-0.21	-0.23	-0.92	-0.57	0.27
1.00	0.48	-0.97	-0.42	-0.46	-0.81	-0.07	-0.59	1.00
-0.95	0.61	-0.29	-1.00	-0.03	0.39	-0.85	0.45	0.29
0.78	0.17	0.87	-0.96	0.21	-0.48	-0.29	0.07	-0.36
0.08	-1.00	0.98	0.85	0.32	-0.24	0.42	-1.00	0.24
0.88	-0.74	-0.28	0.36	0.46	0.64	0.90	0.01	-0.24
0.36	0.01	-0.45	-0.22	-0.29	-0.77	0.40	-1.00	0.15
-0.16	-0.27	-0.27	-0.10	0.69	0.40	-0.08	-0.81	0.17

Вариант 11. $n = 100$; $k = 7$

0.10	0.11	0.78	0.06	0.42	0.22	0.03	0.20	0.01	0.07
0.23	0.17	0.10	0.27	0.46	1.13	0.34	0.96	0.36	0.46
0.08	0.07	0.05	0.53	0.07	1.16	0.04	0.15	0.27	0.15
0.17	0.04	0.07	0.13	0.02	0.00	0.27	0.18	1.14	0.13
0.04	0.02	0.21	1.05	0.37	0.17	0.05	0.05	0.37	0.12
0.03	0.16	0.01	0.09	0.10	0.02	0.00	0.22	0.23	0.24
0.58	0.01	0.10	0.79	0.34	0.50	0.16	0.44	0.39	0.01
0.11	0.48	0.82	0.11	0.02	0.06	0.01	0.36	0.78	0.00
0.37	0.36	0.70	0.04	0.14	0.02	0.20	0.02	0.19	0.02
0.25	0.05	0.29	0.17	0.30	0.03	0.01	0.14	0.30	0.50

Вариант 12. $n = 100$; $k = 7$

0.52	0.08	0.00	0.03	0.42	0.00	0.29	0.13	0.31	0.63
0.03	0.22	0.06	0.04	0.22	0.57	0.18	0.11	0.21	0.54
0.12	0.00	0.03	0.24	0.30	0.04	0.13	0.35	0.32	0.02
0.10	0.03	0.33	0.09	0.27	0.02	0.29	0.25	0.09	0.15
0.29	0.07	0.11	0.06	0.02	0.07	0.06	0.05	0.21	0.10
0.19	0.45	0.15	0.22	0.38	0.54	0.04	0.03	0.03	0.32
0.12	0.04	0.55	0.01	0.01	0.01	0.02	0.49	0.59	0.11
0.22	0.11	0.43	0.19	0.05	0.08	0.56	0.10	0.25	0.91
0.13	0.21	0.16	0.06	0.11	0.11	0.27	0.18	0.13	0.21
0.45	0.15	0.35	0.03	0.08	0.04	0.36	0.01	0.12	0.08

Вариант 13. $n = 99;\ \ k = 5$

22.39	13.89	8.89	3.14	9.14	11.11	12.01	15.16	10.23
5.87	7.09	4.31	10.83	7.02	12.24	11.35	6.16	11.42
2.87	8.75	7.23	8.91	3.73	14.96	14.46	19.13	11.22
10.57	12.11	19.54	2.64	9.00	8.48	12.44	9.60	7.99
5.96	7.39	3.35	14.25	10.29	13.13	15.60	6.20	14.93
16.27	3.75	6.70	17.96	8.92	19.95	14.72	7.36	5.87
6.58	8.69	6.19	10.82	5.68	12.08	16.21	13.01	7.74
5.63	11.75	5.16	14.70	8.37	18.13	9.97	24.35	19.94
7.40	6.22	8.68	8.68	8.96	9.96	3.39	7.89	11.91
10.97	9.11	3.25	12.53	19.96	13.72	6.06	15.21	4.66
10.22	12.53	16.18	10.51	5.20	6.85	7.80	8.86	10.57

Вариант 14. $n = 100;\ \ k = 9$

3.05	3.20	3.57	3.78	2.23	3.86	2.08	3.06	3.75	2.76
2.93	2.12	3.37	3.32	3.14	2.13	3.67	3.47	3.60	3.62
3.86	2.39	3.08	3.41	3.82	3.16	2.77	2.18	2.84	2.74
2.97	2.03	2.11	3.63	3.25	3.42	2.99	3.99	2.53	2.08
2.62	3.09	2.01	3.05	3.04	2.54	3.97	2.84	3.53	3.32
2.23	3.42	2.13	3.35	3.62	2.50	3.26	2.60	3.31	2.67
3.78	4.00	3.00	3.86	2.02	3.58	2.55	3.26	2.26	3.61
2.75	3.60	3.32	3.95	3.68	2.55	3.80	3.70	2.90	2.75
3.19	2.16	3.81	3.29	2.36	2.84	2.30	3.87	2.25	2.79
2.68	3.31	2.11	2.19	3.33	2.18	2.22	2.60	3.04	2.80

Вариант 15. $n = 100;\ \ k = 7$

1.94	0.93	0.21	0.63	1.08	2.16	1.47	2.44	0.52	0.70
2.50	0.77	0.71	1.85	0.35	0.77	1.73	2.75	0.92	1.63
0.35	1.27	0.88	1.21	0.48	1.22	0.34	0.97	1.92	1.59
1.04	1.09	0.26	1.37	0.32	0.51	1.98	2.39	0.45	1.07
1.14	0.50	2.06	0.83	0.39	0.62	1.51	0.55	1.18	0.84
1.68	0.23	0.53	0.16	1.39	0.36	0.93	0.54	1.89	1.19
0.38	2.61	0.67	1.10	0.56	2.09	0.90	1.06	0.55	1.63
1.39	1.81	2.11	1.34	2.63	1.11	0.48	1.40	3.00	0.79
0.59	1.27	0.43	1.67	0.99	3.27	0.63	1.79	0.81	1.06
1.61	1.42	1.37	1.78	1.02	0.93	0.24	1.99	0.86	1.58

Вариант 16. $n = 100; \ k = 5$

0.57	0.15	0.05	2.58	0.08	0.71	0.17	0.16	0.46	0.12
0.12	0.25	1.62	0.60	0.62	1.12	0.39	0.01	1.08	0.10
0.70	0.61	0.24	0.28	0.15	1.00	1.01	0.53	0.66	0.18
1.00	0.11	0.25	0.05	0.05	0.17	0.38	1.25	0.57	0.10
0.89	0.02	0.66	0.90	0.66	2.44	0.24	0.04	0.18	0.42
0.11	0.58	0.89	1.22	0.46	0.43	0.03	0.50	0.82	0.18
0.13	0.24	0.02	1.45	0.45	0.56	0.19	0.76	0.33	2.29
3.25	0.13	0.95	0.39	0.31	0.83	0.02	2.21	0.15	0.34
1.19	0.21	0.45	0.23	1.84	0.72	0.48	0.51	0.71	0.17
2.22	0.51	0.69	1.11	0.37	0.24	0.34	0.51	0.07	0.51

Вариант 17. $n = 99; \ k = 7$

-2.50	-1.21	-1.28	-2.86	-1.41	-1.80	-1.44	-2.80	-2.75
-2.24	-1.49	-1.78	-2.58	-2.70	-2.68	-2.54	-2.46	-2.66
-2.46	-2.45	-1.49	-1.37	-2.71	-1.30	-2.27	-2.47	-2.27
-2.26	-1.19	-1.53	-1.01	-1.05	-2.05	-2.44	-2.89	-1.58
-1.73	-2.68	-2.87	-1.23	-1.11	-2.82	-2.24	-1.10	-2.47
-1.74	-1.40	-2.36	-1.91	-1.76	-2.45	-2.67	-1.09	-2.05
-2.13	-1.52	-2.93	-2.79	-2.04	-2.64	-2.35	-2.69	-2.72
-1.81	-2.66	-2.12	-2.21	-1.42	-1.12	-1.59	-2.72	-2.55
-2.43	-1.82	-2.59	-2.06	-1.17	-2.16	-2.32	-1.41	-1.64
-2.69	-1.94	-1.81	-2.35	-1.86	-1.49	-1.82	-2.54	-1.52
-1.98	-2.03	-2.64	-2.06	-1.67	-2.85	-1.44	-2.43	-1.58

Вариант 18. $n = 99; \ k = 7$

-2.61	-1.43	-3.37	-1.26	-1.83	-2.76	-3.31	-3.47	-2.69
-1.57	-2.86	-3.55	-3.10	-3.48	-3.62	-2.33	-1.99	-3.43
-1.63	-1.52	-1.02	-1.83	-3.09	-2.68	-2.52	-3.26	-3.62
-1.92	-3.25	-1.70	-2.19	-3.03	-2.16	-2.04	-3.52	-2.45
-2.08	-3.86	-2.59	-1.43	-1.98	-1.72	-3.15	-2.42	-1.80
-1.92	-1.01	-2.24	-1.33	-1.66	-3.24	-1.08	-2.26	-1.67
-1.16	-3.45	-1.43	-2.89	-2.89	-2.15	-3.73	-1.80	-1.99
-3.26	-2.09	-1.84	-1.38	-3.51	-2.27	-2.36	-3.36	-3.66
-3.24	-3.92	-3.86	-1.61	-3.71	-2.87	-2.92	-3.90	-2.80
-3.62	-2.52	-1.45	-3.20	-2.34	-3.94	-2.87	-1.93	-2.36
-3.92	-1.01	-2.32	-2.83	-3.39	-2.32	-3.88	-1.30	-3.05

Вариант 19. $n = 96;\ \ k = 5$

-3.98	0.94	1.99	-3.97	5.97	-5.04	-4.70	5.87
-2.27	-3.27	6.58	3.34	-1.15	0.21	3.05	-2.66
4.26	-4.40	1.66	-0.56	-1.23	-5.62	-8.08	-5.68
2.05	6.97	-0.98	4.09	-5.70	-0.42	-5.25	1.53
0.80	-0.81	3.30	1.13	3.17	-4.36	1.35	-6.72
-3.86	-1.93	-1.56	6.00	-0.53	11.80	0.41	-2.32
-3.34	3.32	4.82	6.02	-1.02	-2.04	0.97	2.01
0.71	-0.20	-0.59	-4.42	-2.17	2.42	-2.46	-4.65
4.06	1.89	-2.31	-0.36	0.84	6.12	5.13	-1.90
-5.62	4.13	0.85	3.00	8.63	1.28	-5.99	-1.17
-6.13	-4.51	5.53	-0.34	-0.81	6.48	-0.19	-1.09
2.65	-1.48	4.04	-2.89	1.55	-5.77	1.38	3.69

Вариант 20. $n = 100;\ \ k = 9$

0.59	0.33	0.41	0.73	0.62	0.14	1.32	0.13	0.24	0.77
0.61	0.82	0.74	0.57	0.68	0.60	0.73	0.79	0.30	0.45
0.46	0.09	0.82	0.74	0.57	0.68	0.69	0.73	0.79	0.35
0.31	0.93	0.92	1.25	0.39	0.32	0.24	0.64	0.47	0.75
1.12	0.69	0.51	0.29	0.23	0.83	0.41	0.86	0.53	0.38
0.51	0.20	0.55	0.76	1.30	0.69	0.67	0.20	0.40	0.67
0.97	0.56	0.81	0.38	0.49	1.07	0.37	0.16	0.78	0.51
0.42	0.17	0.65	1.15	0.11	0.91	0.91	0.46	0.41	0.32
0.63	0.42	0.52	0.34	0.56	0.23	0.81	0.75	0.28	0.72
0.85	0.55	0.37	0.53	0.85	0.40	0.29	0.21	0.82	0.35

Вариант 21. $n = 90;\ \ k = 5$

1.39	3.56	1.42	2.81	2.27	0.67	1.90	0.07	0.59	2.35
2.91	0.75	0.81	1.90	2.26	0.03	5.54	0.75	2.66	0.07
0.61	0.01	0.87	0.49	0.65	0.76	0.65	6.89	2.19	1.37
2.88	1.28	1.82	0.02	1.00	2.52	1.00	2.19	3.78	2.02
0.38	0.37	1.61	3.06	3.17	4.40	0.33	0.58	0.35	4.09
1.32	5.09	1.33	0.24	4.68	0.69	4.93	1.91	2.20	0.27
1.81	6.96	5.80	0.69	1.43	6.17	1.69	0.10	3.17	0.45
1.87	1.41	4.14	3.61	4.15	0.27	0.02	0.06	0.06	0.32
3.88	4.27	2.03	1.12	4.40	3.66	2.94	0.13	4.90	0.94

Вариант 22. $n = 88$; $k = 5$

–1.57	–1.39	–1.57	–0.90	0.15	0.62	–1.51	–1.22
–1.55	–0.10	–0.40	–1.35	1.06	1.23	–1.48	–1.28
–1.34	1.56	–1.53	–1.57	–1.56	–1.47	1.29	–0.38
–0.71	–1.57	–0.33	1.37	–1.00	0.52	0.35	–0.21
–1.56	–1.32	–0.58	–0.33	0.05	–1.44	–0.93	0.40
0.30	1.70	1.05	–0.92	–1.05	–1.30	–0.45	–1.40
–1.18	–0.79	0.85	–1.05	–0.97	–1.50	–1.40	0.40
–1.55	–1.32	–0.71	–0.73	–1.41	–0.07	1.56	1.17
–0.39	–0.15	–0.82	–0.49	–0.39	0.10	1.36	–0.78
–1.11	–1.52	–1.30	0.35	–1.32	1.38	–1.56	–1.49
0.64	–1.51	0.90	–0.47	–1.56	–0.22	–1.45	–0.66

В варианте 22- надо сделать сдвиг на +3 единицы.

Вариант 23. $n = 96$; $k = 7$

-1.21	1.31	-1.60	1.29	-2.59	2.76	-0.85	-0.18
-3.25	1.06	-2.04	0.10	-2.05	2.82	-0.47	1.88
1.22	3.24	1.13	0.87	-1.73	3.40	0.87	0.97
-0.87	1.80	1.65	2.20	-3.19	-1.96	0.59	-0.98
2.25	3.10	0.56	0.75	-1.36	-0.24	1.50	-0.77
-1.79	-0.77	0.75	-1.89	-1.16	-0.18	0.60	2.61
-1.21	-2.30	0.84	-5.61	2.64	-3.94	1.73	-1.33
0.95	1.35	2.49	-0.56	0.05	3.17	1.59	-4.59
4.47	3.81	0.27	0.82	1.24	-0.23	-0.32	-1.33
0.52	-0.01	-0.24	-1.87	-0.09	2.81	3.35	1.68
-1.62	3.89	-0.85	-1.59	1.49	1.66	-2.50	-2.53
0.16	-1.46	-0.19	1.99	0.82	-0.52	-2.62	0.27

Вариант 24. $n = 100$; $k = 9$

2.67	2.41	2.54	3.12	2.02	2.68	2.57	1.47	1.38	2.72
3.85	1.91	2.23	3.33	2.66	1.49	2.75	3.87	2.36	3.26
2.48	1.99	2.94	2.90	3.58	1.13	3.36	3.70	2.71	1.42
1.47	3.05	1.13	1.74	2.67	2.43	1.98	3.14	3.39	2.71
1.05	2.47	1.04	2.46	3.16	2.16	1.74	1.26	1.54	1.12
2.68	3.55	2.23	3.17	2.60	2.18	1.72	3.08	2.46	2.57
2.63	3.59	2.46	3.42	2.62	3.51	3.53	2.72	1.11	3.45
3.85	2.38	3.72	1.01	1.98	3.54	1.64	2.71	1.33	2.28
1.57	1.17	1.02	1.20	3.00	2.93	2.76	3.38	1.71	3.65
3.50	1.49	3.29	2.24	1.88	3.43	3.43	3.20	2.07	2.79

Вариант 25. $n = 99;\ \ k = 7$

0.51	-0.93	-0.68	-0.03	0.06	-1.04	-0.12	-0.62	0.63
-0.92	-0.24	-0.34	-0.21	-0.17	1.88	0.19	1.29	-0.18
1.14	0.32	-0.07	-0.18	0.51	-0.46	0.60	0.09	-0.01
1.06	0.25	-0.67	-0.32	-0.77	-0.09	-0.31	-0.46	-0.10
-0.06	0.13	0.60	-0.39	-1.20	0.88	0.13	0.39	-0.58
-0.17	-0.37	-0.14	-0.06	0.51	-0.34	0.62	0.84	0.10
-0.14	-0.40	-0.84	0.05	0.23	-0.21	-0.40	0.37	0.10
0.47	-0.41	0.52	0.02	0.51	0.03	0.69	0.20	-0.60
-0.07	-0.01	-0.57	-0.76	-0.30	-0.06	0.28	0.07	-0.18
0.47	-0.44	0.42	-0.72	0.11	-0.09	1.28	0.43	0.74
-0.25	-0.39	0.31	0.43	0.65	0.31	0.35	0.46	-0.04

Вариант 26. $n = 100;\ \ k = 7$

2.54	11.58	5.22	5.24	9.72	7.36	4.01	3.47	2.08	3.68
14.85	8.12	10.32	2.26	4.28	3.68	6.76	6.71	5.87	7.01
7.01	11.71	4.43	6.83	7.86	8.95	2.67	5.97	0.51	9.41
6.59	4.00	4.84	8.76	5.70	2.41	6.78	1.27	7.28	9.03
4.24	2.45	5.31	9.86	10.17	4.29	7.32	10.21	4.30	1.53
5.13	5.20	4.32	10.74	8.26	0.65	3.89	15.80	6.90	7.57
2.94	4.47	2.21	5.89	3.07	4.61	5.80	9.09	5.04	3.25
3.82	3.87	14.59	4.46	7.63	7.67	14.34	5.59	8.37	7.43
2.85	6.28	9.67	5.17	3.80	3.75	3.49	8.53	5.78	6.52
7.12	4.76	6.88	7.21	13.00	5.77	8.78	4.09	1.09	6.21

Вариант 27. $n = 99\ \ k = 5$

2.38	1.58	0.02	-0.18	-0.51	2.29	-0.55	2.61	-6.28
-2.02	2.75	-2.70	1.61	-0.49	-4.95	-0.36	1.22	-2.77
0.11	-1.45	6.27	-2.70	-0.29	2.79	6.24	-3.26	1.83
0.53	-0.26	1.09	6.20	-0.28	-0.83	1.17	-1.92	-0.39
-2.56	3.45	-2.19	-4.16	-0.16	-3.17	0.45	-0.13	2.90
-2.03	0.42	-3.94	-0.75	2.98	-3.88	5.02	1.02	-4.76
-0.53	-3.59	-2.52	3.86	-3.30	4.52	-2.86	-0.16	-1.49
0.02	-0.19	-3.92	-2.10	-1.74	2.16	-0.99	-2.74	3.78
-2.66	3.32	4.37	-0.01	4.84	-2.32	2.29	1.18	-5.19
-1.34	1.46	0.55	-0.86	-0.85	-2.69	-1.26	-0.69	-1.33
-4.32	0.41	1.85	-1.50	2.58	4.02	4.00	3.20	4.09

Вариант 28. $n = 100; \ k = 9$

7.24	2.47	3.99	0.83	5.20	1.93	2.44	3.62	2.97	4.24
3.40	2.15	1.56	8.00	3.01	6.85	4.67	3.10	5.66	3.47
5.95	0.67	1.83	4.63	4.17	2.94	3.84	2.02	1.81	4.83
7.08	3.27	3.18	2.63	1.37	2.29	4.57	0.91	2.08	5.18
5.24	2.54	2.46	1.43	5.70	3.14	5.18	2.37	4.59	3.55
1.77	0.42	3.17	7.02	5.74	4.30	2.84	4.69	3.42	1.32
2.34	6.08	5.44	5.37	4.60	7.58	2.98	8.24	0.73	3.58
4.18	0.60	4.15	6.11	2.24	2.62	5.31	2.21	5.34	4.33
4.23	5.72	4.38	1.58	0.64	2.88	4.83	5.14	1.83	5.90
1.25	3.95	2.20	5.11	1.69	4.19	3.23	2.75	2.95	1.99

Вариант 29. $n = 99; \ k = 7$

36.36	0.21	26.80	15.85	2.92	20.63	5.01	3.12	12.19
9.02	34.47	4.81	4.49	7.93	17.43	7.98	9.09	15.46
4.24	3.37	11.78	4.64	0.79	11.21	3.94	13.65	9.51
8.81	9.63	5.02	11.30	11.49	33.06	25.32	6.63	11.69
1.93	2.67	4.55	4.75	8.89	7.99	0.87	23.55	15.03
9.15	11.83	0.02	1.20	13.16	3.35	3.20	9.24	10.06
11.84	3.07	8.06	7.92	0.36	14.30	20.67	3.95	2.39
0.88	1.96	14.26	4.69	25.77	13.70	33.80	9.51	22.20
23.86	10.04	7.10	12.10	2.85	3.12	3.65	2.92	15.77
11.46	7.28	7.45	11.49	6.72	22.15	2.57	8.08	18.04
16.67	2.34	8.67	23.75	11.50	2.70	5.27	20.93	2.58

Вариант 30. $n = 100; \ k = 9$

1.46	1.77	6.06	2.25	0.42	8.41	2.04	2.23	4.00	3.11
5.23	2.92	0.04	1.78	0.57	3.75	1.05	7.19	1.88	0.26
1.89	4.24	5.81	3.78	3.18	0.86	6.66	1.05	0.46	2.18
3.90	6.57	2.29	4.76	1.17	2.95	1.45	1.82	2.60	4.56
4.38	3.39	6.97	0.60	0.56	0.73	1.79	9.44	3.96	0.77
0.09	2.76	8.32	0.35	0.86	2.53	2.25	1.29	0.68	0.02
3.79	1.23	0.16	1.49	0.14	0.31	0.77	0.65	2.74	5.89
4.83	0.63	2.70	5.93	6.25	6.60	1.30	4.60	2.56	0.08
4.91	0.44	0.33	0.24	3.94	3.99	4.42	3.00	2.85	2.01
1.08	3.24	0.55	10.24	2.09	0.74	1.05	0.71	0.08	0.32

Вариант 31. $n = 55$

17	19	23	18	21	15	16	13	20	18	15
20	14	20	16	14	20	19	15	19	16	19
15	22	21	12	10	21	18	14	14	17	16
13	19	18	20	24	16	20	19	17	18	18
21	17	19	17	13	17	11	18	19	19	17

Вариант 32. $n = 55$

14	16	20	15	18	12	13	10	17	15	12
17	11	17	13	11	17	16	12	16	13	16
12	19	18	9	7	18	15	11	11	14	13
10	16	15	17	21	13	17	16	14	15	15
18	14	16	14	10	14	8	15	16	16	13

Вариант 33. $n = 55$

15	17	21	16	19	13	14	11	18	16	13
18	12	18	14	12	18	17	13	17	14	17
13	20	19	10	8	19	16	12	12	15	14
19	15	17	15	11	15	9	16	17	17	15
11	17	16	18	22	14	18	17	15	16	16

Вариант 34. $n = 55$

16	18	22	17	20	14	15	12	19	17	14
19	13	19	15	13	19	18	14	18	15	18
14	21	20	11	9	20	17	13	13	16	15
12	18	17	19	23	15	19	18	16	17	17
20	16	18	16	12	16	10	17	18	19	16

Вариант 35. $n = 55$

18	20	24	19	22	16	17	14	21	19	16
21	15	21	17	15	21	20	16	20	17	20
16	23	22	13	11	22	19	15	15	18	17
14	20	19	21	25	17	21	20	18	19	19
22	18	20	18	14	18	12	19	20	20	18

Вариант 36. $n = 55$

20	22	26	21	24	18	19	16	23	21	18
23	17	23	19	17	23	18	22	19	22	18
18	25	24	15	13	24	21	17	17	20	19
16	22	21	23	27	19	23	22	20	21	21
24	20	22	20	16	20	14	21	22	22	20

Вариант 37. $n = 55$

27	19	33	28	31	25	26	23	30	28	25
30	24	30	26	24	30	29	25	29	26	29
25	32	31	22	20	31	28	24	24	27	26
23	29	28	30	34	26	30	29	27	28	28
31	27	29	27	23	27	21	28	29	29	27

Вариант 38. $n = 55$

19	21	25	20	23	17	18	15	22	20	17
22	16	22	18	16	22	21	17	21	18	21
17	24	23	14	12	23	20	16	16	19	18
15	21	20	22	26	18	22	21	19	20	20
23	19	21	19	15	19	13	20	21	21	19

Вариант 39. $n = 55$

26	28	32	27	30	24	25	22	29	27	24
29	23	29	25	23	29	28	24	28	25	28
24	31	30	21	19	30	27	23	23	26	25
22	28	27	29	33	25	29	28	26	27	27
30	26	28	26	22	26	20	27	28	28	26

Вариант 40. $n = 55$

21	23	27	22	25	19	20	17	24	22	19
24	18	24	20	18	24	23	19	23	20	23
19	26	25	16	14	25	22	18	18	21	20
17	23	22	24	28	20	24	23	21	22	22
25	21	23	21	17	21	15	22	23	23	21

Вариант 41. $n = 55$

22	24	28	23	26	20	21	18	25	23	20
25	19	25	21	19	25	24	20	24	21	24
20	27	26	17	15	26	23	19	19	22	21
18	24	23	25	29	21	25	24	22	23	23
26	21	24	22	18	22	16	23	24	24	22

Вариант 42. $n = 55$

25	27	31	26	27	23	24	21	28	26	23
28	22	28	24	22	28	27	23	27	24	27
23	30	29	20	18	29	26	22	22	25	24
21	27	26	28	32	24	28	27	25	26	26
29	25	27	25	21	25	19	26	27	27	25

Вариант 43. $n = 55$

23	25	29	24	27	21	22	19	26	24	21
26	20	26	22	20	26	25	21	25	22	25
21	28	27	18	16	27	24	20	20	23	22
19	25	24	26	30	22	26	25	23	24	24
27	23	25	23	19	23	17	24	25	25	23

Вариант 44. $n = 55$

24	26	30	25	28	22	23	20	27	25	22
27	21	27	23	21	27	26	22	26	23	26
22	29	28	19	17	28	25	21	21	24	23
20	26	25	27	31	23	27	26	24	25	25
28	24	26	24	20	24	18	25	26	26	24

Вариант 45. $n = 55$

28	30	34	29	32	26	27	24	31	29	26
31	25	31	27	25	31	30	26	30	27	30
26	33	32	23	21	32	29	25	25	28	27
24	30	29	31	35	27	31	30	28	29	29
32	28	30	28	24	28	22	29	30	30	28

Вариант 46. $n = 100;\ \ k = 9$

113	50	52	62	41	104	135	268	29	113
6	64	202	188	26	116	175	267	193	257
70	65	25	75	38	265	323	164	144	63
60	152	27	98	3	108	77	106	205	8
23	100	119	7	83	38	19	176	69	138
9	18	294	101	33	142	48	57	38	14
105	26	203	48	211	40	317	116	566	102
172	72	3	137	92	95	18	167	147	141
72	17	36	25	119	99	88	2	12	125
11	217	469	49	197	1	23	136	77	72

Вариант 47. $n = 100;\ \ k = 9$

91	99	106	93	96	91	96	111	86	99
100	94	99	114	101	91	106	101	114	97
97	109	95	101	91	93	98	101	99	115
98	112	100	106	89	103	114	99	108	101
111	116	108	94	90	80	94	98	92	103
102	101	97	88	86	80	83	116	100	103
82	111	101	127	103	94	102	100	86	88
98	118	110	85	85	89	95	107	117	91
112	97	109	100	100	98	96	85	99	101
112	102	89	99	113	89	80	94	93	93

Вариант 48. $n = 100;\ \ k = 9$

365	282	10	218	126	341	340	140	58	378
52	80	384	233	37	18	211	369	96	87
139	274	239	187	301	66	128	87	43	266
181	22	57	89	43	203	314	137	272	50
42	112	228	245	153	139	342	197	290	234
200	299	279	262	227	367	107	384	358	140
392	37	355	253	221	136	232	206	152	296
380	310	246	66	234	236	148	242	156	387
43	393	295	243	148	376	324	318	309	391
58	317	73	353	166	113	42	143	248	45

Вариант 49. $n = 100; \ k = 9$

69	11	39	86	74	46	132	106	13	49
149	10	68	46	70	56	275	25	34	47
34	7	79	38	144	54	28	11	50	7
4	79	55	80	136	12	13	45	25	11
4	10	27	236	70	171	324	75	20	70
123	96	117	16	10	76	40	253	86	3
83	135	92	67	87	20	3	8	167	53
27	51	58	53	8	54	12	59	30	38
28	29	213	96	100	47	114	193	82	5
130	37	15	44	68	82	18	50	112	2

Вариант 50. $n = 100; \ k = 9$

204	218	212	146	209	241	188	241	248	264
177	202	260	235	196	170	154	190	198	137
170	168	199	156	218	174	242	218	177	185
222	196	227	149	244	231	187	193	231	136
177	174	200	142	196	219	224	152	153	214
168	217	139	228	226	240	213	206	173	273
194	216	233	187	172	205	181	156	147	151
212	230	207	179	220	191	203	207	159	213
224	242	194	166	180	142	130	284	189	213ё
131	227	163	221	211	198	219	236	146	189

Вариант 51. $n = 100; \ k = 9$

22	4	164	205	188	27	26	11	351	376
199	83	70	54	308	176	36	159	355	58
140	63	51	31	81	2	15	32	114	24
67	250	455	1	141	19	104	78	309	75
69	274	1	540	83	199	9	143	74	88
235	361	62	115	84	519	182	98	105	277
356	77	128	256	36	75	55	520	307	147
42	23	431	5	1	105	133	122	273	272
765	84	284	157	105	50	92	263	154	139
67	126	198	15	301	26	1	216	185	194

Вариант 52. $n = 100; \ k = 9$

78	75	80	107	92	114	49	79	57	99
112	81	102	85	68	100	78	88	76	60
95	82	98	66	81	81	92	82	107	140
78	93	83	74	75	124	57	78	104	80
61	66	84	84	102	93	95	102	92	93
116	79	79	82	67	119	114	74	88	66
106	84	119	86	105	106	78	104	81	95
87	103	63	86	84	89	100	113	70	87
83	91	103	100	42	96	104	94	145	80
92	93	103	80	87	62	86	110	109	66

Вариант 53. $n = 100; \ k = 9$

59	299	101	80	93	576	78	314	56	122
346	34	227	131	38	215	29	394	159	509
268	8	174	105	331	70	347	433	67	100
137	57	64	276	76	33	196	80	36	466
228	24	42	14	81	337	179	32	50	95
34	179	102	29	277	74	50	11	51	114
29	31	55	2	18	79	48	244	36	9
51	324	71	17	25	9	112	106	529	17
161	4	137	117	161	121	539	25	279	1
153	2	41	241	429	142	152	55	525	201

Вариант 54. $n = 100; \ k = 9$

123	117	152	95	152	128	107	97	170	81
152	85	88	162	152	81	84	169	151	138
127	168	164	187	99	107	116	160	117	113
81	110	122	132	104	148	144	112	138	116
60	90	110	157	112	162	92	113	179	156
101	155	140	139	119	156	115	148	162	197
73	130	155	91	134	121	112	129	117	160
107	125	112	102	154	99	119	136	90	133
107	118	167	161	135	114	131	144	143	133
146	73	70	121	106	61	127	115	176	70

Вариант 55. $n = 100; \ k = 9$

154	191	196	125	219	22	27	98	84	157
134	99	7	9	34	136	261	216	6	168
257	332	27	135	204	192	68	55	51	11
104	91	217	77	152	12	250	196	12	90
112	58	214	147	163	158	36	164	59	31
24	72	104	78	170	97	101	59	20	140
253	6	3	15	14	60	122	40	84	66
178	318	18	75	75	24	19	312	57	32
52	429	466	233	278	237	84	112	97	915
92	210	142	310	89	59	21	182	284	20

Вариант 56. $n = 100; \ k = 9$

89	128	87	109	138	110	119	129	120	137
139	92	89	107	112	106	101	82	121	121
109	100	72	87	108	82	109	134	74	106
97	111	100	113	80	88	102	62	124	74
75	113	113	100	72	106	112	95	68	101
65	89	99	96	98	111	39	92	114	91
103	129	119	139	145	86	136	64	80	122
92	97	82	78	110	66	87	93	136	122
89	128	113	75	83	80	107	133	123	91
102	86	105	80	103	105	102	106	95	59

Вариант 57. $n = 100; \ k = 9$

142	372	274	336	252	172	148	307	325	111
414	272	58	343	192	85	142	283	360	176
442	475	217	318	90	66	100	313	387	244
74	332	170	138	77	476	290	484	182	444
105	199	67	212	385	267	405	385	473	95
305	396	71	272	277	431	225	497	354	308
383	316	126	272	62	423	237	106	173	473
399	142	415	161	209	152	378	406	88	137
310	152	80	338	198	484	106	449	305	304
343	308	475	215	358	234	453	227	150	138

Вариант 58. $n = 100; \ k = 9$

103	153	308	10	180	30	480	18	148	191
188	115	4	42	317	27	25	227	187	136
22	52	152	50	111	81	94	213	523	129
10	556	34	89	26	28	92	65	250	201
63	65	196	479	220	33	197	57	12	190
3	233	26	119	174	78	45	592	54	122
170	104	58	156	180	66	342	124	15	476
87	6	38	45	169	34	115	10	176	13
23	382	56	48	10	17	0	275	563	48
41	251	9	170	144	35	193	118	82	58

Вариант 59. $n = 100; \ k = 9$

92	135	67	85	140	136	79	164	91	96
89	196	110	117	140	233	116	165	100	82
102	169	153	109	162	129	61	53	113	148
43	139	120	156	143	34	139	153	52	107
55	138	96	127	105	146	106	184	139	176
90	117	147	158	128	120	120	160	158	210
120	148	152	122	101	86	153	120	160	149
136	115	112	55	177	195	123	84	111	95
113	166	86	149	144	125	74	162	101	31
119	153	114	58	125	172	105	205	168	97

Вариант 60. $n = 100; \ k = 9$

485	98	116	102	109	357	73	208	49	7
420	105	356	38	266	201	3	94	52	49
82	150	171	12	0	119	84	17	221	3
53	44	233	9	125	202	314	215	27	59
40	42	517	42	96	35	94	9	101	292
25	2	270	38	258	59	17	26	342	73
433	45	69	22	505	108	59	131	22	157
131	90	34	32	133	157	154	42	171	61
130	61	64	40	35	139	20	772	10	8
30	0	28	4	1	48	353	52	25	13

Вариант 61. $n = 100$; $k = 9$

254	284	148	262	267	303	181	303	288	288
265	160	266	200	224	218	301	201	284	165
291	254	317	288	236	269	249	223	286	308
154	307	239	283	105	277	245	306	278	219
302	239	345	270	261	193	232	217	270	229
263	296	129	210	262	245	286	298	302	249
241	204	209	246	193	239	357	223	242	216
269	266	187	247	221	262	176	242	285	197
259	214	287	258	286	274	250	283	205	185
200	257	215	307	259	168	267	203	253	202

Вариант 62. $n = 100$; $k = 9$

102	159	175	168	274	47	61	87	54	40
18	72	22	0	185	13	20	48	65	103
10	293	84	143	39	33	35	12	50	57
62	69	42	3	15	215	123	11	10	14
44	151	508	62	2	76	32	136	87	48
189	52	17	120	75	27	188	186	18	170
39	72	99	157	10	191	52	182	86	37
9	51	204	30	67	18	40	42	100	40
178	89	43	79	193	33	138	111	28	79
47	30	156	130	15	156	68	69	350	104

Вариант 63. $n = 100$; $k = 9$

310	276	378	328	296	333	281	333	224	398
351	337	379	410	302	311	325	308	268	343
406	271	264	402	365	298	363	337	247	329
339	357	335	421	446	372	195	395	279	350
350	369	332	320	304	283	378	252	340	315
431	474	324	405	447	236	300	431	318	227
250	333	299	415	280	410	223	350	313	300
346	361	334	304	321	296	266	299	282	426
394	345	183	299	401	262	278	243	333	287
461	349	352	361	327	286	323	302	369	261

Вариант 64. $n = 100$; $k = 9$

5	133	81	22	59	58	102	10	29	135
109	7	9	19	15	0	118	86	156	25
51	74	84	97	28	38	44	214	185	88
3	57	123	80	205	18	40	101	98	37
134	28	23	115	40	3	56	63	4	172
35	40	16	18	173	14	92	9	78	36
36	183	58	57	18	27	51	2	47	11
177	147	54	4	110	18	28	129	80	68
99	56	31	166	73	31	20	3	63	14
58	56	48	323	22	90	137	406	90	78

Вариант 65. $n = 100$; $k = 9$

352	506	594	565	588	556	663	454	573	465
571	573	330	636	474	651	599	575	619	506
501	593	531	496	507	555	457	530	459	323
552	517	557	663	613	635	372	447	526	614
556	709	447	450	590	503	532	504	495	502
666	594	562	389	648	628	541	659	631	515
365	455	503	433	504	541	419	471	440	599
446	538	732	584	560	723	503	540	646	663
494	509	592	627	635	446	462	615	499	484
593	354	555	397	692	654	555	678	553	612

Вариант 66. $n = 100$; $k = 9$

129	246	11	191	374	75	249	20	55	113
147	67	103	36	142	109	57	128	208	255
89	27	19	34	61	63	117	184	28	71
12	119	5	135	24	131	110	64	1	9
14	35	119	46	1	18	35	117	63	36
11	41	4	150	194	37	22	202	104	71
100	20	89	146	38	70	31	81	32	54
204	61	19	106	15	58	84	23	5	162
410	20	21	271	122	55	355	152	38	203
79	10	1	102	62	54	94	258	10	155

Вариант 67. $n = 100;\ k = 9$

329	255	99	178	316	290	216	217	389	284
260	306	367	253	261	248	309	209	308	207
333	227	250	288	199	322	176	294	267	86
175	150	294	255	335	320	221	242	387	248
166	181	360	217	128	147	221	273	286	393
181	278	280	214	268	278	168	283	272	255
189	315	281	195	283	203	385	201	209	322
352	317	206	267	273	227	212	230	235	156
287	105	287	208	250	328	100	186	279	252
182	254	315	312	381	218	236	220	248	240

ЗАДАЧА 2

Дана корреляционная таблица. По данным таблицы необходимо выполнить следующие две лабораторные работы.

Лабораторная работа №4. Расчет параметров корреляционной зависимости (прямая или обратная линейная зависимость).

1. Найти эмпирические распределения каждой из компонент, сделать их графическое изображение.
2. Вычислить средние значения $\overline{X}$ и $\overline{Y}$ и средние квадратические отклонения σ_X; σ_Y.
3. Вычислить эмпирические линии регрессии $Y = Y(X)$ и $X = X(Y)$, сделать отдельные графики этих линий.
4. Рассчитать параметры теоретических линий регрессии $Y_x = aX + b$ и $X_y = a_1 Y + b_1$. Сделать графики этих линий, располагая каждую на графике соответствующей эмпирической линии.

Лабораторная работа №5 (на усмотрение преподавателя). Расчет параметров корреляционной зависимости (квадратичная зависимость).

Пункты 1 – 3 – те же, что и в лабораторной работе №4.

5. Используя метод наименьших квадратов (см. [10] в учебных пособиях кафедры), вычислить параметры квадратичной зависимости $Y(X) = a_1X^2 + b_1X + c_1$ и $X(Y) = a_2Y^2 + b_2Y + c_2$ (можно только одну зависимость).
6. На графике соответствующей эмпирической зависимости построить график полученной теоретической квадратической зависимости.

Вариант 1. В табл.1 дано распределение 55 однотипных предприятий по количеству выпускаемых изделий X (тыс. шт.) и полным затратам на их производство Y (млн руб.):

Таблица 1

X	Y					
	50 – 80	80 – 110	110 – 140	140 – 170	170 – 200	200 – 230	Итого
25 - 35	5	-	-	-	-	-	5
35 – 45	4	12	-	-	-	-	16
45 – 55	-	8	5	4	-	-	17
55 – 65	-	1	5	7	2	-	15
65 – 75	-	-	-	-	1	1	2
Итого	9	21	10	11	3	1	55

Используя соответствующее уравнение регрессии, оценить среднее значение полных затрат на производство 53 тыс. деталей.

Вариант 2. В табл. 2 дано распределение 100 заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т):

Таблица 2

X	Y						
	5 – 15	15 – 25	25 – 35	35 – 45	45 – 55	55 - 65	Итого
4 – 8	1	3	-	-	-	-	4
8 – 12	2	6	7	1	-	-	16
12 – 16	-	1	9	16	21	10	57
16 – 20	-	-	-	8	4	3	16
20 – 24	-	-	-	-	5	2	7
Итого	3	11	16	25	30	15	100

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 18 млн. руб.

Вариант 3. В табл. 3 дано распределение 60 предприятий по величине основных промышленно-производственных фондов X (млн руб.) и себестоимости продукции Y (руб.)

Таблица 3

X	Y					
	16 – 24	24 – 32	32 – 40	40 – 48	48 – 56	Итого
15 – 30	-	-	-	-	2	2
30 – 45	-	-	4	8	4	16
45 – 60	1	7	12	6	-	26
60 – 75	4	7	2	-	-	13
75 – 90	1	2	-	-	-	3
Итого	6	16	18	14	6	60

Используя соответствующее уравнение регрессии, определить среднюю себестоимость продукции для предприятий с основными промышленно-производственными фондами 65 млн. руб.

Вариант 4. В табл. 4 дано распределение 50 магазинов по уровню издержек обращения X (%) и по годовому объему товарооборота Y (млн руб.)

Таблица 4

X	Y					
	0,5 – 2 ,0	2,0 – 3,5	3,5 – 5,0	5,0 – 6,5	6,5 – 8,0	Итого
4 – 6	-	-	-	3	2	5
6 – 8	-	4	8	8	1	21
8 – 10	2	5	5	2	-	14
10 – 12	3	1	5	-	-	9
12 – 14	1	-	-	-	-	1
Итого	6	10	18	13	3	50

Используя соответствующее уравнение регрессии, определить среднее значение уровня издержек обращения в магазинах с годовым объемом товарооборота 4,2 млн руб.

Вариант 5. В табл. 5 дано распределение 200 совхозов по затратам труда X (чел.-дн. на 1 ц зерна) и себестоимости Y (руб. за 1 ц зерна)

Таблица 5

X	Y					
	7,25 – 9,25	9,25 – 11,25	11,25 – 13,25	13,25 – 15,25	15,25 – 17,25	Итого
0,4- 0,8	14	-	-	-	-	14
0,8 -1,2	22	10	-	-	-	32
1,2-1,6	-	38	30	10	-	78
1,6-2,0	-	6	30	12	2	50
2,0-2,4	-	-	4	8	8	20
2,4-2,8	-	-	-	-	6	6
Итого	36	54	64	30	16	200

Используя соответствующее уравнение регрессии, оценить средние затраты труда на получение 1 ц зерна себестоимостью 15 р. за 1 ц.

Вариант 6. В табл. 6 дано распределение 100 магазинов по величине товарооборота X (млн руб.) и размеру торговой площади магазина Y ($м^2$).

Таблица 6

X	Y					
	100 – 150	150 – 200	200 – 250	250 – 300	300 – 350	Итого
1,0-1,5	4	12	2	-	-	18
1,5-2,0	-	4	9	9	-	22
2,0-2,5	-	2	10	18	-	30
2,5-3,0	-	-	4	9	11	24
3,0-3,5	-	-	-	3	3	6
Итого	4	18	25	39	14	100

Используя соответствующее уравнение регрессии, найти среднюю величину товарооборота магазинов, имеющих торговую площадь 220 $м^2$.

Вариант 7. В табл. 7 дано распределение 100 проб руды, добытой на руднике, по содержанию окиси X (гр) и закиси железа Y (гр):

Таблица 7

X	Y						
	50–80	80–110	110–140	140–170	170 – 200	200 – 230	Итого
40 – 50	-	-	1	6	4	6	17
50 – 60	-	-	2	18	10	2	32
60 – 70	-	6	14	2	2	-	24
70 – 80	-	6	3	-	-	-	9
80 – 90	4	8	-	-	-	-	12
90-100	6	-	-	-	-	-	6
Итого	10	20	20	26	16	8	100

Используя соответствующее уравнение регрессии, оценить среднее количество окиси железа в руде, содержащей 25гр закиси железа.

Вариант 8. Данные о живом весе X (кг) и молочной продуктивности Y (кг) 80 коров приведены в табл. 8.

Таблица 8

X	Y					
	1250-1750	1750-2250	2250-2750	2750-3250	3250-3750	Итого
325-375	3	2	-	-	-	5
375-425	-	8	7	1	-	16
425-475	-	2	5	10	-	17
475-525	-	-	13	10	7	30
525-575	-	-	-	7	5	12
Итого	3	12	25	28	12	80

Используя соответствующее уравнение регрессии, оценить среднюю молочную продуктивность коров весом 450 кг.

Вариант 9. В табл. 9 дано распределение детей четырехлетнего возраста по росту X (см) и весу Y (кг):

Таблица 9

X	Y					
	15,5-16,5	16,5-17,5	17,5-18,5	18,5-19,5	19,5-20,5	Итого
98-100	2	3	-	-	-	5
100-102	3	6	4	-	-	13
102-104	1	4	13	5	-	23
104-106	-	1	14	10	2	27
106-108	-	-	10	8	5	23
108-110	-	-	-	6	3	9
Итого	6	14	41	29	10	100

Используя соответствующее уравнение регрессии, определить средний вес детей ростом 107 см.

Вариант 10. В табл. 10 дано распределение 80 совхозов по числу рабочих X (чел.) на 100 га сельскохозяйственных угодий и объему валовой продукции Y (тыс. руб.).

Таблица 10

X	Y					
	30 – 70	70 – 110	110 – 150	150 – 190	190 – 230	Итого
8 – 16	2	3	1	-	-	6
16 – 24	3	4	5	-	-	12
24 – 32	-	8	16	12	1	37
32 – 40	-	1	8	3	4	16
40 – 48	-	-	1	2	6	9
Итого	5	16	31	17	11	80

Используя соответствующее уравнение регрессии, оценить средний объем валовой продукции на 100 га сельскохозяйственных угодий в совхозе с 35 рабочими.

Вариант 11. В табл. 11 дано распределение заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т)

Таблица 11

X	Y					
	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60	Итого
15 –25	7	20	-	-	-	27
25 – 35	5	23	30	10	-	68
35 – 45	-	-	47	11	9	67
45 – 55	-	-	2	20	7	29
55 – 65	-	-	-	6	3	9
Итого	12	43	79	47	19	200

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 52 млн руб.

Вариант 12. Дана корреляционная табл. 12 результатов измерения перепада давления X (атм.) и расход нефти Y $(м^3/ч)$ в трубопроводе.

Таблица 12

X	Y					
	3000-3200	3200-3400	3400-3600	3600-3800	3800-4000	Итого
24 – 28	2	1	8	-	-	11
28 – 32	2	1	2	-	-	5
32 – 36	-	2	11	6	15	34
36 – 40	-	-	4	2	27	33
40 – 44	-	-	-	3	14	17
Итого	4	4	25	11	56	100

Используя соответствующее уравнение регрессии, оценить средний расход нефти при перепаде давления в 34 атм.

Вариант 13. Распределение растений по весу X каждого из них и по весу семян Y (гс) заданы корреляционной табл. 13.

Таблица 13

X	Y					
	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60	Итого
35 – 45	5	7	-	-	-	12
45 – 55	-	4	16	23	-	43
55 – 65	-	8	20	32	27	87
65 – 75	-	-	11	29	2	42
75 – 85	-	-	-	9	7	16
Итого	5	19	47	93	36	200

Используя соответствующее уравнение регрессии, оценить средний вес семян при весе растения 62 гс.

Вариант 14. В табл. 14 дано распределение прямоугольных плиток по длине Y (см) и по весу X (кгс).

Таблица 14

X	Y					
	5 – 7	7 – 9	9 –11	11 – 13	13 – 15	Итого
27,5-32,5	2	17	9	3	-	31
32,5-37,5	-	10	17	9	-	36
37,5-42,5	-	3	24	16	13	56
42,5-47,5	-	-	6	24	12	42
47,5-52,5	-	-	2	11	22	35
Итого	2	30	58	63	47	200

Используя соответствующее уравнение регрессии, оценить среднюю длину плитки весом 43 кгс.

Вариант 15. В табл. 15 дано распределение 60 однотипных предприятий по количеству выпускаемых изделий X (тыс. шт.) и полным затратам на их производство Y (млн руб.).

Таблица 15

X	Y					
	40 – 70	70 – 100	100 – 130	130 – 160	160 – 190	Итого
20 –30	6	-	-	-	-	6
30 – 40	4	12	-	-	-	16
40 – 50	-	8	6	4	-	18
50 – 60	-	2	6	7	3	18
60 – 70	-	-	-	-	2	2
Итого	10	22	12	11	5	60

Используя соответствующее уравнение регрессии, оценить среднее значение полных затрат на производство 34 тыс. деталей.

Вариант 16. В табл. 16 дано распределение 100 заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т):

Таблица 16

Y	X					
	10	15	20	25	30	35
30	2	6	-	-	-	-
40	-	4	4	-	-	-
50	-	-	7	35	8	-
60	-	-	2	10	8	-
70	-	-	-	5	6	3

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 33 млн руб.

Вариант 17. В табл. 17 дано распределение 100 заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т):

Таблица 17

Y	X					
	15	20	25	30	35	40
5	4	2	-	-	-	-
10	-	6	4	-	-	-
15	-	-	6	45	2	-
20	-	-	2	8	6	-
25	-	-	-	4	7	4

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 28 млн. руб.

Вариант 18. В табл. 18 дано распределение 100 заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т):

Таблица 18

Y	X					
	5	10	15	20	25	30
20	1	5	-	-	-	-
30	-	5	3	-	-	-
40	-	-	9	40	2	-
50	-	-	4	11	6	-
60	-	-	-	4	7	3

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 18 млн руб.

Вариант 19. В табл. 19 дано распределение 100 заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т):

Таблица 19

Y	X					
	10	15	20	25	30	35
6	4	2	-	-	-	-
12	-	6	2	-	-	-
18	-	-	5	40	5	-
24	-	-	2	8	7	-
30	-	-	-	4	7	8

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 21 млн руб.

Вариант 20. В табл. 20 дано распределение 100 заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т):

Таблица 20

Y	X					
	5	10	15	20	25	30
8	2	4	-	-	-	-
12	-	3	7	-	-	-
16	-	-	5	30	10	-
20	-	-	7	10	8	-
24	-	-	-	5	6	3

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 23 млн руб.

Вариант 21. В табл. 21 дано распределение прямоугольных плиток по длине Y (см) и по весу X (кгс).

Таблица 21

Y	X					
	2	7	12	17	22	27
10	2	4	-	-	-	-
20	-	6	2	-	-	-
30	-	-	3	50	-	-
40	-	-	1	10	6	-
50	-	-	-	4	7	3

Используя соответствующее уравнение регрессии, оценить среднюю длину плитки весом 18 (кгс).

Вариант 22. В табл. 22 дано распределение 100 заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т):

Таблица 22

Y	X					
	11	16	21	26	31	36
25	2	4	-	-	-	-
35	-	6	3	-	-	-
45	-	-	6	45	4	-
55	-	-	2	8	6	-
65	-	-	-	4	7	3

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 33 млн руб.

Вариант 23. В табл. 23 дано распределение 80 однотипных предприятий по количеству выпускаемых изделий X (тыс. шт.) и полным затратам на их производство Y (млн руб.).

Таблица 23

Y	X					
	4	9	14	19	24	29
8	3	3	-	-	-	-
18	-	5	4	-	-	-
28	-	-	20	2	8	-
38	-	-	5	10	6	-
48	-	-	-	4	7	3

Используя соответствующее уравнение регрессии, оценить среднее значение полных затрат на производство 27 тыс. деталей.

Вариант 24. В табл. 24 дано распределение 103 однотипных предприятий по количеству выпускаемых изделий X (тыс. шт.) и полным затратам на их производство Y (млн руб.).

Таблица 24

Y	X					
	5	10	15	20	25	30
11	4	2	-	-	-	-
21	-	5	3	-	-	-
31	-	-	5	48	5	-
41	-	-	2	8	7	-
51	-	-	-	4	7	3

Используя соответствующее уравнение регрессии, оценить среднее значение полных затрат на производство 14 тыс. деталей.

Вариант 25. В табл. 25 дано распределение между объемом выполненных работ Y (млн руб.) и накладными расходами X (млн руб.).

Таблица 25

Y	X							
	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
100	-	-	-	-	1	1	2	2
110	-	-	-	1	3	3	5	2
120	-	-	-	2	6	9	2	1
130	-	-	1	8	14	4	1	-
140	-	1	6	15	5	2	-	-
150	-	9	10	7	1	-	-	-
160	1	6	5	2	-	-	-	-

Используя соответствующее уравнение регрессии, оценить средние накладные расходы для объема выполненных работ в 145 млн руб.

Вариант 26. В табл. 26 дана зависимость между относительным уровнем издержек обращения Y(%) и объемом сбыта товаров X (млн руб.).

Таблица 26

Y	X						
	1.22	1.23	1.24	1.25	1.26	1.27	1.28
4.1	-	-	-	-	1	3	1
4.2	-	-	1	2	3	3	-
4.3	-	-	1	6	7	1	-
4.4	-	1	3	10	2	1	-
4.5	-	1	6	5	1	-	-
4.6	1	3	4	3	1	-	-
4.7	1	2	1	1	-	-	-

Используя соответствующее уравнение регрессии, оценить среднее значение сбыта товаров при уровне издержек 4,43 %.

Вариант 27. В табл. 27 дана зависимость между относительным уровнем издержек обращения Y(%) и объемом сбыта товаров X (млн руб.).

Таблица 27

Y	X							
	4.3	4.4	4.5	4.6	4.7	4.8	4.9	5.0
3.9	1	3	2	1	-	-	-	-
4.0	-	1	5	2	1	-	-	-
4.1	-	1	3	8	4	1	-	-
4.2	-	-	1	2	10	3	-	-
4.3	-	-	-	1	3	6	4	1
4.4	-	-	-	-	-	1	3	1
4.5	-	-	-	-	-	-	1	2
4.6	-	-	-	-	-	-	-	1

Используя соответствующее уравнение регрессии, оценить среднее значение сбыта товаров при уровне издержек 4 %.

Вариант 28. В табл. 28 дана зависимость между относительным уровнем издержек обращения Y (%) и объемом сбыта товаров X (млн руб.).

Таблица 28

Y	X						
	11.1	11.2	11.3	11.4	11.5	11.6	11.7
1.2	1	2	3	2	1	-	-
1.3	1	2	4	4	2	1	-
1.4	1	2	6	7	4	3	1
1.5	1	3	6	9	7	2	1
1.6	-	1	3	6	6	3	1
1.7	-	1	2	4	5	3	1
1.8	-	-	1	2	5	3	1

Используя соответствующее уравнение регрессии, оценить среднее значение сбыта товаров при уровне издержек 1,62 %.

Вариант 29. В табл. 29 дана зависимость между относительным уровнем издержек обращения Y (%) и объемом сбыта товаров X (млн руб.).

Таблица 29

Y	X						
	1	2	3	4	5	6	7
0.12	1	-	-	-	-	-	-
0.13	3	1	-	-	-	-	-
0.14	-	6	1	-	-	-	-
0.15	-	1	6	2	-	-	-
0.16	-	-	1	9	1	-	-
0.17	-	-	-	2	7	1	-
0.18	-	-	-	1	1	5	2

Используя соответствующее уравнение регрессии, оценить среднее значение сбыта товаров при уровне издержек 0,154 %.

Вариант 30. В табл. 30 дано распределение совхозов по числу рабочих X (чел.) на 100 га сельскохозяйственных угодий и объему валовой продукции Y (тыс. руб.).

Таблица 30

Y	X								
	125	135	145	155	165	175	185	195	205
11	2	2	1	-	-	-	-	-	-
12	2	5	4	2	1	-	-	-	-
13	1	3	8	6	5	2	-	-	-
14	-	1	5	13	10	5	1	-	-
15	-	-	1	9	20	8	3	1	-
16	-	-	-	3	9	14	5	1	-
17	-	-	-	1	4	7	9	3	1
18	-	-	-	-	2	3	4	6	1
19	-	-	-	-	-	1	1	2	3

Используя соответствующее уравнение регрессии, оценить средний объем валовой продукции на 100 га сельскохозяйственных угодий в совхозе с 140 рабочими.

ПРИЛОЖЕНИЯ

ПРИЛОЖЕНИЕ 1

Таблица значений функции $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-\frac{z^2}{2}}dz$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1	2	3	4	5	6	7	8
0,00	0,0000	0,32	0,1255	0,64	0,2389	0,96	0,3315
0,01	0,0040	0,33	0,1293	0,65	0,2422	0,97	0,3340
0,02	0,0080	0,34	0,1331	0,66	0,2454	0,98	0,3365
0,03	0,0120	0,35	0,1368	0,67	0,2486	0,99	0,3389
0,04	0,0160	0,36	0,1406	0,68	0,2517	1,00	0,3413
0,05	0,0199	0,37	0,1443	0,69	0,2549	1,01	0,3438
0,06	0,0239	0,38	0,1480	0,70	0,2580	1,02	0,3461
0,07	0,0279	0,39	0,1517	0,71	0,2611	1,03	0,3485
0,08	0,0319	0,40	0,1554	0,72	0,2642	1,04	0,3508
0,09	0,0359	0,41	0,1591	0,73	0,2673	1,05	0,3531
0,10	0,0398	0,42	0,1628	0,74	0,2703	1,06	0,3554
0,11	0,0438	0,43	0,1664	0,75	0,2784	1,07	0,3577
0,12	0,0478	0,44	0,1700	0,76	0,2764	1,08	0,3599
0,13	0,0517	0,45	0,1736	0,77	0,2794	1,09	0,3621
0,14	0,0557	0,46	0,1772	0,78	0,2823	1,10	0,3643
0,15	0,0596	0,47	0,1808	0,79	0,2852	1,11	0,3665
0,16	0,0636	0,48	0,1844	0,80	0,2881	1,12	0,3686
0,17	0,0675	0,49	0,1879	0,81	0,2910	1,13	0,3708
0,18	0,0714	0,50	0,1915	0,82	0,2939	1,14	0,3729
0,19	0,0753	0,51	0,1950	0,83	0,2967	1,15	0,3949
0,20	0,0793	0,52	0,1985	0,84	0,2995	1,16	0,3770
0,21	0,0832	0,53	0,2019	0,85	0,3023	1,17	0,3790
0,22	0,0871	0,54	0,2054	0,86	0,3051	1,18	0,3810
0,23	0,0910	0,55	0,2088	0,87	0,3078	1,19	0,3830
0,24	0,0948	0,56	0,2123	0,88	0,3106	1,20	0,3849
0,25	0,0987	0,57	0,2157	0,89	0,3133	1,21	0,3869
0,26	0,1026	0,58	0,2190	0,90	0,3159	1,22	0,3883
0,27	0,1064	0,59	0,2224	0,91	0,3186	1,23	0,3907
0,28	0,1103	0,60	0,2257	0,92	0,3212	1,24	0,3925
0,29	0,1141	0,61	0,2291	0,93	0,3238	1,25	0,3944
0,30	0,1179	0,62	0,2324	0,94	0,3264		
0,31	0,1217	0,63	0,2357	0,95	0,3289		

Продолжение прил. 1

1	2	3	4	5	6	7	8
1,26	0,3962	1,59	0,4441	1,92	0,4726	2,50	0,4938
1,27	0,3980	1,60	0,4452	1,93	0,4732	2,52	0,4941
1,28	0,3997	1,61	0,4463	1,94	0,438	2,54	0,4945
1,29	0,4015	1,62	0,4474	1,95	0,4744	2,56	0,4948
1,30	0,4032	1,63	0,4484	1,96	0,4750	2,58	0,4951
1,31	0,4049	1,64	0,4495	1,97	0,4756	2,60	0,4953
1,32	0,4066	1,65	0,4505	1,98	0,4761	2,62	0,4956
1,33	0,4082	1,66	0,4515	1,99	0,4767	2,64	0,4959
1,34	0,4099	1,67	0,4525	2,00	0,4772	2,66	0,4961
1,35	0,4115	1,68	0,4535	2,02	0,4783	2,68	0,4963
1,36	0,4131	1,69	0,4545	2,04	0,4793	2,70	0,4965
1,37	0,4147	1,70	0,4554	2,06	0,4803	2,72	0,4967
1,38	0,4162	1,71	0,4564	2,08	0,4812	2,74	0,4969
1,39	0,4177	1,72	0,4573	2,10	0,4821	2,76	0,4971
1,40	0,4192	1,73	0,4582	2,12	0,4830	2,78	0,4973
1,41	0,4207	1,74	0,4591	2,14	0,4838	2,80	0,4974
1,42	0,4222	1,75	0,4599	2,16	0,4846	2,82	0,4976
1,43	0,4236	1,76	0,4608	2,18	0,4854	2,84	0,4977
1,44	0,4251	1,77	0,4616	2,20	0,4861	2,86	0,4979
1,45	0,4265	1,78	0,4625	2,22	0,4868	2,88	0,4980
1,46	0,4279	1,79	0,4633	2,24	0,4875	2,90	0,4981
1,47	0,4292	1,80	0,4641	2,26	0,4881	2,92	0,4982
1,48	0,4306	1,81	0,4649	2,28	0,4887	2,94	0,4984
1,49	0,4319	1,82	0,4656	2,30	0,4893	2,96	0,4985
1,50	0,4332	1,83	0,4664	2,32	0,4898	2,98	0,4986
1,51	0,4345	1,84	0,4671	2,34	0,4904	3,00	0,49865
1,52	0,4357	1,85	0,4678	2,36	0,4909	3,20	0,49931
1,53	0,4370	1,86	0,4686	2,38	0,4913	3,40	0,49966
1,54	0,4382	1,87	0,4693	2,40	0,4918	3,60	0,499841
1,55	0,4394	1,88	0,4699	2,42	0,4922	3,80	0,499928
1,56	0,4406	1,89	0,4706	2,44	0,4927	4,00	0,499968
1,57	0,4418	1,90	0,4713	2,46	0,4931	4,50	0,499997
1,58	0,4429	1,91	0,4719	2,48	0,4934	5,00	0,499999

ПРИЛОЖЕНИЕ 2

Таблица значений $t_{\gamma} = t(\gamma, n)$

γ / n	0,95	0,99	0,999	γ / n	0,95	0,99	0,999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,720	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,991	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

ПРИЛОЖЕНИЕ 3

Таблица значений $q = q(\gamma, n)$

γ / n	0,95	0,99	0,999	γ / n	0,95	0,99	0,999
5	1,37	2,67	5,64	20	0,37	0,58	0,88
6	1,09	2,01	3,88	25	0,32	0,49	0,73
7	0,92	1,62	2,98	30	0,28	0,43	0,63
8	0,80	1,38	2,42	35	0,26	0,38	0,56
9	0,71	1,20	2,06	40	0,24	0,35	0,50
10	0,65	1,08	1,80	45	0,22	0,32	0,46
11	0,59	0,98	1,60	50	0,21	0,30	0,43
12	0,55	0,90	1,45	60	0,188	0,269	0,38
13	0,52	0,83	1,33	70	0,174	0,245	0,34
14	0,48	0,78	1,23	80	0,161	0,226	0,31
15	0,46	0,73	1,15	90	0,151	0,211	0,29
16	0,44	0,70	1,07	100	0,143	0,198	0,27
17	0,42	0,66	1,01	150	0,115	0,160	0,211
18	0,40	0,63	0,96	200	0,099	0,136	0,185
19	0,39	0,60	0,92	250	0,089	0,120	0,162

ПРИЛОЖЕНИЕ 4

Критические точки распределения χ^2

Число степеней свободы r	Уровень значимости α					
	0,01	0,025	0,05	0,95	0,975	0,89
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26,1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	7,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63
20	37,6	34,2	31,4	10,9	9,59	8,26
21	38,9	35,5	32,7	11,6	10,3	8,90
22	40,3	36,8	33,9	12,3	11,0	9,54
23	41,6	38,1	35,2	13,1	11,7	10,2
24	43,0	39,4	36,4	13,8	12,4	10,9
25	44,3	40,6	37,7	14,6	13,1	11,5
26	45,6	41,9	38,9	15,4	13,8	12,2
27	47,0	43,2	40,1	16,2	14,6	12,9
28	48,3	44,5	41,3	16,9	15,3	13,6
29	49,6	45,7	42,6	17,7	16,0	14,3
30	50,9	47,0	43,8	18,5	16,8	15,0

СПИСОК ЛИТЕРАТУРЫ

Основная

1. Гмурман, В.Е. Теория вероятностей и математическая статистика / В.Е.Гмурман. – М.: Высшее образование : Юрайт, 2009. - 479с.

2. Гмурман, В.Е. Руководство к решению задач по теории вероятностей и математической статистике / В.Е.Гмурман. –М.: Высшее образование, 2009.- 404 с.

3. Данко, П.Е. Высшая математика в упражнениях и задачах /П.Е.Данко, А.Г.Попов, Т.Я.Кожевникова. - Ч.1, 2. –М.: ОНИКС 21век; Мир и Образование, 2006.-416с.

4. Феллер, В. Введение в теорию вероятностей и ее приложения/ В.Феллер. – Т.1,2. -М.: Мир, 1984.

Дополнительная

5. Вентцель, Е.С. Теория вероятностей и ее инженерные приложения / Е.С.Вентцель, Л.А. Овчаров.– М.: Высшая школа, 2007.-479с.

6. Кремер, Н.Ш. Теория вероятностей и математическая статистика / Н.Ш Кремер. – М.: ЮНИТИ- ДАНА, 2002.-543с.

Учебные пособия кафедры

7. Учебно -методическое пособие по математической статистике/ Т.В.Умергалина, З.Р Хакимова. – Уфа: УГНТУ, 2004.

8. Практикум по математической статистике / Т.В.Умергалина, З.Р. Хакимова. – Уфа: УГНТУ, 2003.

9. Расчетные задания по математической статистике / Т.В.Умергалина, З.Р. Хакимова. – Уфа: УГНТУ, 2007.

10. Функции нескольких переменных: учебно-методическое пособие / Р.Г.Гимаев, Т.В.Умергалина.- Уфа: УГНТУ, 2005 (Раздел «Метод наименьших квадратов»).

Учебное издание

Бахтизин Рамиль Назифович, Галиуллин Марат Мидхатович, Галиакбарова Эмилия Вильевна, Гимаев Роберт Гильмутдинович, Зарипов Раиль Муталлапович, Исламгулова Галия Файзеевна, Ковалева Элла Александровна, Лазарев Владимир Анатольевич, Майский Равиль Анварович, Мухаметзянов Ирек Зирягович, Нагаева Зиля Мунировна, Сахарова Лидия Александровна, Сокова Инна Александровна, Сулейманов Игорь Нугуманович, Умергалина Татьяна Васильевна, Фаткуллин Николай Юрьевич, Хайбуллин Радик Яруллович, Хакимова Зульфия Разифовна, Чернятьева Рита Раисовна, Юлдыбаев Лев Хадиевич, Шамшович Валентина Федоровна, Шварева Елена Николаевна, Якупов Вагизьян Минигалиевич, Янчушка Анна Павловна, Абзалимов Рамиль Рафикович, Акмадиева Танслу Рифхатовна, Аносова Елизавета Петровна, Ахтямов Наиль Тагирович, Байрамгулова Римма Сагитовна

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

2-е издание, исправленное и дополненное

2. Методические указания для студентов

2.1. ЛАБОРАТОРНАЯ РАБОТА № 1
«Первичная обработка статистических данных»

Из данных, входящих в выборку $X=\{x_1, x_2, \ldots, x_n\}$ (см. табл.1.1), находим x_{min} и x_{max}, соответственно наименьшее и наибольшее значения выборки, и вычисляем число $d = x_{max} - x_{min}$, называемое ***размахом выборки***. Размах выборки – это длина основного интервала, в который попадают все значения выборки. Далее значения x_i $\left(i=\overline{1,n}\right)$ можно упорядочить, то есть расположить в порядке возрастания. Тогда выборка $X=\{x_1, x_2, \ldots, x_n\}$, записанная по возрастанию, называется ***вариационным рядом*** и её значения x_i - вариантами. По формуле $k = 1 + 4\cdot[\lg n]$, где n- объем выборки X, $[\lg n]$– целая часть числа $\lg n$, определим число k. Данное число k задает количество подынтервалов, на которые разбиваем основной интервал $[x_{min}; x_{max}]$. Вычисляем длину подынтервалов по формуле

$$h = \frac{d}{k} \tag{2.1}$$

и затем – границы подынтервалов:

$$a_0 = x_{min};\ \ a_1 = a_0 + h; \ldots; a_j = a_{j-1} + h; \ldots; a_k = x_{max}. \tag{2.2}$$

Находим m_j $\left(j=\overline{1,k}\right)$– частоты и $\mu_j = \frac{m_j}{n}\left(j=\overline{1,k}\right)$ относительные частоты попадания значений выборки X в $j-$й подынтервал. Причем должно быть $\sum_{j=1}^{k} m_j = m_1 + m_2 + \ldots + m_k = n$; для относительных частот: $\sum_{j=1}^{k} \mu_j = 1.$

В результате проведенных расчетов получаем две таблицы:

Таблица 2.1

X	$[a_0;a_1)$	$[a_1;a_2)$	$\ldots$	$[a_{k-1};a_k]$
m	m_1	m_2	$\ldots$	m_k

Таблица 2.2

X	$[a_0;a_1)$	$[a_1;a_2)$	$\ldots$	$[a_{k-1};a_k]$
μ	μ_1	μ_2	$\ldots$	μ_k

Далее, если найти середины подынтервалов:

$$b_1 = \frac{a_1 + a_0}{2};\ b_2 = \frac{a_2 + a_1}{2}; \ldots ; b_k = \frac{a_k + a_{k-1}}{2}, \qquad (2.3)$$

Таблица 2.3

X	b_1	b_2	$\ldots$	b_k
μ	μ_1	μ_2	$\ldots$	μ_k

В целях наглядности полученных в табл. 2.1, 2.2, 2.3 данных пользуются различными способами их графического изображения. К ним относятся гистограмма и полигон.

Для построения гистограммы относительных частот используем данные табл.2.2. В декартовой системе координат на оси $0X$ находим значения x_{min} и x_{max} и тем самым находим границы основного интервала, в который попадают все значения выборки. Затем на этом интервале откладываем границы подынтервалов. По оси $0Y$ откладываем значения относительных частот μ_j $\left(j = \overline{1,\, k}\right)$. Тогда гистограммой относительных частот назовем ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные подынтервалы длины h, а высоты равны числам μ_j $\left(j = \overline{1,\, k}\right)$. Аналогично, по данным табл. 2.1, строится гистограмма частот.

Для построения полигона относительных частот используем данные табл. 2.3. В декартовой системе координат на оси $0X$ находим x_{min} и x_{max}, то есть изображаем границы основного интервала. Затем наносим значения середин подынтервалов b_j. По оси $0Y$ откладываем значения, соответствующие относительным частотам μ_j.

Полигоном относительных частот называют ломаную, отрезки которой соединяют точки $(b_1, \mu_1);\ (b_2, \mu_2); \ldots ; (b_k, \mu_k)$.

Данные табл. 2.3 представляют эмпирический закон распределения выборки, а полигон относительных частот есть его визуальное представление.

Эмпирической функцией распределения (функцией распределения выборки) называют функцию $F^*(x)$, определяющую для каждого значения x относительную частоту события $X < x$.

Таким образом, $F^*(x) = \dfrac{n_x}{n}$, где n_x – число вариант, меньших x, n – объем выборки.

Для каждой реализации выборки эмпирическая функция распределения однозначно определена и обладает всеми свойствами теоретической функции распределения:

1) значения эмпирической функции принадлежат отрезку $[0;1]$;

2) $F^*(x)$ – не убывающая функция;

3) если $x_1 -$ наименьшая варианта, то $F^*(x) = 0$ при $x \le x_1$; если $x_k -$наибольшая варианта, то $F^*(x) = 1$ при $x > x_k$.

Эмпирическая функция распределения выборки является оценкой теоретической функции распределения генеральной совокупности.

ПРИМЕР 2.1

Дана выборка X из генеральной совокупности объема n=100.

Таблица 2.4

254	1158	522	524	972	736	401	347	208	368
1485	812	1032	226	428	368	676	671	587	701
701	1171	443	683	786	895	267	597	51	941
659	400	484	876	570	241	678	127	728	903
424	245	531	986	1017	429	732	1021	430	153
513	520	221	1074	826	65	389	1180	504	325
294	447	1459	589	307	461	1434	559	837	743
382	387	967	446	763	767	349	853	578	652
285	628	688	517	380	375	878	409	109	621
712	476	432	721	1300	577	580	909	690	757

1. Находим из выборки x_{min} и x_{max}, рассчитываем размах выборки d:

$$x_{max} = 1485; \quad x_{min} = 51; \quad d = x_{max} - x_{min} = 1485 - 51 = 1434.$$

2. Составим вариационный ряд, для чего всю последовательность выборки расположим в порядке возрастания

51	65	109	127	153	208	221	226	241	245
254	267	285	294	307	325	347	349	368	368
375	380	382	387	389	400	401	409	424	428
429	430	432	443	446	447	461	476	484	504
513	517	520	522	524	531	559	570	577	578
580	587	589	597	621	628	652	659	671	676
678	683	688	690	701	701	712	721	728	732
736	743	757	763	767	786	812	826	837	853
876	878	895	903	909	941	967	972	986	1017
1021	1032	1074	1158	1171	1180	1300	1434	1459	1485

3. Вычисляем число $k -$ количество частичных подынтервалов, на которое разбиваем нашу выборку X: $k = 1 + 4 \cdot \lg 100 = 9$. Исходя из этого вычисляем длину подынтервалов $h = \dfrac{1434}{9} \approx 159{,}333$ и границы подынтервалов a_j, $j = \overline{1{,}9}$: $a_0 = x_{min} = 51;$

$$a_1 = a_0 + h = 51 + 159{,}333 = 210{,}333;$$

$$a_2 = a_1 + h = 210{,}333 + 159{,}333 = 369{,}666;$$

$a_3 = a_2 + h = 369{,}666 + 159{,}333 = 528{,}999$;

$a_4 = a_3 + h = 528{,}999 + 159{,}333 = 688{,}332$;

$a_5 = a_4 + h = 688{,}332 + 159{,}333 = 847{,}665$;

$a_6 = a_5 + h = 847{,}665 + 159{,}333 = 1006{,}998$;

$a_7 = a_6 + h = 1006{,}998 + 159{,}333 = 1166{,}331$;

$a_8 = a_7 + h = 1166{,}331 + 159{,}333 = 1325{,}664$;

$a_9 = a_8 + h = x_{max} = 1485$.

4. Рассчитываем частоты – число попаданий в подынтервалы значений из выборки, то есть $m_1, m_2, \ldots, m_9$.

$[51;\ 210{,}333)\quad m_1 = 5$

$[210{,}333;\ 369{,}666)\quad m_2 = 14$

$[369{,}666;\ 528{,}999)\quad m_3 = 25$

$[528{,}999;\ 688{,}332)\quad m_4 = 18$

$[688{,}332;\ 847{,}665)\quad m_5 = 16$

$[847{,}665;\ 1006{,}998)\quad m_6 = 10$

$[1006{,}998;\ 1166{,}331)\quad m_7 = 5$

$[1166{,}331;\ 1325{,}664)\quad m_8 = 3$

$[1325{,}664; 1485]\quad m_9 = 3$.

Контроль: $\sum_{j=1}^{9} m_j = n \qquad 5 + 14 + 25 + 18 + 16 + 10 + 5 + 3 = 100$.

На основе полученных данных заполняем таблицу:

X	[51; 210,333)	[210,333; 369,666)	[369,666; 528,999)	[528,999; 688,332)	[688,332; 847,665)
m	6	14	25	18	16

X	[847,665; 1006,998)	[1006,998; 1166,331)	[1166,331; 1325,664)	[1325,664; 1485]
m	10	5	3	3

5. Считаем середины подынтервалов $b_1, b_2, \cdots, b_9$ и относительные частоты $\mu_1, \mu_2, \ldots, \mu_9$.

$$b_j = \frac{a_{j-1} + a_j}{2}; \quad \mu_j = \frac{m_j}{N}, \quad j = \overline{1,9};$$

$$b_1 = \frac{a_0 + a_1}{2} = \frac{51 + 210{,}333}{2} = 130{,}666;$$

$$b_2 = \frac{a_1 + a_2}{2} = \frac{210{,}333 + 369{,}666}{2} = 290;$$

$$b_3 = \frac{a_2 + a_3}{2} = \frac{369{,}666 + 528{,}999}{2} = 449{,}3;$$

$$b_4 = \frac{a_3 + a_4}{2} = \frac{528{,}999 + 688{,}332}{2} = 608{,}6;$$

$$b_5 = \frac{a_4 + a_5}{2} = \frac{688{,}332 + 847{,}665}{2} = 768;$$

$$b_6 = \frac{a_5 + a_6}{2} = \frac{847{,}665 + 1006{,}998}{2} = 927{,}3;$$

$$b_7 = \frac{a_6 + a_7}{2} = \frac{1006{,}998 + 1166{,}331}{2} = 1086{,}66;$$

$$b_8 = \frac{a_7 + a_8}{2} = \frac{1166{,}331 + 1325{,}664}{2} = 1246;$$

$$b_9 = \frac{a_8 + a_9}{2} = \frac{1325{,}664 + 1485}{2} = 1405{,}3.$$

Относительные частоты:

$$\mu_1 = \frac{m_1}{n} = \frac{6}{100} = 0{,}06; \qquad \mu_6 = \frac{m_6}{n} = \frac{10}{100} = 0{,}1;$$

$$\mu_2 = \frac{m_2}{n} = \frac{14}{100} = 0{,}14; \qquad \mu_7 = \frac{m_7}{n} = \frac{5}{100} = 0{,}05;$$

$$\mu_3 = \frac{m_3}{n} = \frac{25}{100} = 0{,}25; \qquad \mu_8 = \frac{m_8}{n} = \frac{3}{100} = 0{,}03;$$

$$\mu_4 = \frac{m_4}{n} = \frac{18}{100} = 0{,}18; \qquad \mu_9 = \frac{m_9}{n} = \frac{3}{100} = 0{,}03.$$

$$\mu_5 = \frac{m_5}{n} = \frac{16}{100} = 0{,}16;$$

Контроль:

$$\sum_{j=1}^{9} \mu_j = 1 \qquad 0{,}06 + 0{,}14 + 0{,}25 + 0{,}18 + 0{,}16 + 0{,}1 + 0{,}05 + 0{,}03 + 0{,}03 = 1.$$

В результате имеем таблицу:

Таблица 2.5

X	130,6	290	449,3	608,6	768	927,3	1086,6	1246	1405
μ_i	0,06	0,14	0,25	0,18	0,16	0,1	0,05	0,03	0,03

6. Из данных табл. 2.5 получим эмпирический закон распределения относительных частот и визуальное его представление, то есть строим гистограмму (рис. 2.1) и полигон распределения относительных частот (рис. 2.2).

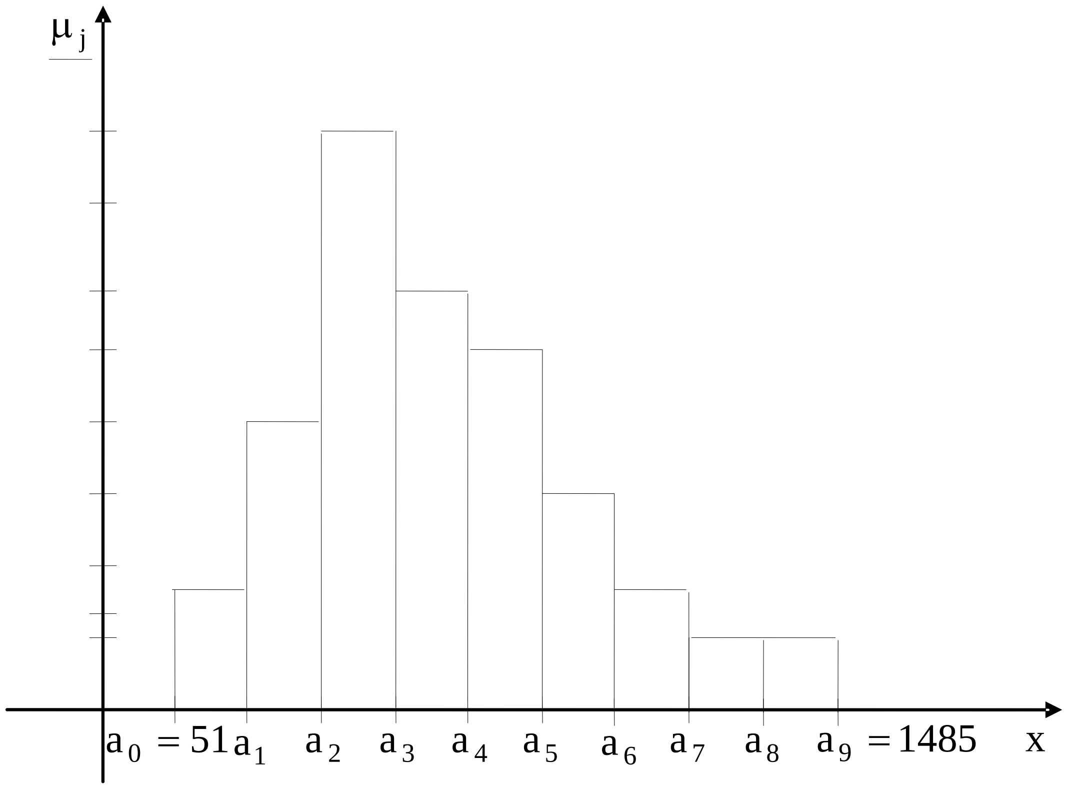

Рис. 2.1

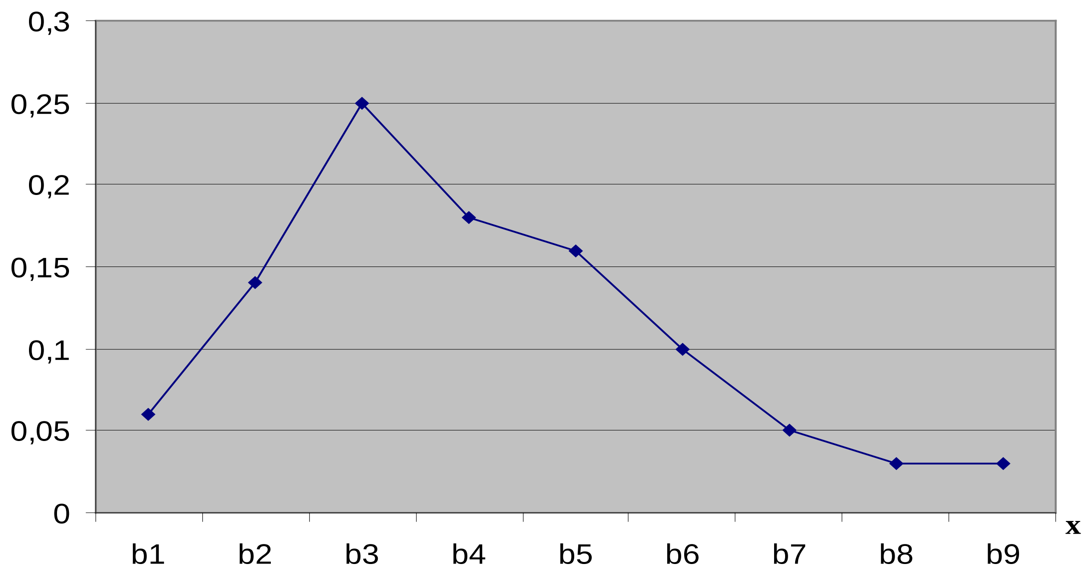

Рис. 2.2

7. Составим эмпирическую функцию распределения. По определению, для группированного статистического ряда $F^*(x)$ имеет вид

$$F^*(x)=\begin{cases} 0 & x \le 51 \\ 0,06 & \text{при } 51 < x \le 210,333 \\ 0,2 & \text{при } 210,333 < x \le 369,666 \\ 0,45 & \text{при } 369,666 < x \le 528,999 \\ 0,63 & \text{при } 528,999 < x \le 688,332 \\ 0,79 & \text{при } 688,332 < x \le 847,665 \\ 0,89 & \text{при } 847,665 < x \le 1006,998 \\ 0,94 & \text{при } 1006,998 < x \le 1166,331 \\ 0,97 & \text{при } 1166,331 < x \le 1325,664 \\ 0,99 & \text{при } 1325,664 < x \le 1485 \\ 1 & \text{при } x > 1485 \end{cases}$$

Построим график $F^*(x)$.

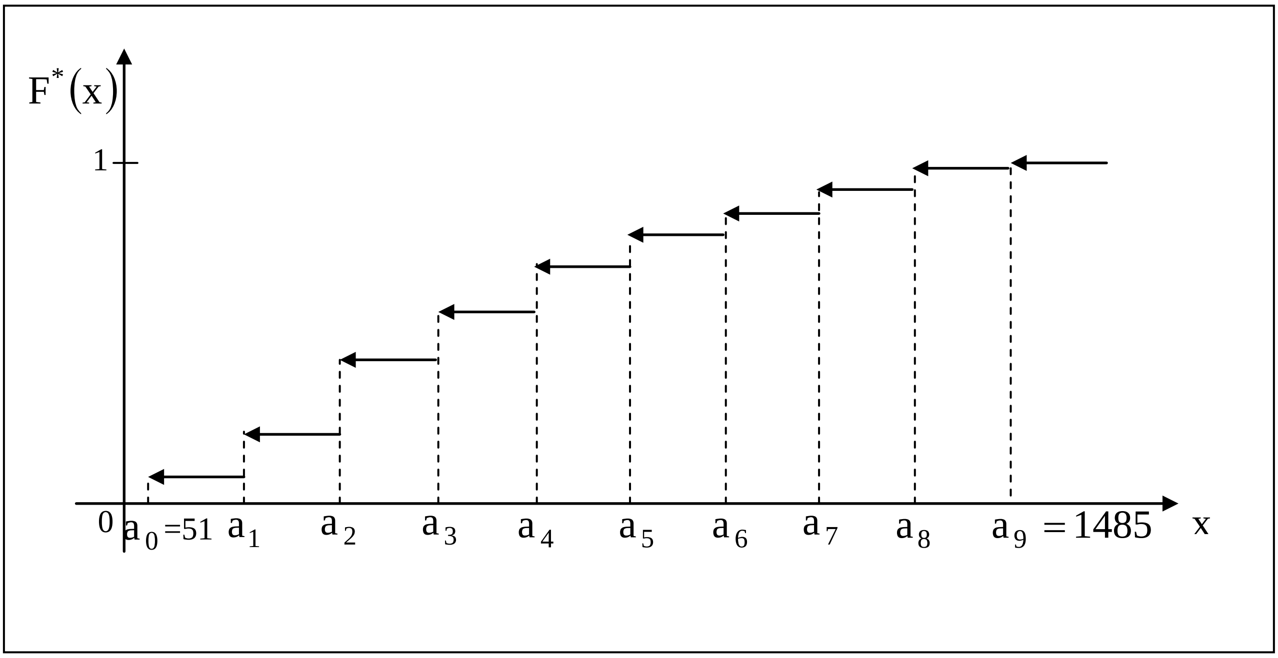

Рис.2.3

2.2. ЛАБОРАТОРНАЯ РАБОТА № 2
«Расчет точечных и интервальных оценок генерального математического ожидания и дисперсии»

Пусть $\theta_г$ – некоторый параметр генеральной совокупности, который невозможно вычислить. Но знать его значение (хотя бы приближенное, оценочное) надо! Поэтому по выборочным данным производят расчет статистических оценок данного генерального параметра.

Точечной называют статистическую оценку генерального параметра $\theta_г$, которая определяется одним числом $\tilde{\theta}$. Точечная оценка $\tilde{\theta}$ может быть несмещенной и смещенной.

Несмещенной называют такую точечную оценку $\tilde{\theta}$, математическое ожидание которой равно оцениваемому генеральному параметру при любом объеме выборки, то есть

$$M\left[\tilde{\theta}\right] = \theta_г. \tag{2.4}$$

Если равенство (2.4) нарушается, то в этом случае точечная оценка $\theta_в$ называется ***смещенной.***

Несмещенной оценкой генеральной средней (математического ожидания генеральной совокупности) служит выборочная средняя:

$$\bar{x}_в = \sum_{j=1}^{k} b_j \cdot \frac{m_j}{n} = \sum_{j=1}^{k} b_j \cdot \mu_j, \tag{2.5}$$

которую считаем по данным табл. 2.3.

Смещенной оценкой генеральной дисперсии $D_г$ служит выборочная дисперсия:

$$D_в = \sum_{j=1}^{k} \left(b_j - \bar{x}_в\right)^2 \cdot \frac{m_j}{n} = \sum_{j=1}^{k} \left(b_j - \bar{x}_в\right)^2 \cdot \mu_j, \tag{2.6}$$

где b_j; μ_j – из табл.2.3. Иногда более удобно пользоваться другой формулой для вычисления выборочной дисперсии:

$$D_в = \sum_{j=1}^{k} b_j^2 \cdot \mu_j - \left(\bar{x}_в\right)^2. \tag{2.6а}$$

<u>Замечание.</u> Поскольку $D_в$ является смещенной оценкой, то ее «исправляют» следующим образом:

$$s^2 = \frac{n \cdot D_в}{n-1}. \tag{2.7}$$

Полученная оценка s^2 – это несмещенная дисперсия, а s – выборочное среднее квадратическое отклонение.

При выборке малого объема точечная оценка может значительно отличаться от оцениваемого генерального параметра, то есть приводит к грубым ошибкам, поэтому при небольшом объеме выборки следует пользоваться интервальными оценками.

Интервальной называют оценку, которая определяется двумя числами – концами интервала, покрывающего оцениваемый генеральный параметр $\theta_г$.

Доверительным называют интервал, который с заданной надежностью (доверительной вероятностью) γ покрывает оцениваемый генеральный параметр, то есть с которой осуществляется неравенство $\left|\theta_г - \tilde{\theta}\right| < \delta$.

Обычно надежность оценки (доверительная вероятность γ) задается. Причем в качестве γ берут число, близкое к единице (0,95; 0,99; 0,999).

Итак, пусть вероятность того, что $\left|\theta_г - \tilde{\theta}\right| < \delta$ равна γ, то есть

$$P\left(\left|\theta_г - \tilde{\theta}\right| < \delta\right) = \gamma, \tag{2.8}$$

или

$$P\left(\tilde{\theta} - \delta < \theta_г < \tilde{\theta} + \delta\right) = \gamma, \tag{2.8a}$$

тогда интервал $\left(\tilde{\theta} - \delta; \tilde{\theta} + \delta\right)$ и есть доверительный интервал.

Для оценки математического ожидания $M_г(x)$ нормально распределенной генеральной совокупности X по выборочной средней $\overline{x}_в$ при известном среднем квадратическом отклонении $\sigma_г = \sqrt{D_г}$ служит доверительный интервал.

$$\overline{x}_в - t\frac{\sigma_г}{\sqrt{n}} < M_г(x) < \overline{x}_в + t\cdot\frac{\sigma_г}{\sqrt{n}}, \tag{2.9}$$

где $t\frac{\sigma_г}{\sqrt{n}} = \delta -$ точность оценки; $n -$ объем выборки; $t -$ это такое значение аргумента функции Лапласа $\Phi(t)$ (прил. 1), при котором $\Phi(t) = \frac{\gamma}{2}$.

Для оценки математического ожидания нормально распределенной генеральной совокупности X по выборочной средней $\overline{x}_в$ при неизвестном среднем квадратическом отклонении $\sigma_г = \sqrt{D_г}$ (при объеме выборки $N > 30$) служит доверительный интервал

$$\overline{x}_в - t_\gamma\cdot\frac{s}{\sqrt{n}} < M_г(x) < \overline{x}_в + t_\gamma\cdot\frac{s}{\sqrt{n}}, \tag{2.10}$$

где t_γ находим по таблице (приложение 2) по заданным n и γ.

Для оценки среднего квадратического отклонения $\sigma_г$ нормально распределенной генеральной совокупности X с доверительной вероятностью служат доверительные интервалы:

$$\sigma_в(1-q) < \sigma_г < \sigma_в(1+q) \quad (\text{при} \quad q<1)$$
$$0 < \sigma_г < \sigma_в(1+q) \quad (\text{при} \quad q>1), \tag{2.11}$$

где $q-$ находим по таблице (приложение 3) при заданных n и γ.

Замечание. Для $M_г(x)$ предлагается построить доверительные интервалы для двух значений вероятности $\gamma = 0{,}95$ и $0{,}99$. Провести анализ, как меняются границы интервалов с увеличением доверительной вероятности.

ПРИМЕР 2.2. Найти точечные и интервальные оценки генерального математического ожидания и генеральной дисперсии, исходя из данных примера 2.1.

1. По данным табл. 2.5 рассчитываем выборочное математическое ожидание и выборочную дисперсию, среднее квадратическое отклонение.

$$\bar{x}_в = 0{,}06\cdot 130{,}666 + 0{,}14\cdot 290 + 0{,}25\cdot 449{,}333 + 0{,}18\cdot 608{,}666 +$$
$$+ 0{,}16\cdot 768 + 0{,}1\cdot 927{,}332 + 0{,}05\cdot 1086{,}66 + 0{,}03\cdot 1246 +$$
$$+ 0{,}03\cdot 1405{,}332 = 619{,}81.$$

$$D_в = (130{,}666-619{,}81)^2\cdot 0{,}06 + (290{,}0-619{,}81)^2\cdot 0{,}14 +$$
$$+ (449{,}333-619{,}81)^2\cdot 0{,}25 + (608{,}666-619{,}81)^2\cdot 0{,}18 +$$
$$+ (768-619{,}81)^2\cdot 0{,}16 + (927{,}332-619{,}81)^2\cdot 0{,}1 +$$
$$+ (1086{,}66-619{,}81)^2\cdot 0{,}05 + (1246-619{,}81)^2\cdot 0{,}03 +$$
$$+ (1405{,}332-619{,}81)^2\cdot 0{,}03 = 91015.$$

$$\sigma_в = \sqrt{D_в} = \sqrt{91015} \approx 301{,}69.$$

По данным табл. 2.4 вычисляем еще одну точечную характеристику - среднее арифметическое значение нашей выборки $\overline{X} = \sum_{i=1}^{n} x_i/n$: $\overline{X} = 615{,}80$.

2. Делаем расчет интервальных оценок, то есть будем строить доверительные интервалы с доверительной вероятностью γ.

а) $\gamma = 0{,}95$; $\sigma_в = 301{,}69$; $\bar{x}_в = 619{,}81$.

Ищем соответствующее значение t_γ по таблице в прил. 1 $t_\gamma = 1{,}96$.

Точность оценки $\delta = \dfrac{1{,}96\cdot 301{,}69}{\sqrt{100}} \approx 59{,}3$. Тогда

$$p(619{,}81-59{,}3 < M_г(x) < 619{,}81+59{,}3) = 0{,}95;$$
$$p(560{,}51 < M_г(x) < 679{,}11) = 0{,}95.$$

б) $\gamma = 0{,}99$; $t_\gamma = 2{,}57$.

$$\delta = \frac{2{,}57 \cdot 301{,}69}{\sqrt{100}} = 77{,}53.$$

$$p(619{,}81 - 77{,}53 < m < 619{,}81 + 77{,}53) = 0{,}99;$$

$$p(542{,}28 < M_г(x) < 697{,}34) = 0{,}99.$$

Строим полученные интервалы на полигоне распределения относительных частот.

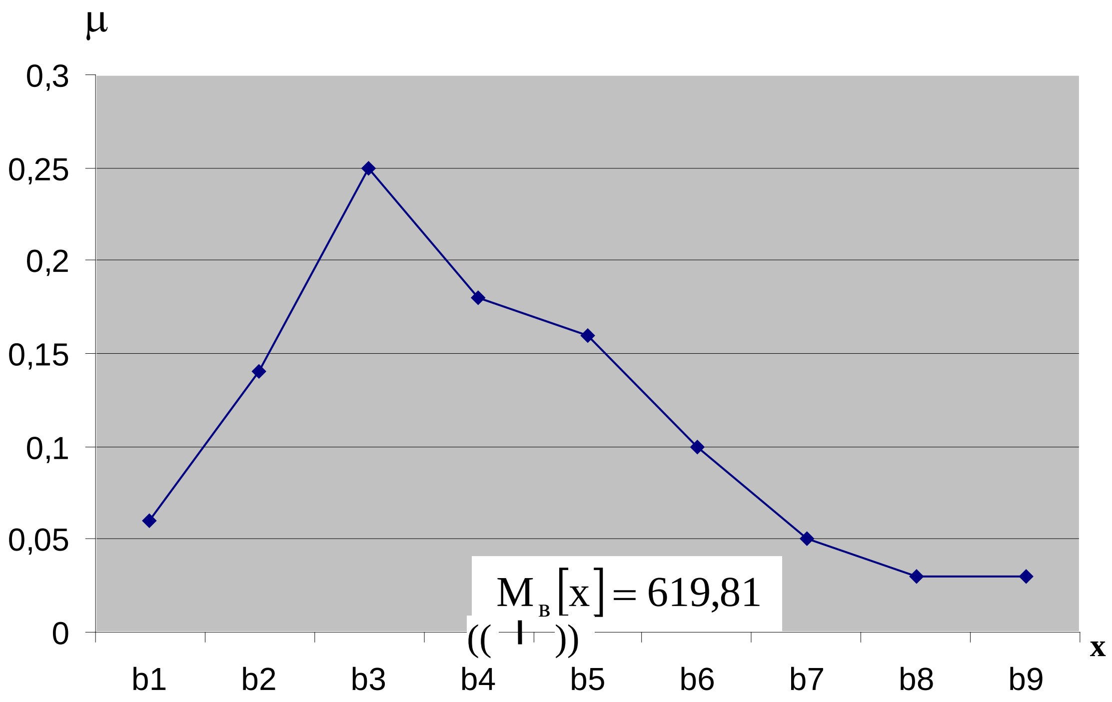

Рис.2.4

2.3. ЛАБОРАТОРНАЯ РАБОТА № 3

«Проверка гипотезы о законе распределения генеральной совокупности»

В лабораторной работе №1 в результате первичной обработки исходных данных получено эмпирическое распределение (табл. 2.3) и по данным этой таблицы построен полигон относительных частот. Относительные частоты иногда называют эмпирическими вероятностями. Из визуального наблюдения полигона можно сделать один из следующих выводов:

Гипотеза $A_о$**:** Генеральная совокупность распределена по нормальному закону;

Гипотеза $B_о$**:** Генеральная совокупность распределена по показательному закону;

Гипотеза C_o**:** Генеральная совокупность распределена по равномерному закону.

Гипотеза A_o

Для того чтобы при заданном уровне значимости α ($\gamma -$ доверительная вероятность, то есть вероятность принять верную гипотезу; $\alpha -$ это вероятность отвергнуть верную гипотезу $(\alpha + \gamma = 1)$) проверить гипотезу о нормальном распределении генеральной совокупности, надо:

1. Вычислить $\overline{x}_в$, $\sigma_в$ (лабораторная работа № 2).

2. Вычислить теоретические вероятности p_j $\left(j = \overline{1,\ k}\right)$. Поскольку плотность распределения для нормального закона есть

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-(x - M[x])^2 / (2D[x])}. \tag{2.12}$$

Тогда

$$p_j = p\left(a_{j-1} < x < a_j\right) = \frac{1}{\sqrt{2\pi}\,\sigma_в} \cdot \int\limits_{a_{j-1}}^{a_j} e^{-(x - \overline{x}_в)^2 / (2D_в)} dx =$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_в} \cdot e^{-\left(b_j - \overline{x}_в\right)^2 / (2D_в)} \cdot \int\limits_{a_{j-1}}^{a_j} dx = \frac{1}{\sqrt{2\pi}\,\sigma_в} \cdot e^{-\left(b_j - \overline{x}_в\right)^2 / (2D_в)} \cdot h, \tag{2.13}$$

$$j = \overline{1, k}$$

где $a_j -$ границы частичных интервалов;

b_j – середина $j-$го частичного интервала;

$h -$длина частичного интервала (см. формулу (2.2)).

3. Составить сводную таблицу на основе данных табл. 2.3 и рассчитанных теоретических вероятностей:

Таблица 2.6

X	b_1	b_2	…	b_j	…	b_k	
μ	μ_1	μ_2	…	μ_j	…	μ_k	← эмпирические вероятности
p	p_1	p_2	…	p_j	…	p_k	← теоретические вероятности

4. Оценку отклонения эмпирических вероятностей μ_j $\left(j = \overline{1, k}\right)$ от теоретических вероятностей p_j $\left(j = \overline{1, k}\right)$ произвести с помощью критерия Пирсона χ^2:

$$\chi^2_{набл.} = \sum_{j=1}^{k} \frac{\left(\mu_j - p_j\right)^2 \cdot n}{p_j}. \tag{2.14}$$

5. По таблице критических точек распределения χ^2 (приложение 4) по заданному уровню значимости α и числу степеней свободы $r = k - s - 1$ ($k -$

количество подынтервалов, s - число параметров распределения) найти критическое значение $\chi^2_{кр.}\ (\alpha, r)$ правосторонней критической области.

<u>Правило 2.1.</u> Если $\chi^2_{набл.} < \chi^2_{кр.}$, тогда нет оснований отвергать гипотезу A_0 о нормальном распределении генеральной совокупности (то есть эмпирические и теоретические частоты различаются незначимо (случайно).

<u>Правило 2.2.</u> Если $\chi^2_{набл.} > \chi^2_{кр.}$, тогда гипотезу A_0 отвергаем.

Гипотеза B_0

Для того чтобы при заданном уровне значимости α проверить гипотезу о показательном распределении генеральной совокупности, надо:

1. Вычислить $\bar{x}_в$ (лабораторная работа № 1). Принять в качестве оценки параметра λ показательного распределения величину, обратную выборочной средней:

$$\lambda = \frac{1}{\bar{x}_в}. \tag{2.15}$$

2. Вычислить теоретические вероятности $p_j \ \left(j = \overline{1, k}\right)$. Поскольку плотность распределения для показательного (экспоненциального) закона есть

$$f(x) = \begin{cases} 0, & \text{при } x < 0, \\ \lambda \cdot e^{-\lambda x}, & \text{при } x \geq 0, \end{cases} \tag{2.16}$$

тогда

$$p_j = p\left(a_{j-1} < x < a_j\right) = \lambda \int_{a_{j-1}}^{a_j} e^{-\lambda \cdot x}\, dx =$$

$$= -e^{-\lambda \cdot x}\Big|_{a_{j-1}}^{a_j} = -\left(e^{-\lambda \cdot a_j} - e^{-\lambda \cdot a_{j-1}} \right) = e^{-\lambda \cdot a_{j-1}} - e^{-\lambda \cdot a_j},$$

где $a_j \ \left(j = 1, \bar{k}\right) -$ границы частичных интервалов;

$\lambda -$ вычисляем по формуле (2.15).

3. Составить сводную таблицу на основе данных табл. 2.3 и рассчитанных теоретических вероятностей (см. табл. 2.6).

4. Оценку отклонения эмпирических вероятностей $\mu_j \left(j = \overline{1, k}\right)$ от теоретических вероятностей $p_j \ \left(j = \overline{1, k}\right)$ произвести с помощью критерия Пирсона χ^2 (формула (2.14)).

5. По таблице критических точек распределения χ^2 по заданному уровню значимости α и по числу степеней свободы $r = k - s - 1$ ($k -$ количество подынтервалов, s - число параметров распределения) найти критическое значение $\chi^2_{кр.}\ (\alpha, r)$правосторонней критической области (см. прил.4).

Далее необходимо проанализировать в соответствии с правилами 2.1 и 2.2 (для предыдущей гипотезы).

Гипотеза C_0

Для того чтобы при заданном уровне значимости α проверить гипотезу о равномерном распределении генеральной совокупности, надо:

1. Оценить параметры a и c – концы интервала, в котором наблюдались возможные значения x, по формулам (через a^* и c^* обозначены оценки параметров):

$$a^* = \overline{x}_в - \sqrt{3}\cdot\sigma_в, \quad c^* = \overline{x}_в + \sqrt{3}\cdot\sigma_в. \tag{2.17}$$

2. Вычислить теоретические частоты $p_j \left(j = \overline{1,k}\right)$. Поскольку плотность распределения для равномерного закона есть

$$f(x) = \frac{1}{c^* - a^*}, \tag{2.18}$$

тогда

$$p_j = p\left(a_{j-1} < x < a_j\right) = \int\limits_{a_{j-1}}^{a_j} \frac{1}{c^* - a^*}dx = \frac{1}{c^* - a^*}\cdot\int\limits_{a_{j-1}}^{a_j}dx = \frac{1}{c^* - a^*}\cdot h, \tag{2.19}$$

где $a_j \left(j = \overline{1,k}\right)$– границы частичных интервалов;

h – длина частичных интервалов.

Получаем, что все $p_j \left(j = \overline{1,k}\right)$ равны одному числу $\dfrac{h}{c^* - a^*}$.

3. Составить сводную таблицу на основе эмпирических вероятностей и рассчитанных теоретических вероятностей (см. табл. 2.6).

4. Оценку отклонения эмпирических вероятностей $\mu_j \left(j = \overline{1,k}\right)$ от теоретических вероятностей $p_j \left(j = \overline{1,k}\right)$ произвести с помощью критерия Пирсона χ^2 (формула (2.14)).

5. По таблице критических точек распределения χ^2 по заданному уровню значимости α и по числу степеней свободы $r = k - s - 1$ (k – количество подынтервалов, s - число параметров распределений) найти критические значения $\chi^2_{кр.}(\alpha, r)$ правосторонней критической области.

Далее анализируем в соответствии с правилами 2.1 и 2.2 (см. гипотезу А).

Замечание 1. После составления табл.5 необходимо сделать на одном рисунке два графика: ломаную эмпирических вероятностей и кривую теоретических вероятностей.

Замечание 2. Здесь же на этом рисунке рекомендуется нанести:

а) $\overline{x}_в$; б) доверительный интервал, построенный для одной из доверительных вероятностей γ (например, для $\lambda = 0{,}95$); в) интервал, построенный по правилу «3-х сигм».

Замечание 3. Данный рисунок является наглядным результатом работы, проделанной в лабораторных работах 1,2,3.

ПРИМЕР 2.3. Используя критерий согласия Пирсона при уровне значимости $\alpha = 0{,}01$, проверить гипотезу о нормальном распределении генеральной совокупности.

1. Переносим из лабораторной работы № 1 полигон (рис.2.2) распределения относительных частот, табл. 2.5, из лабораторной работы №2 (пример 2.2) $\overline{x}_в$ =619,81, $\sigma_в$ =301,69.

X	130,6	290	449,3	608,6	768	927,3	1086,6	1246	1405
μ	0,06	0,14	0,25	0,18	0,16	0,1	0,05	0,03	0,03

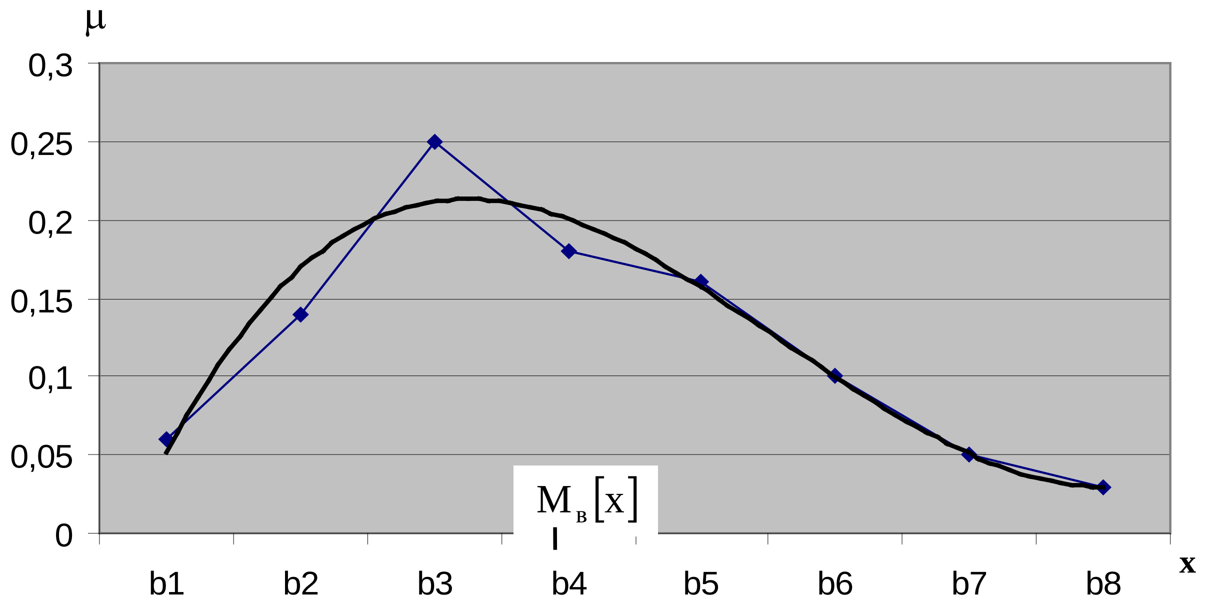

Рис. 2.5

2. Из визуального наблюдения ломаной делаем предположение (ставим гипотезу) о законе распределения генеральной совокупности, то есть ставим гипотезу H_0: выборка распределена по нормальному закону.

3. Вычислим теоретические вероятности p_j. Для этого записываем функцию плотности $f(x)$ для нормального закона:

$$f(x)=\frac{1}{\sqrt{2\pi}\cdot\sigma_в}\cdot e^{-\frac{(x-\bar{x}_в)^2}{2\sigma_в^2}}.$$

Соответственно $f(x)=\frac{1}{\sqrt{2\cdot 3{,}14}\cdot 301{,}69}\cdot e^{-\frac{(x-619{,}81)^2}{2\cdot(301{,}69)^2}}$ или

$f(x)=0{,}001\cdot e^{-\frac{(x-619{,}81)^2}{2\cdot(301{,}69)^2}}$.

Тогда $p_j=\frac{1}{\sqrt{2\pi}\cdot 301{,}69}\cdot e^{-\frac{(b_j-619{,}81)^2}{2\cdot(301{,}69)^2}}\cdot 159{,}333.$

$$p_1=0{,}003\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(130{,}666-619{,}81)^2}{2\cdot(301{,}69)^2}}\cdot 159{,}333=$$

$$=0{,}478\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(1{,}62)^2}{2}}=0{,}478\cdot 0{,}1109=0{,}05;$$

$$p_2=0{,}478\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(290-619{,}81)^2}{2\cdot(301{,}69)^2}}=0{,}478\cdot 0{,}2203=0{,}105;$$

$$p_3=0{,}478\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(449{,}333-619{,}819)^2}{2\cdot 301{,}69^2}}=0{,}478\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{(0{,}565)^2}{2}}=$$

$$=0{,}478\cdot 0{,}3391=0{,}162;$$

$$p_4=0{,}478\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(608{,}666-619{,}819)^2}{2\cdot 301{,}69^2}}=0{,}478\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{(0{,}037)^2}{2}}=$$

$$=0{,}478\cdot 0{,}3986=0{,}19;$$

$$p_5=0{,}478\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(768-619{,}819)^2}{2\cdot 301{,}69^2}}=0{,}169;$$

$$p_6=0{,}478\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(927{,}332-619{,}819)^2}{2\cdot 301{,}69^2}}=0{,}11;$$

$$p_7=0{,}478\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(1086{,}66-619{,}819)^2}{2\cdot 301{,}69^2}}=0{,}06;$$

$$p_8 = 0,478 \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(1246-619,819)^2}{2\cdot 301,69^2}} = 0,02;$$

$$p_9 = 0,478 \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(1405,332-619,819)^2}{2\cdot 301,69^2}} = 0,007.$$

4. Составляем табл. 2.7 распределения теоретических вероятностей.

Таблица 2.7

X	130,6	290	449,3	608,6	768	927,3	1086,6	1246	1405
μ	0,06	0,14	0,25	0,18	0,16	0,1	0,05	0,03	0,03
p	0,05	0,105	0,162	0,19	0,17	0,11	0,06	0,02	0,007

Отметим теоретические вероятности на полигоне относительных частот.

5. Рассчитаем значение критерия $\chi^2_{набл.}$ (см. формулу (2.14)):

$$\chi^2_{набл.} = \frac{(0,06-0,05)^2}{0,05}\cdot 100 + \frac{(0,14-0,11)^2}{0,11}\cdot 100 + \frac{(0,25-0,16)^2}{0,16}\cdot 100 +$$

$$+\frac{(0,18-0,19)^2}{0,19}\cdot 100 + \frac{(0,16-0,17)^2}{0,17}\cdot 100 + \frac{(0,1-0,11)^2}{0,11}\cdot 100 +$$

$$+\frac{(0,03-0,03)^2}{0,02}\cdot 100 + \frac{(0,03-0,007)^2}{0,007}\cdot 100 = 0,2+0,8+5,06+0,05+$$

$$+0,06+0,09+0,17+0,5+7,56 = 14,48.$$

6. Из таблицы «Критические точки распределения χ^2» (прил. 4) находим соответствующее нашим значениям $\chi^2_{кр.}(\alpha, r)$.

$$\alpha = 0,01;\ r = k-3$$

$$k = 9 \quad \Rightarrow \ r = 6.\ \ \chi^2_{кр.}(0,01;6) = 16,8.$$

Сравниваем $\chi^2_{набл.}$ и $\chi^2_{кр.}$. Так как $\chi^2_{набл.} < \chi^2_{кр.}$ $(14,48 < 16,8)$, то гипотеза A_0 (выборка распределена по нормальному закону) принимается по правилу 2.1.

2.4 ЛАБОРАТОРНАЯ РАБОТА № 4

«Расчет параметров корреляционной зависимости. Вывод линейной зависимости».

Две случайные величины могут быть связаны функциональной зависимостью, либо зависимостью другого рода, называемой статистической, либо быть независимыми. Строгая функциональная зависимость для случайных величин реализуется редко, так как обе величины (или одна из них) подвержены различным случайным факторам (даже если среди этих факторов есть общие).

Статистической называют зависимость, при которой изменение одной из величин влечет изменение распределения другой. Если при изменении одной из величин изменяется среднее значение другой, то в этом случае статистическую зависимость называют ***корреляционной.*** Значит, корреляционная зависимость есть частный случай статистической зависимости.

Чтобы установить наличие и характер статистической связи между двумя случайными величинами X и Y, нужно привести к удобному виду исходный цифровой материал. Наглядной (удобной) формой представления данных является корреляционная таблица:

Таблица 2.8

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$	y_ℓ	m_x
x_1	m_{11}	m_{12}	$\cdots$	m_{1j}	$\cdots$	$m_{1\ell}$	m_{x1}
x_2	m_{21}	m_{22}	$\cdots$	m_{2j}	$\cdots$	$m_{2\ell}$	m_{x2}
…	…	…	…	…	…	…	…
x_i	m_{i1}	m_{i2}	$\cdots$	m_{ij}	$\cdots$	$m_{i\ell}$	m_{xi}
…	…	…	…	…	…	…	…
x_k	m_{k1}	m_{k2}	$\cdots$	m_{kj}	$\cdots$	$m_{k\ell}$	m_{xk}
m_y	m_{y1}	m_{y2}	$\cdots$	m_{yj}	$\cdots$	$m_{y\ell}$	m

Здесь $x_1, x_2, \ldots, x_k; y_1, y_2, \ldots, y_\ell$ – середины подинтервалов сгруппированных выборок X и Y (см. лабораторную работу №1); m_{ij} – частота, с которой встречается пара $(x_i; y_j)$. В последнем столбце и в последней строке таблицы помещены суммарные частоты, соответствующие значению $X = x_i$ и, соответственно, $Y = y_j$, то есть

$m_{xi} = m_{i1} + m_{i2} + \ldots + m_{i\ell}$;

$m_{yj} = m_{1j} + m_{2j} + \ldots + m_{kj}$, тогда должно быть

$$m_{x_1} + m_{x_2} + \ldots + m_{x_k} = \sum_{i=1}^{k} m_{x_i} = m \text{ и}$$

$$m_{y_1} + m_{y_2} + \ldots + m_{y_j} + \ldots + m_{y_\ell} = \sum_{j=1}^{\ell} m_{y_j} = m;$$

$m -$ общее количество пар значений $(x_i; y_j)$.

Каждая $i -$я строка табл. 2.8 представляет собой совместно с первой строкой некоторое распределение случайной величины Y, соответствующей данному значению случайной величины $X = x_i$. Такое распределение называется ***условным распределением***. Последняя строка табл. 2.8 совместно с первой строкой образует безусловное распределение случайной величины Y (ее эмпирический закон распределения):

Таблица 2.9

Y	y_1	y_2	$\cdots$	y_j	$\cdots$	y_ℓ
m_y	m_{y_1}	m_{y_2}	$\cdots$	m_{y_j}	$\cdots$	m_{y_ℓ}

Каждый $j -$й столбец табл. 2.8 представляет собой совместно с первым столбцом некоторое распределение случайной величины X, соответствующее данному значению случайной величины $Y = y_j$ (то есть условное распределение). Последний столбец табл. 2.8 совместно с первым столбцом образует безусловное распределение случайной величины X (ее эмпирический закон распределения):

Таблица 2.10

X	x_1	x_2	$\cdots$	x_i	$\cdots$	x_k
m_x	m_{x_1}	m_{x_2}	$\cdots$	m_{x_i}	$\cdots$	m_{x_k}

По данным табл. 2.9 и. 2.10 вычисляем средние значения:

$$\overline{Y} = \left(\sum_{j=1}^{\ell} y_j \cdot m_{yj} \right) \Big/ m; \quad \overline{X} = \left(\sum_{i=1}^{k} x_i \cdot m_{xi} \right) \Big/ m \qquad (2.20)$$

и средние квадратические отклонения:

$$\sigma_Y^2 = \left(\sum_{j=1}^{\ell} \left(y_j - \overline{Y} \right)^2 \cdot m_{yj} \right) \Big/ m; \ \sigma_X^2 = \left(\sum_{i=1}^{k} \left(x_i - \overline{X} \right)^2 \cdot m_{xi} \right) \Big/ m. \qquad (2.21)$$

<u>*Замечание.*</u> Рекомендуется сделать два рисунка – это графические изображения эмпирических законов распределения случайных величин X и Y в

виде полигонов распределения частот. На рисунках нанести средние значения $\overline{X}$ и $\overline{Y}$.

Уточним определение корреляционной зависимости. Для этого введем понятие условной средней. Для каждой $i-$й строки табл. 2.8 (совместно с первой строкой) можно вычислить среднее значение случайной величины Y (по формуле 2.20), которое называется ***условным средним***

$\overline{Y}_{x_i} = \overline{Y}(X = x_i)$ $(i = 1,2,\ldots,k)$.

Так как каждому значению x_i соответствует одно значение условной средней, то очевидно, условная средняя $\overline{Y}_{x_i}$ есть функция от X. В этом случае говорят, что случайная величина Y зависит от X корреляционно.

Корреляционной зависимостью Y от X называют функциональную зависимость условной средней $\overline{Y}_x$ от X:

$$\overline{Y}_x = f(X). \qquad (2.22)$$

Уравнение (2.22) называется уравнением ***регрессии*** Y на X; функция $f(x)$ называется ***регрессией*** Y на X; график функции $f(x)$ - ***линией регрессии*** Y на X.

Аналогично, для каждого $j-$го столбца табл. 2.8 (совместно с первым столбцом) можно вычислить среднее значение случайной величины X по формуле (2.20), которое называется ***условным средним***

$$\overline{X}_{y_j} = \overline{X}(Y = y_j) \ (j = 1,2,\ldots,\ell).$$

Тогда корреляционной зависимостью X от Y называется функциональная зависимость средней $\overline{X}_{y_i}$ от Y:

$$\overline{X}_y = \varphi(Y), \qquad (2.23)$$

уравнение (2.23) называется ***уравнением регрессии*** X на Y; функция $\varphi(Y)-$ ***регрессией*** X на Y; график функции $\varphi(Y)-$***линией регрессии*** X на Y.

Замечание. Рассматриваемые два уравнения регрессии существенно различны и не могут быть получены одно из другого.

Изучение корреляционной связи будем проводить при решении двух основных задач:

- определение формы корреляционной связи, то есть вида теоретической функции регрессии (она может быть линейной и нелинейной);
- определение тесноты (силы) корреляционной связи.

Наиболее простой и важный случай корреляционной зависимости – линейная регрессия. В этом случае теоретическое уравнение линейной регрессии Y на X (формула 2.22) имеет вид

$$\overline{Y}_x = aX + b. \qquad (2.24)$$

Коэффициент a в уравнении (2.24) называют ***коэффициентом регрессии*** X на Y и обозначают $\rho_{YX}\left(a=\rho_{YX}\right)$. Оценки неизвестных параметров ρ_{YX} и b рассчитаем применяя данные табл. 2.8:

$$a=\rho_{YX}=\frac{\left(\sum\limits_{i=1}^{k}\ \sum\limits_{j=1}^{\ell} m_{ij}\cdot x_i\cdot y_j\right)\Big/ m-\overline{X}\cdot\overline{Y}}{\left(\sum\limits_{i=1}^{k} x_i^2\cdot m_{x_i}\right)\Big/ m-\left(\overline{X}\right)^2}. \quad (2.25)$$

$$b=\overline{Y}-\rho_{YX}\cdot\overline{X}, \quad (2.26)$$

где $\overline{X}$ и $\overline{Y}$ - средние значения случайных величин X и Y, вычисленные по формулам (2.20).

Сделаем графическое изображение так называемой эмпирической линии регрессии Y на X и теоретической линии регрессии Y на X. Для этого в декартовой системе координат по оси $0X$ откладываем значения $x_1, x_2, \ldots, x_k$ из табл. 2.8, по оси $0Y$ откладываем значения условных средних $\overline{Y_{x_i}}$. Тогда ломаная, соединяющая точки $\left(x_1, \overline{Y}_{x_1}\right); \left(x_2, \overline{Y}_{x_2}\right); \ldots \left(x_k, \overline{Y}_{x_k}\right)$, и будет эмпирической линией регрессии Y на X. Здесь же на данном графике строим теоретическую линию регрессии, то есть прямую $\overline{Y}_x=\rho_{YX}\cdot X+b$ с вычисленными коэффициентами.

<u>*Замечание .*</u> Поскольку формулы (2.25) и (2.26) получены по методу наименьших квадратов, то по сути этого метода теоретическая линия регрессии должна на графике быть в «середине» ломаной линии.

Аналогично можно поставить вопрос о нахождении теоретического уравнения линейной регрессии X на Y (формула 2.23), которое имеет вид

$$\overline{X}_y=a_1 Y+b_1. \quad (2.27)$$

Коэффициент a_1 в уравнении (2.27) называется ***коэффициентом регрессии*** X на Y и обозначается $\rho_{XY}\left(a_1=\rho_{XY}\right)$. Оценки неизвестных параметров ρ_{XY} и b_1 рассчитываются по данным табл. 2.7:

$$a_1=\rho_{XY}=\frac{\left(\sum\limits_{i=1}^{k}\ \sum\limits_{j=1}^{\ell} m_{ij}\, x_i\cdot y_j\right)\Big/ m-\overline{X}\,\overline{Y}}{\left(\sum\limits_{j=1}^{\ell} y_j^2\cdot m_{y_j}\right)\Big/ m-\left(\overline{Y}\right)^2}; \quad (2.28)$$

$$b_1=\overline{X}-\rho_{YX}\cdot\overline{Y}, \quad (2.29)$$

где $\overline{X}$ и $\overline{Y}$ – средние значения случайных величин X и Y, вычисленные по формулам (2.20).

Далее целесообразно сделать графическое изображение эмпирической и теоретической линий регрессии X на Y аналогично вышеизложенному.

В случае линейной регрессии задача определения тесноты связи сводится к вычислению эмпирического (выборочного) коэффициента корреляции, который можно вычислить по одной из формул:

$$r_B = \rho_{YX} \cdot \frac{\sigma_x}{\sigma_Y} \quad \text{или} \quad r_B = \rho_{XY} \cdot \frac{\sigma_Y}{\sigma_x}, \tag{2.30}$$

где σ_X, σ_Y – значения средних квадратических отклонений, вычисленных по формуле (2.21).

Приведем свойства выборочного коэффициента корреляции:

1. $|r_B| \le 1$ или $-1 \le r_B \le 1$.

2. Если $r_B = 0$, тогда X и Y не связаны линейной корреляционной зависимостью (но могут быть связаны нелинейной корреляционной или даже функциональной зависимостью).

3. С возрастанием абсолютной величины выборочного коэффициента корреляции линейная корреляционная зависимость становится более тесной и при $|r_B| = 1$ переходит в функциональную зависимость.

4. Если $r_B = +1$ $(r_B = -1)$, тогда X и Y связаны прямой (обратной) линейной зависимостью.

Замечание. Однако эмпирический коэффициент корреляции является весьма условным показателем даже линейной связи, так как он является средней пропорциональной величиной между коэффициентами регрессии. В теории корреляции существует понятие корреляционного отношения, которое является более естественным и общим показателем степени тесноты связи, так как не связано с формой зависимости. Но в тему данного практикума корреляционные отношения не входят.

ПРИМЕР 2.4. Дана таблица распределения заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т).

X \ Y	6	12	18	24	30	m_x
10	4	-	-	-	-	4
15	2	6	-	-	-	8
20	-	2	5	2	-	9
25	-	-	40	8	4	52
30	-	-	5	7	7	19
35	-	-	-	-	8	8
m_y	6	8	50	17	19	$m = 100$

Находим эмпирические распределения каждой из компонент, их графические изображения, средние значения и средние квадратические отклонения (2.20) – (2.21).

X	10	15	20	25	30	35
$\mu_X = \dfrac{m_x}{m}$	$\dfrac{4}{100}$	$\dfrac{8}{100}$	$\dfrac{9}{100}$	$\dfrac{52}{100}$	$\dfrac{19}{100}$	$\dfrac{8}{100}$

Y	6	12	18	24	30
$\mu_y = \dfrac{m_y}{m}$	$\dfrac{6}{100}$	$\dfrac{8}{100}$	$\dfrac{50}{100}$	$\dfrac{17}{100}$	$\dfrac{19}{100}$

$$\overline{X} = 10\cdot\frac{4}{100} + 15\cdot\frac{8}{100} + 20\cdot\frac{9}{100} + 25\cdot\frac{52}{100} + 30\cdot\frac{19}{100} + 35\cdot\frac{8}{100} =$$
$$= (10\cdot 4 + 15\cdot 8 + 20\cdot 9 + 25\cdot 52 + 30\cdot 19 + 35\cdot 8)/100 = 24{,}9;$$

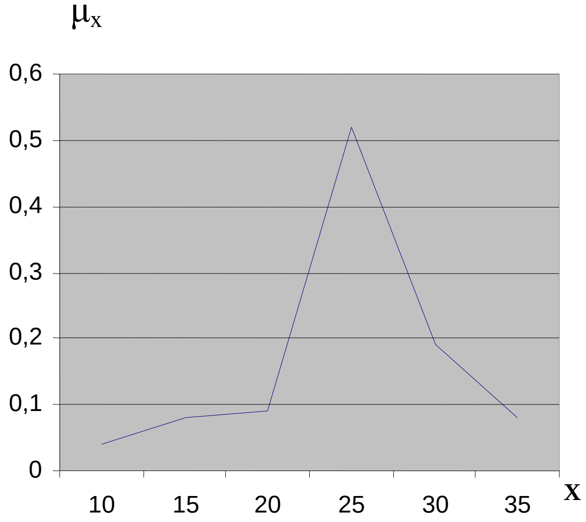

Рис. 2.6

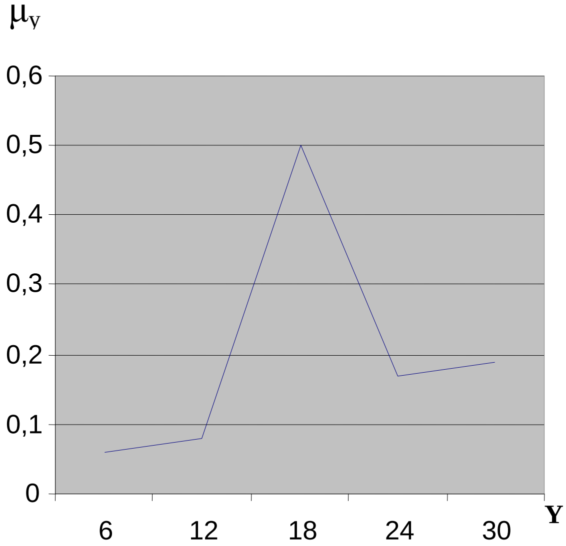

Рис. 2.7

$$\overline{Y} = 6\cdot\frac{6}{100} + 12\cdot\frac{8}{100} + 18\cdot\frac{50}{100} + 24\cdot\frac{17}{100} + 30\cdot\frac{19}{100} =$$
$$= (6\cdot 6 + 12\cdot 8 + 18\cdot 50 + 24\cdot 17 + 30\cdot 19)/100 = 20{,}1.$$
$$\sigma_X^2 = \left((10-24{,}9)^2\cdot 4 + (15-24{,}9)^2\cdot 8 + (20-24{,}9)^2\cdot 9 +\right.$$
$$\left. + (25-24{,}9)^2\cdot 52 + (30-24{,}9)^2\cdot 19 + (35-24{,}9)^2\cdot 8\right)/100 = 31{,}99$$
$$\sigma_Y^2 = \left((6-20{,}1)^2\cdot 6 + (12-20{,}1)^2\cdot 8 + (18-20{,}1)^2\cdot 5 +\right.$$
$$\left. + (24-20{,}1)^2\cdot 17 + (30-20{,}1)^2\cdot 19\right)/100 = 40{,}6;$$
$$\sigma_X = \sqrt{31{,}99} = 5{,}66; \qquad \sigma_Y = \sqrt{40{,}6} = 6{,}37.$$

Перейдем к расчету данных для построения эмпирических линий регрессии Y на X и X на Y, то есть к расчету условных средних $\overline{Y}(X = x_i) = \left(\sum_{j=1}^{\ell} y_i\, m_{i\,j}\right) \Big/ m_{x_i}$.

$\overline{Y}(X=10)=6\cdot 4/4=6;\qquad \overline{Y}(X=15)=(6\cdot 2+12\cdot 6)/8=10{,}5;$

$\overline{Y}(X=20)=(12\cdot 2+18\cdot 5+24\cdot 2)/9=18;$

$\overline{Y}(X=25)=(18\cdot 40+24\cdot 8+30\cdot 4)/52=19{,}846;$

$\overline{Y}(X=30)=(18\cdot 5+24\cdot 7+30\cdot 7)/19=24{,}632;$

$\overline{Y}(X=35)=30\cdot 8/8=30,$

и условных средних $\overline{X}(Y=y_j)=\left(\sum_{i=1}^{k} x_i\, m_{ij}\right)\Big/ m_{y_j}$

$\overline{X}(Y=6)=(10\cdot 4+15\cdot 2)/6=11{,}7;$

$\overline{X}(Y=12)=(15\cdot 6+20\cdot 2)/8=16{,}25;$

$\overline{X}(Y=18)=(20\cdot 5+25\cdot 40+30\cdot 5)/50=25;$

$\overline{X}(Y=24)=(20\cdot 2+25\cdot 8+30\cdot 7)/17=26{,}5;$

$\overline{X}(Y=30)=(25\cdot 4+30\cdot 7+35\cdot 8)/19=31{,}1.$

Сделаем графическое изображение эмпирических линий регрессии:

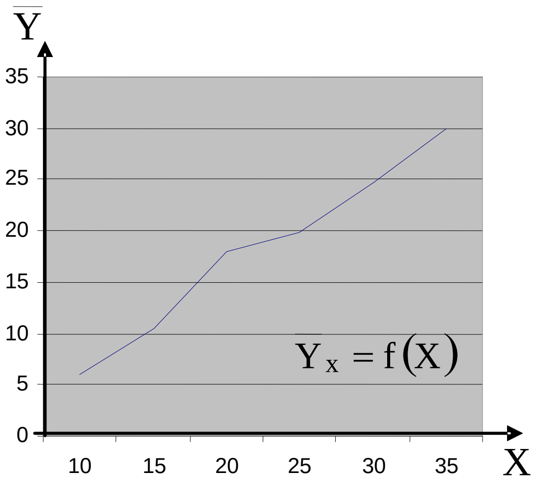

Рис. 2.8

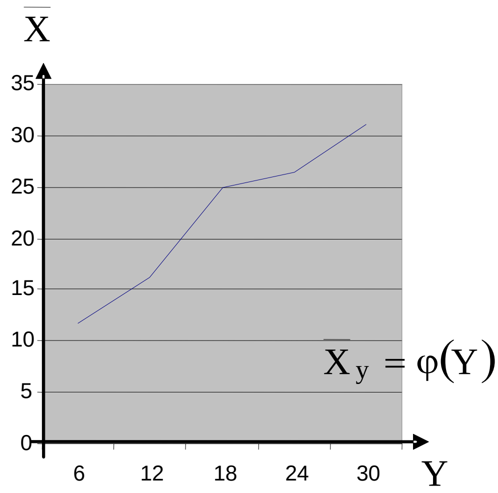

Рис. 2.9

Перейдем к расчету параметров для теоретических линий регрессии $\overline{Y}_x = a\,X + b$ и $\overline{X}_y = a_1 Y + b_1$. Прежде чем воспользоваться формулами (2.25) и (2.28), подготовим их общее значение выражения:

$$C=\sum_{i=1}^{k}\ \sum_{j=1}^{\ell}\left(x_i\, y_j\, m_{ij}\right)/m=$$

$$=(10\cdot 6\cdot 4+(15\cdot 6\cdot 2+15\cdot 12\cdot 6)+(20\cdot 12\cdot 2+20\cdot 18\cdot 5+20\cdot 24\cdot 2)+$$
$$+(25\cdot 18\cdot 40+25\cdot 24\cdot 8+25\cdot 30\cdot 4)+(30\cdot 18\cdot 5+30\cdot 24\cdot 7+30\cdot 30\cdot 7)+$$
$$+35\cdot 30\cdot 8)/100=(240+1260+3240+25800+14040+$$
$$+8400)/100=529{,}8.$$

Знаменатели формул (2.25) и (2.28) есть, соответственно, σ_X^2 и σ_Y^2, которые у нас уже рассчитаны, поэтому:

$$a=\rho_{YX}=\frac{529{,}8-24{,}9\cdot 20{,}1}{31{,}99}=\frac{529{,}8-500{,}49}{31{,}99}=\frac{29{,}31}{31{,}99}=0{,}916;$$

$$b = 20{,}1 - 0{,}916 \cdot 24{,}9 = 20{,}1 - 22{,}81 = -2{,}71;$$

$$a_1 = \rho_{XY} = \frac{529{,}8 - 24{,}9 \cdot 20{,}1}{40{,}6} = \frac{529{,}8 - 500{,}49}{40{,}6} = \frac{29{,}31}{40{,}6} = 0{,}722;$$

$$b_1 = 24{,}9 - 0{,}722 \cdot 20{,}1 = 24{,}9 - 14{,}51 = 10{,}39.$$

Тогда $\overline{Y} = 0{,}916\,X - 2{,}71$ и $\overline{X} = 0{,}722\ Y + 10{,}39$. Для графического изображения полученных прямых линий вновь вернемся к рисункам 2.3 и 2.4, чтобы совместить на одном графике эмпирическую и соответствующую ей теоретическую линии регрессии.

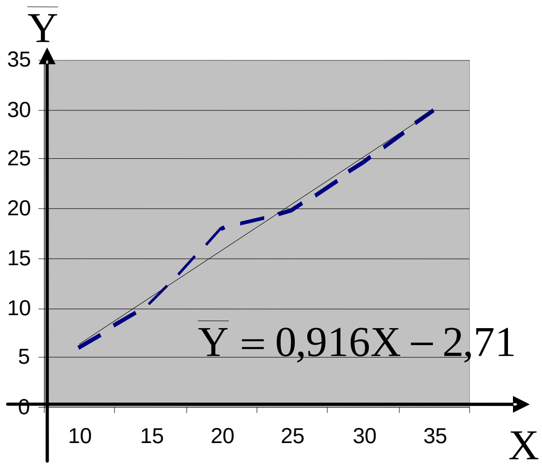

Рис. 2.10

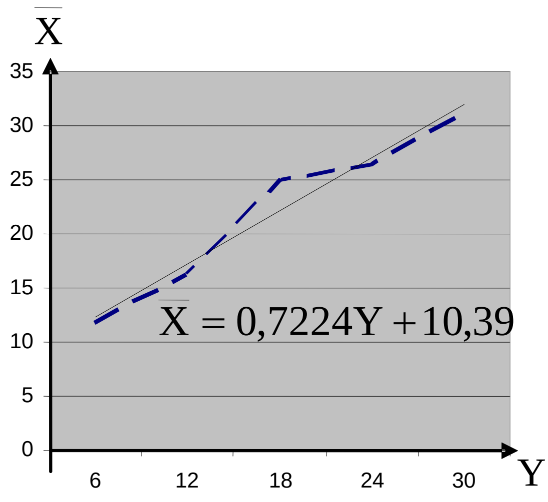

Рис. 2.11

Перейдем к вычислению эмпирического (выборочного) коэффициента корреляции по формуле (2.30):

$$r_B = \rho_{YX} \cdot \frac{\sigma_x}{\sigma_y} = \frac{0{,}916 \cdot 5{,}66}{6{,}37} \approx 0{,}81.$$

Итак: 1. Поскольку r_B достаточно близок к единице, то можно утверждать, что между случайными величинами X и Y существует линейная зависимость, причем достаточно тесная.

2. Поскольку $r_B > 0$, то эта линейная зависимость прямая.

Данные выводы можно сделать и из рис. 2.10, 2.11.

3. МАТЕРИАЛЫ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

3.1. ЛАБОРАТОРНЫЕ РАБОТЫ

ЗАДАЧА 1

Дана выборка из генеральной совокупности объема n. По выборке необходимо выполнить следующие три лабораторные работы.

Лабораторная работа №1. Первичная обработка статистических данных

1. Построить вариационный ряд.
2. Построить группированную выборку с числом интервалов $k = 5 \div 9$.
3. Построить гистограмму и полигон частот и (или) относительных частот.
4. Найти эмпирическую функцию распределения и построить ее график.

Лабораторная работа №2. Расчет точечных и интервальных оценок генерального математического ожидания и дисперсии.

1. По сгруппированной выборке вычислить точечные оценки математического ожидания и среднеквадратического отклонения.
2. Построить доверительные интервалы для генерального математического ожидания с доверительными вероятностями $\gamma_1 = 0{,}95$ и $\gamma_2 = 0{,}99$.

Лабораторная работа №3. Проверка гипотезы о законе распределения генеральной совокупности.

1. Перенести из Л.Р. №1 график полигона относительных частот.
2. Из визуального наблюдения полигона выбрать один из законов распределения (равномерный, нормальный, показательный) в качестве предполагаемого (теоретического) распределения.
3. Найти параметры теоретического распределения.
4. Построить на одном графике полигон относительных частот (выборочное распределение) и кривую теоретического распределения генеральной выборки.
5. Проверить гипотезу о том, что выборка имеет выбранное теоретическое распределение. Принять уровень значимости $\alpha = 0{,}01$.

Вариант 1. $n = 100; \;\; k = 9$

1.70	1.51	1.35	1.11	1.45	1.94	1.15	1.79	1.35	1.49
1.52	1.97	1.27	1.43	1.10	1.38	1.90	1.07	1.29	1.61
1.29	1.29	1.59	1.32	1.00	1.33	1.29	1.17	1.09	1.50
1.58	1.32	1.70	1.17	1.89	1.97	1.40	1.19	1.24	1.06
1.25	1.41	1.33	1.87	1.13	1.35	1.57	1.48	1.94	1.54
1.03	1.91	1.74	1.56	1.66	1.11	1.56	1.73	1.54	1.10
1.94	1.69	1.78	1.65	1.00	1.37	1.08	1.79	1.72	1.68
1.39	1.45	1.10	1.40	1.90	1.61	1.42	1.63	1.51	1.51
1.82	1.78	1.24	1.78	1.11	1.86	1.73	1.70	1.36	1.92
1.59	1.47	1.29	1.04	1.20	1.90	1.40	1.71	1.04	1.32

Вариант 2. $n = 100; \;\; k = 7$

0.64	0.39	0.00	0.56	0.05	1.07	0.02	0.01	0.00	0.31
0.06	0.02	0.35	0.13	0.14	0.04	0.42	0.22	0.44	0.08
0.54	1.13	0.47	0.20	0.09	0.25	0.14	0.20	0.01	0.60
0.31	0.36	0.28	0.36	0.91	0.17	0.49	0.08	0.46	0.29
0.63	0.12	0.45	0.07	0.43	0.51	0.84	0.83	0.28	0.37
0.17	0.07	0.04	0.31	0.40	0.12	0.20	0.21	0.60	0.09
0.06	0.24	0.43	0.32	0.16	0.31	0.29	0.60	0.77	0.76
0.22	0.26	0.09	0.18	0.24	0.17	0.10	0.33	0.24	0.44
0.03	0.43	0.08	0.96	0.40	0.22	0.01	0.03	0.94	0.14
0.30	0.27	0.05	0.24	0.03	0.15	0.52	0.16	0.48	0.38

Вариант 3. $n = 96; \;\; k = 5$

12.57	5.83	7.82	14.30	14.12	8.73	6.91	10.79
10.25	12.86	6.27	11.14	7.90	7.46	6.99	9.69
12.29	6.56	14.16	14.28	6.66	10.91	14.12	12.60
8.60	10.32	8.66	7.04	11.74	7.98	12.30	7.21
6.44	11.20	9.03	9.02	8.14	12.70	11.05	8.56
9.32	13.57	6.73	5.98	13.19	14.63	5.80	5.03
8.32	9.11	7.18	8.02	11.97	14.91	10.16	12.41
8.59	5.58	14.20	12.03	14.69	8.57	6.39	13.10
13.98	11.54	13.76	6.19	8.93	7.73	13.20	14.75
12.98	12.14	5.40	5.10	10.45	11.28	7.76	8.49
6.69	14.85	12.98	11.71	8.89	10.48	8.02	5.83
10.70	13.66	14.72	10.71	6.43	8.53	7.14	9.88

Вариант 4. $n = 96;\ \ k = 7$

1.37	0.42	-0.96	-0.88	1.47	3.14	-1.76	-1.35
1.61	0.02	-0.91	-2.59	1.48	-3.00	2.54	1.51
-1.71	1.60	0.00	-0.73	-2.06	-0.67	0.50	0.14
0.98	-0.14	1.93	-1.09	3.12	2.42	-2.09	-2.44
1.19	1.21	-0.09	1.50	2.71	0.09	-3.05	-0.12
2.30	-1.24	2.74	-3.07	-1.98	-0.86	3.12	-2.47
-0.12	-3.14	-2.91	-1.47	-0.74	-1.04	-1.48	-0.65
3.11	1.72	2.92	0.06	-0.05	-0.20	2.41	0.95
1.25	1.68	3.08	1.11	-1.46	2.29	2.65	-0.59
1.92	-1.60	-1.72	1.05	0.64	-2.51	-2.61	-1.14
1.11	-2.95	0.54	0.12	-2.29	2.61	-2.34	0.95
1.12	2.19	2.36	0.79	2.65	-0.63	1.06	3.12

Вариант 5. $n = 88;\ \ k = 5$

-3.55	-1.52	-6.23	5.94	5.26	-5.79	-4.96	3.87
-2.61	-5.33	-2.73	-5.90	-3.11	2.26	4.94	5.69
2.25	-5.40	1.10	4.60	-0.31	-4.69	-3.87	2.82
-4.75	-5.42	-6.19	5.85	3.78	-6.04	-4.86	3.77
-6.28	-6.02	-0.16	-4.45	-5.15	-5.86	5.44	1.71
-1.83	3.03	4.98	3.98	0.88	0.14	3.11	6.28
-1.65	-5.46	1.72	-1.57	-5.36	-5.93	6.07	0.12
4.96	-6.25	6.07	2.62	5.81	-0.63	4.37	4.64
-5.52	-5.85	1.62	2.89	-6.17	-0.38	-6.21	-4.71
-6.23	-6.23	-3.57	3.59	2.02	3.13	1.65	-6.16
5.22	4.23	-4.62	-4.62	6.04	1.80	-6.28	3.80

Вариант 6. $n = 99;\ \ k = 7$

2.17	2.39	2.80	2.54	1.79	1.77	2.09	1.65	2.91
2.53	1.92	2.82	2.10	2.90	1.64	2.93	2.58	1.72
2.23	2.12	1.70	2.61	2.38	1.88	2.21	2.13	2.09
1.97	3.10	2.54	2.61	1.60	2.61	1.88	3.15	2.42
2.72	3.41	2.47	2.53	3.31	2.74	1.67	2.97	2.46
2.36	3.45	2.66	2.56	3.43	1.80	3.48	2.45	2.65
2.53	2.09	2.45	2.11	1.98	2.19	2.59	3.20	2.33
2.58	2.75	2.77	2.41	2.08	2.98	2.08	3.26	2.64
1.93	1.59	1.91	2.84	3.12	2.29	1.65	2.31	3.32
3.23	3.32	2.17	3.35	1.88	3.30	3.06	2.07	2.16
2.87	2.18	3.25	2.01	1.93	3.35	1.78	3.37	2.32

Вариант 7. $n = 100; \ \ k = 9$

3.32	5.41	9.86	5.64	5.97	4.74	6.15	3.60	2.20	9.09
11.78	3.69	5.97	5.02	3.25	2.57	4.94	4.86	7.82	6.72
6.67	3.51	7.72	4.44	3.26	8.03	4.15	3.57	2.83	3.01
5.58	7.48	4.20	1.74	2.33	10.88	5.02	5.39	1.54	9.13
2.81	7.54	3.27	1.80	2.21	3.91	4.74	12.32	6.06	5.92
4.16	2.83	8.91	7.96	7.69	1.42	6.12	2.97	0.80	2.35
7.36	4.83	9.56	7.18	10.13	1.13	6.79	3.98	4.63	1.34
2.99	3.71	6.17	5.97	1.72	4.58	1.02	4.65	1.03	9.55
5.42	7.79	8.13	3.32	6.04	6.04	6.58	8.23	3.46	3.22
2.84	3.61	3.78	8.56	4.35	5.74	4.96	5.01	2.10	5.67

Вариант 8. $n = 96; \ \ k = 7$

-0.99	1.38	0.18	0.47	-0.72	-0.21	1.43	-1.62
-0.82	1.74	-1.41	1.65	0.84	1.28	0.32	-1.43
0.42	1.62	0.83	-0.25	-0.37	-0.54	0.76	1.55
1.02	-0.98	0.24	1.03	-0.14	0.21	2.16	-1.50
-0.96	-0.57	-0.66	-0.39	0.50	-1.14	-1.32	-1.44
1.66	-1.13	-1.10	1.21	0.05	-0.20	-2.02	1.47
1.02	0.31	-0.20	-0.31	0.83	-1.26	-0.15	-0.11
-1.34	0.20	-0.09	-1.16	1.50	2.95	-1.09	1.50
-1.42	-0.51	-0.48	-0.58	-1.10	-0.05	-0.29	-0.66
0.24	-0.83	-0.05	0.21	0.75	0.61	0.10	0.38
0.34	0.58	0.50	0.28	-0.38	0.92	-1.63	-0.09
-0.13	-0.27	-1.68	-0.46	1.01	-0.29	0.39	0.34

Вариант 9. $n = 100; \ \ k = 9$

0.79	0.03	0.18	2.51	2.88	0.59	0.40	0.08	0.02	0.37
0.54	1.06	0.12	0.17	1.42	3.27	0.29	1.57	1.89	0.70
0.28	0.54	0.74	1.28	0.23	1.01	0.41	0.93	0.66	0.05
0.26	1.02	0.43	0.37	0.90	0.16	2.21	0.69	1.06	3.13
0.52	0.86	2.53	0.59	0.76	0.01	1.36	3.20	2.59	1.12
1.84	1.12	1.32	1.65	0.34	1.14	0.61	0.18	0.36	1.88
1.37	0.35	1.11	1.17	0.04	0.15	0.83	2.75	0.20	0.32
0.05	1.03	0.51	1.36	0.23	0.05	0.40	0.87	0.31	0.40
2.02	0.95	0.55	0.24	2.16	0.20	0.12	1.21	1.49	0.97
0.09	1.49	0.52	1.63	3.33	0.52	0.12	0.14	3.19	0.42

Вариант 10. $n = 99$; $k = 7$

-1.00	0.92	0.06	-0.21	0.21	0.44	-0.14	-0.67	0.44
0.65	-0.33	0.19	-0.12	-0.77	0.15	0.67	-0.99	0.59
0.28	0.24	0.13	-0.37	0.14	-0.09	0.79	-0.64	0.30
-0.83	-0.17	-1.00	0.10	-0.21	-0.23	-0.92	-0.57	0.27
1.00	0.48	-0.97	-0.42	-0.46	-0.81	-0.07	-0.59	1.00
-0.95	0.61	-0.29	-1.00	-0.03	0.39	-0.85	0.45	0.29
0.78	0.17	0.87	-0.96	0.21	-0.48	-0.29	0.07	-0.36
0.08	-1.00	0.98	0.85	0.32	-0.24	0.42	-1.00	0.24
0.88	-0.74	-0.28	0.36	0.46	0.64	0.90	0.01	-0.24
0.36	0.01	-0.45	-0.22	-0.29	-0.77	0.40	-1.00	0.15
-0.16	-0.27	-0.27	-0.10	0.69	0.40	-0.08	-0.81	0.17

Вариант 11. $n = 100$; $k = 7$

0.10	0.11	0.78	0.06	0.42	0.22	0.03	0.20	0.01	0.07
0.23	0.17	0.10	0.27	0.46	1.13	0.34	0.96	0.36	0.46
0.08	0.07	0.05	0.53	0.07	1.16	0.04	0.15	0.27	0.15
0.17	0.04	0.07	0.13	0.02	0.00	0.27	0.18	1.14	0.13
0.04	0.02	0.21	1.05	0.37	0.17	0.05	0.05	0.37	0.12
0.03	0.16	0.01	0.09	0.10	0.02	0.00	0.22	0.23	0.24
0.58	0.01	0.10	0.79	0.34	0.50	0.16	0.44	0.39	0.01
0.11	0.48	0.82	0.11	0.02	0.06	0.01	0.36	0.78	0.00
0.37	0.36	0.70	0.04	0.14	0.02	0.20	0.02	0.19	0.02
0.25	0.05	0.29	0.17	0.30	0.03	0.01	0.14	0.30	0.50

Вариант 12. $n = 100$; $k = 7$

0.52	0.08	0.00	0.03	0.42	0.00	0.29	0.13	0.31	0.63
0.03	0.22	0.06	0.04	0.22	0.57	0.18	0.11	0.21	0.54
0.12	0.00	0.03	0.24	0.30	0.04	0.13	0.35	0.32	0.02
0.10	0.03	0.33	0.09	0.27	0.02	0.29	0.25	0.09	0.15
0.29	0.07	0.11	0.06	0.02	0.07	0.06	0.05	0.21	0.10
0.19	0.45	0.15	0.22	0.38	0.54	0.04	0.03	0.03	0.32
0.12	0.04	0.55	0.01	0.01	0.01	0.02	0.49	0.59	0.11
0.22	0.11	0.43	0.19	0.05	0.08	0.56	0.10	0.25	0.91
0.13	0.21	0.16	0.06	0.11	0.11	0.27	0.18	0.13	0.21
0.45	0.15	0.35	0.03	0.08	0.04	0.36	0.01	0.12	0.08

Вариант 13. $n = 99$; $k = 5$

22.39	13.89	8.89	3.14	9.14	11.11	12.01	15.16	10.23
5.87	7.09	4.31	10.83	7.02	12.24	11.35	6.16	11.42
2.87	8.75	7.23	8.91	3.73	14.96	14.46	19.13	11.22
10.57	12.11	19.54	2.64	9.00	8.48	12.44	9.60	7.99
5.96	7.39	3.35	14.25	10.29	13.13	15.60	6.20	14.93
16.27	3.75	6.70	17.96	8.92	19.95	14.72	7.36	5.87
6.58	8.69	6.19	10.82	5.68	12.08	16.21	13.01	7.74
5.63	11.75	5.16	14.70	8.37	18.13	9.97	24.35	19.94
7.40	6.22	8.68	8.68	8.96	9.96	3.39	7.89	11.91
10.97	9.11	3.25	12.53	19.96	13.72	6.06	15.21	4.66
10.22	12.53	16.18	10.51	5.20	6.85	7.80	8.86	10.57

Вариант 14. $n = 100$; $k = 9$

3.05	3.20	3.57	3.78	2.23	3.86	2.08	3.06	3.75	2.76
2.93	2.12	3.37	3.32	3.14	2.13	3.67	3.47	3.60	3.62
3.86	2.39	3.08	3.41	3.82	3.16	2.77	2.18	2.84	2.74
2.97	2.03	2.11	3.63	3.25	3.42	2.99	3.99	2.53	2.08
2.62	3.09	2.01	3.05	3.04	2.54	3.97	2.84	3.53	3.32
2.23	3.42	2.13	3.35	3.62	2.50	3.26	2.60	3.31	2.67
3.78	4.00	3.00	3.86	2.02	3.58	2.55	3.26	2.26	3.61
2.75	3.60	3.32	3.95	3.68	2.55	3.80	3.70	2.90	2.75
3.19	2.16	3.81	3.29	2.36	2.84	2.30	3.87	2.25	2.79
2.68	3.31	2.11	2.19	3.33	2.18	2.22	2.60	3.04	2.80

Вариант 15. $n = 100$; $k = 7$

1.94	0.93	0.21	0.63	1.08	2.16	1.47	2.44	0.52	0.70
2.50	0.77	0.71	1.85	0.35	0.77	1.73	2.75	0.92	1.63
0.35	1.27	0.88	1.21	0.48	1.22	0.34	0.97	1.92	1.59
1.04	1.09	0.26	1.37	0.32	0.51	1.98	2.39	0.45	1.07
1.14	0.50	2.06	0.83	0.39	0.62	1.51	0.55	1.18	0.84
1.68	0.23	0.53	0.16	1.39	0.36	0.93	0.54	1.89	1.19
0.38	2.61	0.67	1.10	0.56	2.09	0.90	1.06	0.55	1.63
1.39	1.81	2.11	1.34	2.63	1.11	0.48	1.40	3.00	0.79
0.59	1.27	0.43	1.67	0.99	3.27	0.63	1.79	0.81	1.06
1.61	1.42	1.37	1.78	1.02	0.93	0.24	1.99	0.86	1.58

Вариант 16. $n = 100; \quad k = 5$

0.57	0.15	0.05	2.58	0.08	0.71	0.17	0.16	0.46	0.12
0.12	0.25	1.62	0.60	0.62	1.12	0.39	0.01	1.08	0.10
0.70	0.61	0.24	0.28	0.15	1.00	1.01	0.53	0.66	0.18
1.00	0.11	0.25	0.05	0.05	0.17	0.38	1.25	0.57	0.10
0.89	0.02	0.66	0.90	0.66	2.44	0.24	0.04	0.18	0.42
0.11	0.58	0.89	1.22	0.46	0.43	0.03	0.50	0.82	0.18
0.13	0.24	0.02	1.45	0.45	0.56	0.19	0.76	0.33	2.29
3.25	0.13	0.95	0.39	0.31	0.83	0.02	2.21	0.15	0.34
1.19	0.21	0.45	0.23	1.84	0.72	0.48	0.51	0.71	0.17
2.22	0.51	0.69	1.11	0.37	0.24	0.34	0.51	0.07	0.51

Вариант 17. $n = 99; \quad k = 7$

-2.50	-1.21	-1.28	-2.86	-1.41	-1.80	-1.44	-2.80	-2.75
-2.24	-1.49	-1.78	-2.58	-2.70	-2.68	-2.54	-2.46	-2.66
-2.46	-2.45	-1.49	-1.37	-2.71	-1.30	-2.27	-2.47	-2.27
-2.26	-1.19	-1.53	-1.01	-1.05	-2.05	-2.44	-2.89	-1.58
-1.73	-2.68	-2.87	-1.23	-1.11	-2.82	-2.24	-1.10	-2.47
-1.74	-1.40	-2.36	-1.91	-1.76	-2.45	-2.67	-1.09	-2.05
-2.13	-1.52	-2.93	-2.79	-2.04	-2.64	-2.35	-2.69	-2.72
-1.81	-2.66	-2.12	-2.21	-1.42	-1.12	-1.59	-2.72	-2.55
-2.43	-1.82	-2.59	-2.06	-1.17	-2.16	-2.32	-1.41	-1.64
-2.69	-1.94	-1.81	-2.35	-1.86	-1.49	-1.82	-2.54	-1.52
-1.98	-2.03	-2.64	-2.06	-1.67	-2.85	-1.44	-2.43	-1.58

Вариант 18. $n = 99; \quad k = 7$

-2.61	-1.43	-3.37	-1.26	-1.83	-2.76	-3.31	-3.47	-2.69
-1.57	-2.86	-3.55	-3.10	-3.48	-3.62	-2.33	-1.99	-3.43
-1.63	-1.52	-1.02	-1.83	-3.09	-2.68	-2.52	-3.26	-3.62
-1.92	-3.25	-1.70	-2.19	-3.03	-2.16	-2.04	-3.52	-2.45
-2.08	-3.86	-2.59	-1.43	-1.98	-1.72	-3.15	-2.42	-1.80
-1.92	-1.01	-2.24	-1.33	-1.66	-3.24	-1.08	-2.26	-1.67
-1.16	-3.45	-1.43	-2.89	-2.89	-2.15	-3.73	-1.80	-1.99
-3.26	-2.09	-1.84	-1.38	-3.51	-2.27	-2.36	-3.36	-3.66
-3.24	-3.92	-3.86	-1.61	-3.71	-2.87	-2.92	-3.90	-2.80
-3.62	-2.52	-1.45	-3.20	-2.34	-3.94	-2.87	-1.93	-2.36
-3.92	-1.01	-2.32	-2.83	-3.39	-2.32	-3.88	-1.30	-3.05

Вариант 19. $n = 96; \; k = 5$

-3.98	0.94	1.99	-3.97	5.97	-5.04	-4.70	5.87
-2.27	-3.27	6.58	3.34	-1.15	0.21	3.05	-2.66
4.26	-4.40	1.66	-0.56	-1.23	-5.62	-8.08	-5.68
2.05	6.97	-0.98	4.09	-5.70	-0.42	-5.25	1.53
0.80	-0.81	3.30	1.13	3.17	-4.36	1.35	-6.72
-3.86	-1.93	-1.56	6.00	-0.53	11.80	0.41	-2.32
-3.34	3.32	4.82	6.02	-1.02	-2.04	0.97	2.01
0.71	-0.20	-0.59	-4.42	-2.17	2.42	-2.46	-4.65
4.06	1.89	-2.31	-0.36	0.84	6.12	5.13	-1.90
-5.62	4.13	0.85	3.00	8.63	1.28	-5.99	-1.17
-6.13	-4.51	5.53	-0.34	-0.81	6.48	-0.19	-1.09
2.65	-1.48	4.04	-2.89	1.55	-5.77	1.38	3.69

Вариант 20. $n = 100; \; k = 9$

0.59	0.33	0.41	0.73	0.62	0.14	1.32	0.13	0.24	0.77
0.61	0.82	0.74	0.57	0.68	0.60	0.73	0.79	0.30	0.45
0.46	0.09	0.82	0.74	0.57	0.68	0.69	0.73	0.79	0.35
0.31	0.93	0.92	1.25	0.39	0.32	0.24	0.64	0.47	0.75
1.12	0.69	0.51	0.29	0.23	0.83	0.41	0.86	0.53	0.38
0.51	0.20	0.55	0.76	1.30	0.69	0.67	0.20	0.40	0.67
0.97	0.56	0.81	0.38	0.49	1.07	0.37	0.16	0.78	0.51
0.42	0.17	0.65	1.15	0.11	0.91	0.91	0.46	0.41	0.32
0.63	0.42	0.52	0.34	0.56	0.23	0.81	0.75	0.28	0.72
0.85	0.55	0.37	0.53	0.85	0.40	0.29	0.21	0.82	0.35

Вариант 21. $n = 90; \; k = 5$

1.39	3.56	1.42	2.81	2.27	0.67	1.90	0.07	0.59	2.35
2.91	0.75	0.81	1.90	2.26	0.03	5.54	0.75	2.66	0.07
0.61	0.01	0.87	0.49	0.65	0.76	0.65	6.89	2.19	1.37
2.88	1.28	1.82	0.02	1.00	2.52	1.00	2.19	3.78	2.02
0.38	0.37	1.61	3.06	3.17	4.40	0.33	0.58	0.35	4.09
1.32	5.09	1.33	0.24	4.68	0.69	4.93	1.91	2.20	0.27
1.81	6.96	5.80	0.69	1.43	6.17	1.69	0.10	3.17	0.45
1.87	1.41	4.14	3.61	4.15	0.27	0.02	0.06	0.06	0.32
3.88	4.27	2.03	1.12	4.40	3.66	2.94	0.13	4.90	0.94

Вариант 22. $n = 88;\ \ k = 5$

–1.57	–1.39	–1.57	–0.90	0.15	0.62	–1.51	–1.22
–1.55	–0.10	–0.40	–1.35	1.06	1.23	–1.48	–1.28
–1.34	1.56	–1.53	–1.57	–1.56	–1.47	1.29	–0.38
–0.71	–1.57	–0.33	1.37	–1.00	0.52	0.35	–0.21
–1.56	–1.32	–0.58	–0.33	0.05	–1.44	–0.93	0.40
0.30	1.70	1.05	–0.92	–1.05	–1.30	–0.45	–1.40
–1.18	–0.79	0.85	–1.05	–0.97	–1.50	–1.40	0.40
–1.55	–1.32	–0.71	–0.73	–1.41	–0.07	1.56	1.17
–0.39	–0.15	–0.82	–0.49	–0.39	0.10	1.36	–0.78
–1.11	–1.52	–1.30	0.35	–1.32	1.38	–1.56	–1.49
0.64	–1.51	0.90	–0.47	–1.56	–0.22	–1.45	–0.66

В варианте 22- надо сделать сдвиг на +3 единицы.

Вариант 23. $n = 96;\ \ k = 7$

-1.21	1.31	-1.60	1.29	-2.59	2.76	-0.85	-0.18
-3.25	1.06	-2.04	0.10	-2.05	2.82	-0.47	1.88
1.22	3.24	1.13	0.87	-1.73	3.40	0.87	0.97
-0.87	1.80	1.65	2.20	-3.19	-1.96	0.59	-0.98
2.25	3.10	0.56	0.75	-1.36	-0.24	1.50	-0.77
-1.79	-0.77	0.75	-1.89	-1.16	-0.18	0.60	2.61
-1.21	-2.30	0.84	-5.61	2.64	-3.94	1.73	-1.33
0.95	1.35	2.49	-0.56	0.05	3.17	1.59	-4.59
4.47	3.81	0.27	0.82	1.24	-0.23	-0.32	-1.33
0.52	-0.01	-0.24	-1.87	-0.09	2.81	3.35	1.68
-1.62	3.89	-0.85	-1.59	1.49	1.66	-2.50	-2.53
0.16	-1.46	-0.19	1.99	0.82	-0.52	-2.62	0.27

Вариант 24. $n = 100;\ \ k = 9$

2.67	2.41	2.54	3.12	2.02	2.68	2.57	1.47	1.38	2.72
3.85	1.91	2.23	3.33	2.66	1.49	2.75	3.87	2.36	3.26
2.48	1.99	2.94	2.90	3.58	1.13	3.36	3.70	2.71	1.42
1.47	3.05	1.13	1.74	2.67	2.43	1.98	3.14	3.39	2.71
1.05	2.47	1.04	2.46	3.16	2.16	1.74	1.26	1.54	1.12
2.68	3.55	2.23	3.17	2.60	2.18	1.72	3.08	2.46	2.57
2.63	3.59	2.46	3.42	2.62	3.51	3.53	2.72	1.11	3.45
3.85	2.38	3.72	1.01	1.98	3.54	1.64	2.71	1.33	2.28
1.57	1.17	1.02	1.20	3.00	2.93	2.76	3.38	1.71	3.65
3.50	1.49	3.29	2.24	1.88	3.43	3.43	3.20	2.07	2.79

Вариант 25. $n = 99; \ \ k = 7$

0.51	-0.93	-0.68	-0.03	0.06	-1.04	-0.12	-0.62	0.63
-0.92	-0.24	-0.34	-0.21	-0.17	1.88	0.19	1.29	-0.18
1.14	0.32	-0.07	-0.18	0.51	-0.46	0.60	0.09	-0.01
1.06	0.25	-0.67	-0.32	-0.77	-0.09	-0.31	-0.46	-0.10
-0.06	0.13	0.60	-0.39	-1.20	0.88	0.13	0.39	-0.58
-0.17	-0.37	-0.14	-0.06	0.51	-0.34	0.62	0.84	0.10
-0.14	-0.40	-0.84	0.05	0.23	-0.21	-0.40	0.37	0.10
0.47	-0.41	0.52	0.02	0.51	0.03	0.69	0.20	-0.60
-0.07	-0.01	-0.57	-0.76	-0.30	-0.06	0.28	0.07	-0.18
0.47	-0.44	0.42	-0.72	0.11	-0.09	1.28	0.43	0.74
-0.25	-0.39	0.31	0.43	0.65	0.31	0.35	0.46	-0.04

Вариант 26. $n = 100; \ \ k = 7$

2.54	11.58	5.22	5.24	9.72	7.36	4.01	3.47	2.08	3.68
14.85	8.12	10.32	2.26	4.28	3.68	6.76	6.71	5.87	7.01
7.01	11.71	4.43	6.83	7.86	8.95	2.67	5.97	0.51	9.41
6.59	4.00	4.84	8.76	5.70	2.41	6.78	1.27	7.28	9.03
4.24	2.45	5.31	9.86	10.17	4.29	7.32	10.21	4.30	1.53
5.13	5.20	4.32	10.74	8.26	0.65	3.89	15.80	6.90	7.57
2.94	4.47	2.21	5.89	3.07	4.61	5.80	9.09	5.04	3.25
3.82	3.87	14.59	4.46	7.63	7.67	14.34	5.59	8.37	7.43
2.85	6.28	9.67	5.17	3.80	3.75	3.49	8.53	5.78	6.52
7.12	4.76	6.88	7.21	13.00	5.77	8.78	4.09	1.09	6.21

Вариант 27. $n = 99 \ \ k = 5$

2.38	1.58	0.02	-0.18	-0.51	2.29	-0.55	2.61	-6.28
-2.02	2.75	-2.70	1.61	-0.49	-4.95	-0.36	1.22	-2.77
0.11	-1.45	6.27	-2.70	-0.29	2.79	6.24	-3.26	1.83
0.53	-0.26	1.09	6.20	-0.28	-0.83	1.17	-1.92	-0.39
-2.56	3.45	-2.19	-4.16	-0.16	-3.17	0.45	-0.13	2.90
-2.03	0.42	-3.94	-0.75	2.98	-3.88	5.02	1.02	-4.76
-0.53	-3.59	-2.52	3.86	-3.30	4.52	-2.86	-0.16	-1.49
0.02	-0.19	-3.92	-2.10	-1.74	2.16	-0.99	-2.74	3.78
-2.66	3.32	4.37	-0.01	4.84	-2.32	2.29	1.18	-5.19
-1.34	1.46	0.55	-0.86	-0.85	-2.69	-1.26	-0.69	-1.33
-4.32	0.41	1.85	-1.50	2.58	4.02	4.00	3.20	4.09

Вариант 28. $n = 100; \ k = 9$

7.24	2.47	3.99	0.83	5.20	1.93	2.44	3.62	2.97	4.24
3.40	2.15	1.56	8.00	3.01	6.85	4.67	3.10	5.66	3.47
5.95	0.67	1.83	4.63	4.17	2.94	3.84	2.02	1.81	4.83
7.08	3.27	3.18	2.63	1.37	2.29	4.57	0.91	2.08	5.18
5.24	2.54	2.46	1.43	5.70	3.14	5.18	2.37	4.59	3.55
1.77	0.42	3.17	7.02	5.74	4.30	2.84	4.69	3.42	1.32
2.34	6.08	5.44	5.37	4.60	7.58	2.98	8.24	0.73	3.58
4.18	0.60	4.15	6.11	2.24	2.62	5.31	2.21	5.34	4.33
4.23	5.72	4.38	1.58	0.64	2.88	4.83	5.14	1.83	5.90
1.25	3.95	2.20	5.11	1.69	4.19	3.23	2.75	2.95	1.99

Вариант 29. $n = 99; \ k = 7$

36.36	0.21	26.80	15.85	2.92	20.63	5.01	3.12	12.19
9.02	34.47	4.81	4.49	7.93	17.43	7.98	9.09	15.46
4.24	3.37	11.78	4.64	0.79	11.21	3.94	13.65	9.51
8.81	9.63	5.02	11.30	11.49	33.06	25.32	6.63	11.69
1.93	2.67	4.55	4.75	8.89	7.99	0.87	23.55	15.03
9.15	11.83	0.02	1.20	13.16	3.35	3.20	9.24	10.06
11.84	3.07	8.06	7.92	0.36	14.30	20.67	3.95	2.39
0.88	1.96	14.26	4.69	25.77	13.70	33.80	9.51	22.20
23.86	10.04	7.10	12.10	2.85	3.12	3.65	2.92	15.77
11.46	7.28	7.45	11.49	6.72	22.15	2.57	8.08	18.04
16.67	2.34	8.67	23.75	11.50	2.70	5.27	20.93	2.58

Вариант 30. $n = 100; \ k = 9$

1.46	1.77	6.06	2.25	0.42	8.41	2.04	2.23	4.00	3.11
5.23	2.92	0.04	1.78	0.57	3.75	1.05	7.19	1.88	0.26
1.89	4.24	5.81	3.78	3.18	0.86	6.66	1.05	0.46	2.18
3.90	6.57	2.29	4.76	1.17	2.95	1.45	1.82	2.60	4.56
4.38	3.39	6.97	0.60	0.56	0.73	1.79	9.44	3.96	0.77
0.09	2.76	8.32	0.35	0.86	2.53	2.25	1.29	0.68	0.02
3.79	1.23	0.16	1.49	0.14	0.31	0.77	0.65	2.74	5.89
4.83	0.63	2.70	5.93	6.25	6.60	1.30	4.60	2.56	0.08
4.91	0.44	0.33	0.24	3.94	3.99	4.42	3.00	2.85	2.01
1.08	3.24	0.55	10.24	2.09	0.74	1.05	0.71	0.08	0.32

Вариант 31. $n = 55$

17	19	23	18	21	15	16	13	20	18	15
20	14	20	16	14	20	19	15	19	16	19
15	22	21	12	10	21	18	14	14	17	16
13	19	18	20	24	16	20	19	17	18	18
21	17	19	17	13	17	11	18	19	19	17

Вариант 32. $n = 55$

14	16	20	15	18	12	13	10	17	15	12
17	11	17	13	11	17	16	12	16	13	16
12	19	18	9	7	18	15	11	11	14	13
10	16	15	17	21	13	17	16	14	15	15
18	14	16	14	10	14	8	15	16	16	13

Вариант 33. $n = 55$

15	17	21	16	19	13	14	11	18	16	13
18	12	18	14	12	18	17	13	17	14	17
13	20	19	10	8	19	16	12	12	15	14
19	15	17	15	11	15	9	16	17	17	15
11	17	16	18	22	14	18	17	15	16	16

Вариант 34. $n = 55$

16	18	22	17	20	14	15	12	19	17	14
19	13	19	15	13	19	18	14	18	15	18
14	21	20	11	9	20	17	13	13	16	15
12	18	17	19	23	15	19	18	16	17	17
20	16	18	16	12	16	10	17	18	19	16

Вариант 35. $n = 55$

18	20	24	19	22	16	17	14	21	19	16
21	15	21	17	15	21	20	16	20	17	20
16	23	22	13	11	22	19	15	15	18	17
14	20	19	21	25	17	21	20	18	19	19
22	18	20	18	14	18	12	19	20	20	18

Вариант 36. $n = 55$

20	22	26	21	24	18	19	16	23	21	18
23	17	23	19	17	23	18	22	19	22	18
18	25	24	15	13	24	21	17	17	20	19
16	22	21	23	27	19	23	22	20	21	21
24	20	22	20	16	20	14	21	22	22	20

Вариант 37. $n = 55$

27	19	33	28	31	25	26	23	30	28	25
30	24	30	26	24	30	29	25	29	26	29
25	32	31	22	20	31	28	24	24	27	26
23	29	28	30	34	26	30	29	27	28	28
31	27	29	27	23	27	21	28	29	29	27

Вариант 38. $n = 55$

19	21	25	20	23	17	18	15	22	20	17
22	16	22	18	16	22	21	17	21	18	21
17	24	23	14	12	23	20	16	16	19	18
15	21	20	22	26	18	22	21	19	20	20
23	19	21	19	15	19	13	20	21	21	19

Вариант 39. $n = 55$

26	28	32	27	30	24	25	22	29	27	24
29	23	29	25	23	29	28	24	28	25	28
24	31	30	21	19	30	27	23	23	26	25
22	28	27	29	33	25	29	28	26	27	27
30	26	28	26	22	26	20	27	28	28	26

Вариант 40. $n = 55$

21	23	27	22	25	19	20	17	24	22	19
24	18	24	20	18	24	23	19	23	20	23
19	26	25	16	14	25	22	18	18	21	20
17	23	22	24	28	20	24	23	21	22	22
25	21	23	21	17	21	15	22	23	23	21

Вариант 41. $n = 55$

22	24	28	23	26	20	21	18	25	23	20
25	19	25	21	19	25	24	20	24	21	24
20	27	26	17	15	26	23	19	19	22	21
18	24	23	25	29	21	25	24	22	23	23
26	21	24	22	18	22	16	23	24	24	22

Вариант 42. $n = 55$

25	27	31	26	27	23	24	21	28	26	23
28	22	28	24	22	28	27	23	27	24	27
23	30	29	20	18	29	26	22	22	25	24
21	27	26	28	32	24	28	27	25	26	26
29	25	27	25	21	25	19	26	27	27	25

Вариант 43. $n = 55$

23	25	29	24	27	21	22	19	26	24	21
26	20	26	22	20	26	25	21	25	22	25
21	28	27	18	16	27	24	20	20	23	22
19	25	24	26	30	22	26	25	23	24	24
27	23	25	23	19	23	17	24	25	25	23

Вариант 44. $n = 55$

24	26	30	25	28	22	23	20	27	25	22
27	21	27	23	21	27	26	22	26	23	26
22	29	28	19	17	28	25	21	21	24	23
20	26	25	27	31	23	27	26	24	25	25
28	24	26	24	20	24	18	25	26	26	24

Вариант 45. $n = 55$

28	30	34	29	32	26	27	24	31	29	26
31	25	31	27	25	31	30	26	30	27	30
26	33	32	23	21	32	29	25	25	28	27
24	30	29	31	35	27	31	30	28	29	29
32	28	30	28	24	28	22	29	30	30	28

Вариант 46. $n = 100; \ \ k = 9$

113	50	52	62	41	104	135	268	29	113
6	64	202	188	26	116	175	267	193	257
70	65	25	75	38	265	323	164	144	63
60	152	27	98	3	108	77	106	205	8
23	100	119	7	83	38	19	176	69	138
9	18	294	101	33	142	48	57	38	14
105	26	203	48	211	40	317	116	566	102
172	72	3	137	92	95	18	167	147	141
72	17	36	25	119	99	88	2	12	125
11	217	469	49	197	1	23	136	77	72

Вариант 47. $n = 100; \ \ k = 9$

91	99	106	93	96	91	96	111	86	99
100	94	99	114	101	91	106	101	114	97
97	109	95	101	91	93	98	101	99	115
98	112	100	106	89	103	114	99	108	101
111	116	108	94	90	80	94	98	92	103
102	101	97	88	86	80	83	116	100	103
82	111	101	127	103	94	102	100	86	88
98	118	110	85	85	89	95	107	117	91
112	97	109	100	100	98	96	85	99	101
112	102	89	99	113	89	80	94	93	93

Вариант 48. $n = 100; \ \ k = 9$

365	282	10	218	126	341	340	140	58	378
52	80	384	233	37	18	211	369	96	87
139	274	239	187	301	66	128	87	43	266
181	22	57	89	43	203	314	137	272	50
42	112	228	245	153	139	342	197	290	234
200	299	279	262	227	367	107	384	358	140
392	37	355	253	221	136	232	206	152	296
380	310	246	66	234	236	148	242	156	387
43	393	295	243	148	376	324	318	309	391
58	317	73	353	166	113	42	143	248	45

Вариант 49. n = 100; k = 9

69	11	39	86	74	46	132	106	13	49
149	10	68	46	70	56	275	25	34	47
34	7	79	38	144	54	28	11	50	7
4	79	55	80	136	12	13	45	25	11
4	10	27	236	70	171	324	75	20	70
123	96	117	16	10	76	40	253	86	3
83	135	92	67	87	20	3	8	167	53
27	51	58	53	8	54	12	59	30	38
28	29	213	96	100	47	114	193	82	5
130	37	15	44	68	82	18	50	112	2

Вариант 50. n = 100; k = 9

204	218	212	146	209	241	188	241	248	264
177	202	260	235	196	170	154	190	198	137
170	168	199	156	218	174	242	218	177	185
222	196	227	149	244	231	187	193	231	136
177	174	200	142	196	219	224	152	153	214
168	217	139	228	226	240	213	206	173	273
194	216	233	187	172	205	181	156	147	151
212	230	207	179	220	191	203	207	159	213
224	242	194	166	180	142	130	284	189	213ё
131	227	163	221	211	198	219	236	146	189

Вариант 51. n = 100; k = 9

22	4	164	205	188	27	26	11	351	376
199	83	70	54	308	176	36	159	355	58
140	63	51	31	81	2	15	32	114	24
67	250	455	1	141	19	104	78	309	75
69	274	1	540	83	199	9	143	74	88
235	361	62	115	84	519	182	98	105	277
356	77	128	256	36	75	55	520	307	147
42	23	431	5	1	105	133	122	273	272
765	84	284	157	105	50	92	263	154	139
67	126	198	15	301	26	1	216	185	194

Вариант 52. $n = 100$; $k = 9$

78	75	80	107	92	114	49	79	57	99
112	81	102	85	68	100	78	88	76	60
95	82	98	66	81	81	92	82	107	140
78	93	83	74	75	124	57	78	104	80
61	66	84	84	102	93	95	102	92	93
116	79	79	82	67	119	114	74	88	66
106	84	119	86	105	106	78	104	81	95
87	103	63	86	84	89	100	113	70	87
83	91	103	100	42	96	104	94	145	80
92	93	103	80	87	62	86	110	109	66

Вариант 53. $n = 100$; $k = 9$

59	299	101	80	93	576	78	314	56	122
346	34	227	131	38	215	29	394	159	509
268	8	174	105	331	70	347	433	67	100
137	57	64	276	76	33	196	80	36	466
228	24	42	14	81	337	179	32	50	95
34	179	102	29	277	74	50	11	51	114
29	31	55	2	18	79	48	244	36	9
51	324	71	17	25	9	112	106	529	17
161	4	137	117	161	121	539	25	279	1
153	2	41	241	429	142	152	55	525	201

Вариант 54. $n = 100$; $k = 9$

123	117	152	95	152	128	107	97	170	81
152	85	88	162	152	81	84	169	151	138
127	168	164	187	99	107	116	160	117	113
81	110	122	132	104	148	144	112	138	116
60	90	110	157	112	162	92	113	179	156
101	155	140	139	119	156	115	148	162	197
73	130	155	91	134	121	112	129	117	160
107	125	112	102	154	99	119	136	90	133
107	118	167	161	135	114	131	144	143	133
146	73	70	121	106	61	127	115	176	70

Вариант 55. $n = 100;\ k = 9$

154	191	196	125	219	22	27	98	84	157
134	99	7	9	34	136	261	216	6	168
257	332	27	135	204	192	68	55	51	11
104	91	217	77	152	12	250	196	12	90
112	58	214	147	163	158	36	164	59	31
24	72	104	78	170	97	101	59	20	140
253	6	3	15	14	60	122	40	84	66
178	318	18	75	75	24	19	312	57	32
52	429	466	233	278	237	84	112	97	915
92	210	142	310	89	59	21	182	284	20

Вариант 56. $n = 100;\ k = 9$

89	128	87	109	138	110	119	129	120	137
139	92	89	107	112	106	101	82	121	121
109	100	72	87	108	82	109	134	74	106
97	111	100	113	80	88	102	62	124	74
75	113	113	100	72	106	112	95	68	101
65	89	99	96	98	111	39	92	114	91
103	129	119	139	145	86	136	64	80	122
92	97	82	78	110	66	87	93	136	122
89	128	113	75	83	80	107	133	123	91
102	86	105	80	103	105	102	106	95	59

Вариант 57. $n = 100;\ k = 9$

142	372	274	336	252	172	148	307	325	111
414	272	58	343	192	85	142	283	360	176
442	475	217	318	90	66	100	313	387	244
74	332	170	138	77	476	290	484	182	444
105	199	67	212	385	267	405	385	473	95
305	396	71	272	277	431	225	497	354	308
383	316	126	272	62	423	237	106	173	473
399	142	415	161	209	152	378	406	88	137
310	152	80	338	198	484	106	449	305	304
343	308	475	215	358	234	453	227	150	138

Вариант 58. $n = 100; \;\; k = 9$

103	153	308	10	180	30	480	18	148	191
188	115	4	42	317	27	25	227	187	136
22	52	152	50	111	81	94	213	523	129
10	556	34	89	26	28	92	65	250	201
63	65	196	479	220	33	197	57	12	190
3	233	26	119	174	78	45	592	54	122
170	104	58	156	180	66	342	124	15	476
87	6	38	45	169	34	115	10	176	13
23	382	56	48	10	17	0	275	563	48
41	251	9	170	144	35	193	118	82	58

Вариант 59. $n = 100; \;\; k = 9$

92	135	67	85	140	136	79	164	91	96
89	196	110	117	140	233	116	165	100	82
102	169	153	109	162	129	61	53	113	148
43	139	120	156	143	34	139	153	52	107
55	138	96	127	105	146	106	184	139	176
90	117	147	158	128	120	120	160	158	210
120	148	152	122	101	86	153	120	160	149
136	115	112	55	177	195	123	84	111	95
113	166	86	149	144	125	74	162	101	31
119	153	114	58	125	172	105	205	168	97

Вариант 60. $n = 100; \;\; k = 9$

485	98	116	102	109	357	73	208	49	7
420	105	356	38	266	201	3	94	52	49
82	150	171	12	0	119	84	17	221	3
53	44	233	9	125	202	314	215	27	59
40	42	517	42	96	35	94	9	101	292
25	2	270	38	258	59	17	26	342	73
433	45	69	22	505	108	59	131	22	157
131	90	34	32	133	157	154	42	171	61
130	61	64	40	35	139	20	772	10	8
30	0	28	4	1	48	353	52	25	13

Вариант 61. $n = 100; \ k = 9$

254	284	148	262	267	303	181	303	288	288
265	160	266	200	224	218	301	201	284	165
291	254	317	288	236	269	249	223	286	308
154	307	239	283	105	277	245	306	278	219
302	239	345	270	261	193	232	217	270	229
263	296	129	210	262	245	286	298	302	249
241	204	209	246	193	239	357	223	242	216
269	266	187	247	221	262	176	242	285	197
259	214	287	258	286	274	250	283	205	185
200	257	215	307	259	168	267	203	253	202

Вариант 62. $n = 100; \ k = 9$

102	159	175	168	274	47	61	87	54	40
18	72	22	0	185	13	20	48	65	103
10	293	84	143	39	33	35	12	50	57
62	69	42	3	15	215	123	11	10	14
44	151	508	62	2	76	32	136	87	48
189	52	17	120	75	27	188	186	18	170
39	72	99	157	10	191	52	182	86	37
9	51	204	30	67	18	40	42	100	40
178	89	43	79	193	33	138	111	28	79
47	30	156	130	15	156	68	69	350	104

Вариант 63. $n = 100; \ k = 9$

310	276	378	328	296	333	281	333	224	398
351	337	379	410	302	311	325	308	268	343
406	271	264	402	365	298	363	337	247	329
339	357	335	421	446	372	195	395	279	350
350	369	332	320	304	283	378	252	340	315
431	474	324	405	447	236	300	431	318	227
250	333	299	415	280	410	223	350	313	300
346	361	334	304	321	296	266	299	282	426
394	345	183	299	401	262	278	243	333	287
461	349	352	361	327	286	323	302	369	261

Вариант 64. $n = 100;\ \ k = 9$

5	133	81	22	59	58	102	10	29	135
109	7	9	19	15	0	118	86	156	25
51	74	84	97	28	38	44	214	185	88
3	57	123	80	205	18	40	101	98	37
134	28	23	115	40	3	56	63	4	172
35	40	16	18	173	14	92	9	78	36
36	183	58	57	18	27	51	2	47	11
177	147	54	4	110	18	28	129	80	68
99	56	31	166	73	31	20	3	63	14
58	56	48	323	22	90	137	406	90	78

Вариант 65. $n = 100;\ \ k = 9$

352	506	594	565	588	556	663	454	573	465
571	573	330	636	474	651	599	575	619	506
501	593	531	496	507	555	457	530	459	323
552	517	557	663	613	635	372	447	526	614
556	709	447	450	590	503	532	504	495	502
666	594	562	389	648	628	541	659	631	515
365	455	503	433	504	541	419	471	440	599
446	538	732	584	560	723	503	540	646	663
494	509	592	627	635	446	462	615	499	484
593	354	555	397	692	654	555	678	553	612

Вариант 66. $n = 100;\ \ k = 9$

129	246	11	191	374	75	249	20	55	113
147	67	103	36	142	109	57	128	208	255
89	27	19	34	61	63	117	184	28	71
12	119	5	135	24	131	110	64	1	9
14	35	119	46	1	18	35	117	63	36
11	41	4	150	194	37	22	202	104	71
100	20	89	146	38	70	31	81	32	54
204	61	19	106	15	58	84	23	5	162
410	20	21	271	122	55	355	152	38	203
79	10	1	102	62	54	94	258	10	155

Вариант 67. $n = 100;\ k = 9$

329	255	99	178	316	290	216	217	389	284
260	306	367	253	261	248	309	209	308	207
333	227	250	288	199	322	176	294	267	86
175	150	294	255	335	320	221	242	387	248
166	181	360	217	128	147	221	273	286	393
181	278	280	214	268	278	168	283	272	255
189	315	281	195	283	203	385	201	209	322
352	317	206	267	273	227	212	230	235	156
287	105	287	208	250	328	100	186	279	252
182	254	315	312	381	218	236	220	248	240

ЗАДАЧА 2

Дана корреляционная таблица. По данным таблицы необходимо выполнить следующие две лабораторные работы.

Лабораторная работа №4. Расчет параметров корреляционной зависимости (прямая или обратная линейная зависимость).

1. Найти эмпирические распределения каждой из компонент, сделать их графическое изображение.

2. Вычислить средние значения $\overline{X}$ и $\overline{Y}$ и средние квадратические отклонения $\sigma_X;\ \sigma_Y$.

3. Вычислить эмпирические линии регрессии $Y = Y(X)$ и $X = X(Y)$, сделать отдельные графики этих линий.

4. Рассчитать параметры теоретических линий регрессии $Y_x = aX + b$ и $X_y = a_1 Y + b_1$. Сделать графики этих линий, располагая каждую на графике соответствующей эмпирической линии.

Лабораторная работа №5 (на усмотрение преподавателя). Расчет параметров корреляционной зависимости (квадратичная зависимость).

Пункты 1 – 3 – те же, что и в лабораторной работе №4.

5. Используя метод наименьших квадратов (см. [10] в учебных пособиях кафедры), вычислить параметры квадратичной зависимости $Y(X) = a_1 X^2 + b_1 X + c_1$ и $X(Y) = a_2 Y^2 + b_2 Y + c_2$ (можно только одну зависимость).

6. На графике соответствующей эмпирической зависимости построить график полученной теоретической квадратической зависимости.

Вариант 1. В табл.1 дано распределение 55 однотипных предприятий по количеству выпускаемых изделий X (тыс. шт.) и полным затратам на их производство Y (млн руб.):

Таблица 1

X	Y						
	50 – 80	80 – 110	110 – 140	140 – 170	170 – 200	200 – 230	Итого
25 - 35	5	-	-	-	-	-	5
35 – 45	4	12	-	-	-	-	16
45 – 55	-	8	5	4	-	-	17
55 – 65	-	1	5	7	2	-	15
65 – 75	-	-	-	-	1	1	2
Итого	9	21	10	11	3	1	55

Используя соответствующее уравнение регрессии, оценить среднее значение полных затрат на производство 53 тыс. деталей.

Вариант 2. В табл. 2 дано распределение 100 заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т):

Таблица 2

X	Y						
	5 – 15	15 – 25	25 – 35	35 – 45	45 – 55	55 - 65	Итого
4 – 8	1	3	-	-	-	-	4
8 – 12	2	6	7	1	-	-	16
12 – 16	-	1	9	16	21	10	57
16 – 20	-	-	-	8	4	3	16
20 – 24	-	-	-	-	5	2	7
Итого	3	11	16	25	30	15	100

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 18 млн. руб.

Вариант 3. В табл. 3 дано распределение 60 предприятий по величине основных промышленно-производственных фондов X (млн руб.) и себестоимости продукции Y (руб.)

Таблица 3

X	Y					
	16 – 24	24 – 32	32 – 40	40 – 48	48 – 56	Итого
15 – 30	-	-	-	-	2	2
30 – 45	-	-	4	8	4	16
45 – 60	1	7	12	6	-	26
60 – 75	4	7	2	-	-	13
75 – 90	1	2	-	-	-	3
Итого	6	16	18	14	6	60

Используя соответствующее уравнение регрессии, определить среднюю себестоимость продукции для предприятий с основными промышленно-производственными фондами 65 млн. руб.

Вариант 4. В табл. 4 дано распределение 50 магазинов по уровню издержек обращения X (%) и по годовому объему товарооборота Y (млн руб.)

Таблица 4

X	Y					
	0,5 – 2 ,0	2,0 – 3,5	3,5 – 5,0	5,0 – 6,5	6,5 – 8,0	Итого
4 – 6	-	-	-	3	2	5
6 – 8	-	4	8	8	1	21
8 – 10	2	5	5	2	-	14
10 – 12	3	1	5	-	-	9
12 – 14	1	-	-	-	-	1
Итого	6	10	18	13	3	50

Используя соответствующее уравнение регрессии, определить среднее значение уровня издержек обращения в магазинах с годовым объемом товарооборота 4,2 млн руб.

Вариант 5. В табл. 5 дано распределение 200 совхозов по затратам труда X (чел.-дн. на 1 ц зерна) и себестоимости Y (руб. за 1 ц зерна)

Таблица 5

X	Y					
	7,25 – 9,25	9,25 – 11,25	11,25 – 13,25	13,25 – 15,25	15,25 – 17,25	Итого
0,4- 0,8	14	-	-	-	-	14
0,8 -1,2	22	10	-	-	-	32
1,2-1,6	-	38	30	10	-	78
1,6-2,0	-	6	30	12	2	50
2,0-2,4	-	-	4	8	8	20
2,4-2,8	-	-	-	-	6	6
Итого	36	54	64	30	16	200

Используя соответствующее уравнение регрессии, оценить средние затраты труда на получение 1 ц зерна себестоимостью 15 р. за 1 ц.

Вариант 6. В табл. 6 дано распределение 100 магазинов по величине товарооборота X (млн руб.) и размеру торговой площади магазина Y ($м^2$).

Таблица 6

X	Y					
	100 – 150	150 – 200	200 – 250	250 – 300	300 – 350	Итого
1,0-1,5	4	12	2	-	-	18
1,5-2,0	-	4	9	9	-	22
2,0-2,5	-	2	10	18	-	30
2,5-3,0	-	-	4	9	11	24
3,0-3,5	-	-	-	3	3	6
Итого	4	18	25	39	14	100

Используя соответствующее уравнение регрессии, найти среднюю величину товарооборота магазинов, имеющих торговую площадь 220 $м^2$.

Вариант 7. В табл. 7 дано распределение 100 проб руды, добытой на руднике, по содержанию окиси X (гр) и закиси железа Y (гр):

Таблица 7

X	Y						
	50–80	80–110	110–140	140–170	170 – 200	200 – 230	Итого
40 – 50	-	-	1	6	4	6	17
50 – 60	-	-	2	18	10	2	32
60 – 70	-	6	14	2	2	-	24
70 – 80	-	6	3	-	-	-	9
80 – 90	4	8	-	-	-	-	12
90-100	6	-	-	-	-	-	6
Итого	10	20	20	26	16	8	100

Используя соответствующее уравнение регрессии, оценить среднее количество окиси железа в руде, содержащей 25гр закиси железа.

Вариант 8. Данные о живом весе X (кг) и молочной продуктивности Y (кг) 80 коров приведены в табл. 8.

Таблица 8

X	Y					
	1250-1750	1750-2250	2250-2750	2750-3250	3250-3750	Итого
325-375	3	2	-	-	-	5
375-425	-	8	7	1	-	16
425-475	-	2	5	10	-	17
475-525	-	-	13	10	7	30
525-575	-	-	-	7	5	12
Итого	3	12	25	28	12	80

Используя соответствующее уравнение регрессии, оценить среднюю молочную продуктивность коров весом 450 кг.

Вариант 9. В табл. 9 дано распределение детей четырехлетнего возраста по росту X (см) и весу Y (кг):

Таблица 9

X	Y					
	15,5-16,5	16,5-17,5	17,5-18,5	18,5-19,5	19,5-20,5	Итого
98-100	2	3	-	-	-	5
100-102	3	6	4	-	-	13
102-104	1	4	13	5	-	23
104-106	-	1	14	10	2	27
106-108	-	-	10	8	5	23
108-110	-	-	-	6	3	9
Итого	6	14	41	29	10	100

Используя соответствующее уравнение регрессии, определить средний вес детей ростом 107 см.

Вариант 10. В табл. 10 дано распределение 80 совхозов по числу рабочих X (чел.) на 100 га сельскохозяйственных угодий и объему валовой продукции Y (тыс. руб.).

Таблица 10

X	Y					
	30 – 70	70 – 110	110 – 150	150 – 190	190 – 230	Итого
8 – 16	2	3	1	-	-	6
16 – 24	3	4	5	-	-	12
24 – 32	-	8	16	12	1	37
32 – 40	-	1	8	3	4	16
40 – 48	-	-	1	2	6	9
Итого	5	16	31	17	11	80

Используя соответствующее уравнение регрессии, оценить средний объем валовой продукции на 100 га сельскохозяйственных угодий в совхозе с 35 рабочими.

Вариант 11. В табл. 11 дано распределение заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т)

Таблица 11

X	Y					
	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60	Итого
15 –25	7	20	-	-	-	27
25 – 35	5	23	30	10	-	68
35 – 45	-	-	47	11	9	67
45 – 55	-	-	2	20	7	29
55 – 65	-	-	-	6	3	9
Итого	12	43	79	47	19	200

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 52 млн руб.

Вариант 12. Дана корреляционная табл. 12 результатов измерения перепада давления X (атм.) и расход нефти Y $\left(м^3/ч\right)$ в трубопроводе.

Таблица 12

X	Y					
	3000-3200	3200-3400	3400-3600	3600-3800	3800-4000	Итого
24 – 28	2	1	8	-	-	11
28 – 32	2	1	2	-	-	5
32 – 36	-	2	11	6	15	34
36 – 40	-	-	4	2	27	33
40 – 44	-	-	-	3	14	17
Итого	4	4	25	11	56	100

Используя соответствующее уравнение регрессии, оценить средний расход нефти при перепаде давления в 34 атм.

Вариант 13. Распределение растений по весу X каждого из них и по весу семян Y (гс) заданы корреляционной табл. 13.

Таблица 13

X	Y					
	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60	Итого
35 – 45	5	7	-	-	-	12
45 – 55	-	4	16	23	-	43
55 – 65	-	8	20	32	27	87
65 – 75	-	-	11	29	2	42
75 – 85	-	-	-	9	7	16
Итого	5	19	47	93	36	200

Используя соответствующее уравнение регрессии, оценить средний вес семян при весе растения 62 гс.

Вариант 14. В табл. 14 дано распределение прямоугольных плиток по длине Y (см) и по весу X (кгс).

Таблица 14

X	Y					
	5 – 7	7 – 9	9 –11	11 – 13	13 – 15	Итого
27,5-32,5	2	17	9	3	-	31
32,5-37,5	-	10	17	9	-	36
37,5-42,5	-	3	24	16	13	56
42,5-47,5	-	-	6	24	12	42
47,5-52,5	-	-	2	11	22	35
Итого	2	30	58	63	47	200

Используя соответствующее уравнение регрессии, оценить среднюю длину плитки весом 43 кгс.

Вариант 15. В табл. 15 дано распределение 60 однотипных предприятий по количеству выпускаемых изделий X (тыс. шт.) и полным затратам на их производство Y (млн руб.).

Таблица 15

X	Y					
	40 – 70	70 – 100	100 – 130	130 – 160	160 – 190	Итого
20 –30	6	-	-	-	-	6
30 – 40	4	12	-	-	-	16
40 – 50	-	8	6	4	-	18
50 – 60	-	2	6	7	3	18
60 – 70	-	-	-	-	2	2
Итого	10	22	12	11	5	60

Используя соответствующее уравнение регрессии, оценить среднее значение полных затрат на производство 34 тыс. деталей.

Вариант 16. В табл. 16 дано распределение 100 заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т):

Таблица 16

Y	X					
	10	15	20	25	30	35
30	2	6	-	-	-	-
40	-	4	4	-	-	-
50	-	-	7	35	8	-
60	-	-	2	10	8	-
70	-	-	-	5	6	3

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 33 млн руб.

Вариант 17. В табл. 17 дано распределение 100 заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т):

Таблица 17

Y	X					
	15	20	25	30	35	40
5	4	2	-	-	-	-
10	-	6	4	-	-	-
15	-	-	6	45	2	-
20	-	-	2	8	6	-
25	-	-	-	4	7	4

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 28 млн. руб.

Вариант 18. В табл. 18 дано распределение 100 заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т):

Таблица 18

Y	X					
	5	10	15	20	25	30
20	1	5	-	-	-	-
30	-	5	3	-	-	-
40	-	-	9	40	2	-
50	-	-	4	11	6	-
60	-	-	-	4	7	3

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 18 млн руб.

Вариант 19. В табл. 19 дано распределение 100 заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т):

Таблица 19

Y	X					
	10	15	20	25	30	35
6	4	2	-	-	-	-
12	-	6	2	-	-	-
18	-	-	5	40	5	-
24	-	-	2	8	7	-
30	-	-	-	4	7	8

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 21 млн руб.

Вариант 20. В табл. 20 дано распределение 100 заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т):

Таблица 20

Y	X					
	5	10	15	20	25	30
8	2	4	-	-	-	-
12	-	3	7	-	-	-
16	-	-	5	30	10	-
20	-	-	7	10	8	-
24	-	-	-	5	6	3

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 23 млн руб.

Вариант 21. В табл. 21 дано распределение прямоугольных плиток по длине Y (см) и по весу X (кгс).

Таблица 21

Y	X					
	2	7	12	17	22	27
10	2	4	-	-	-	-
20	-	6	2	-	-	-
30	-	-	3	50	-	-
40	-	-	1	10	6	-
50	-	-	-	4	7	3

Используя соответствующее уравнение регрессии, оценить среднюю длину плитки весом 18 (кгс).

Вариант 22. В табл. 22 дано распределение 100 заводов по объему основных производственных фондов X (млн руб.) и по суточной выработке продукции Y (т):

Таблица 22

Y	X					
	11	16	21	26	31	36
25	2	4	-	-	-	-
35	-	6	3	-	-	-
45	-	-	6	45	4	-
55	-	-	2	8	6	-
65	-	-	-	4	7	3

Используя соответствующее уравнение регрессии, оценить среднюю суточную выработку продукции на заводах с объемом основных фондов 33 млн руб.

Вариант 23. В табл. 23 дано распределение 80 однотипных предприятий по количеству выпускаемых изделий X (тыс. шт.) и полным затратам на их производство Y (млн руб.).

Таблица 23

Y	X					
	4	9	14	19	24	29
8	3	3	-	-	-	-
18	-	5	4	-	-	-
28	-	-	20	2	8	-
38	-	-	5	10	6	-
48	-	-	-	4	7	3

Используя соответствующее уравнение регрессии, оценить среднее значение полных затрат на производство 27 тыс. деталей.

Вариант 24. В табл. 24 дано распределение 103 однотипных предприятий по количеству выпускаемых изделий X (тыс. шт.) и полным затратам на их производство Y (млн руб.).

Таблица 24

Y	X					
	5	10	15	20	25	30
11	4	2	-	-	-	-
21	-	5	3	-	-	-
31	-	-	5	48	5	-
41	-	-	2	8	7	-
51	-	-	-	4	7	3

Используя соответствующее уравнение регрессии, оценить среднее значение полных затрат на производство 14 тыс. деталей.

Вариант 25. В табл. 25 дано распределение между объемом выполненных работ Y (млн руб.) и накладными расходами X (млн руб.).

Таблица 25

Y	X							
	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
100	-	-	-	-	1	1	2	2
110	-	-	-	1	3	3	5	2
120	-	-	-	2	6	9	2	1
130	-	-	1	8	14	4	1	-
140	-	1	6	15	5	2	-	-
150	-	9	10	7	1	-	-	-
160	1	6	5	2	-	-	-	-

Используя соответствующее уравнение регрессии, оценить средние накладные расходы для объема выполненных работ в 145 млн руб.

Вариант 26. В табл. 26 дана зависимость между относительным уровнем издержек обращения Y(%) и объемом сбыта товаров X (млн руб.).

Таблица 26

Y	X						
	1.22	1.23	1.24	1.25	1.26	1.27	1.28
4.1	-	-	-	-	1	3	1
4.2	-	-	1	2	3	3	-
4.3	-	-	1	6	7	1	-
4.4	-	1	3	10	2	1	-
4.5	-	1	6	5	1	-	-
4.6	1	3	4	3	1	-	-
4.7	1	2	1	1	-	-	-

Используя соответствующее уравнение регрессии, оценить среднее значение сбыта товаров при уровне издержек 4,43 %.

Вариант 27. В табл. 27 дана зависимость между относительным уровнем издержек обращения Y(%) и объемом сбыта товаров X (млн руб.).

Таблица 27

Y	X							
	4.3	4.4	4.5	4.6	4.7	4.8	4.9	5.0
3.9	1	3	2	1	-	-	-	-
4.0	-	1	5	2	1	-	-	-
4.1	-	1	3	8	4	1	-	-
4.2	-	-	1	2	10	3	-	-
4.3	-	-	-	1	3	6	4	1
4.4	-	-	-	-	-	1	3	1
4.5	-	-	-	-	-	-	1	2
4.6	-	-	-	-	-	-	-	1

Используя соответствующее уравнение регрессии, оценить среднее значение сбыта товаров при уровне издержек 4 %.

Вариант 28. В табл. 28 дана зависимость между относительным уровнем издержек обращения Y (%) и объемом сбыта товаров X (млн руб.).

Таблица 28

Y	X						
	11.1	11.2	11.3	11.4	11.5	11.6	11.7
1.2	1	2	3	2	1	-	-
1.3	1	2	4	4	2	1	-
1.4	1	2	6	7	4	3	1
1.5	1	3	6	9	7	2	1
1.6	-	1	3	6	6	3	1
1.7	-	1	2	4	5	3	1
1.8	-	-	1	2	5	3	1

Используя соответствующее уравнение регрессии, оценить среднее значение сбыта товаров при уровне издержек 1,62 %.

Вариант 29. В табл. 29 дана зависимость между относительным уровнем издержек обращения Y (%) и объемом сбыта товаров X (млн руб.).

Таблица 29

Y	X						
	1	2	3	4	5	6	7
0.12	1	-	-	-	-	-	-
0.13	3	1	-	-	-	-	-
0.14	-	6	1	-	-	-	-
0.15	-	1	6	2	-	-	-
0.16	-	-	1	9	1	-	-
0.17	-	-	-	2	7	1	-
0.18	-	-	-	1	1	5	2

Используя соответствующее уравнение регрессии, оценить среднее значение сбыта товаров при уровне издержек 0,154 %.

Вариант 30. В табл. 30 дано распределение совхозов по числу рабочих X (чел.) на 100 га сельскохозяйственных угодий и объему валовой продукции Y (тыс. руб.).

Таблица 30

Y	X								
	125	135	145	155	165	175	185	195	205
11	2	2	1	-	-	-	-	-	-
12	2	5	4	2	1	-	-	-	-
13	1	3	8	6	5	2	-	-	-
14	-	1	5	13	10	5	1	-	-
15	-	-	1	9	20	8	3	1	-
16	-	-	-	3	9	14	5	1	-
17	-	-	-	1	4	7	9	3	1
18	-	-	-	-	2	3	4	6	1
19	-	-	-	-	-	1	1	2	3

Используя соответствующее уравнение регрессии, оценить средний объем валовой продукции на 100 га сельскохозяйственных угодий в совхозе с 140 рабочими.

ПРИЛОЖЕНИЯ

ПРИЛОЖЕНИЕ 1

Таблица значений функции $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{z^2}{2}}dz$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1	2	3	4	5	6	7	8
0,00	0,0000	0,32	0,1255	0,64	0,2389	0,96	0,3315
0,01	0,0040	0,33	0,1293	0,65	0,2422	0,97	0,3340
0,02	0,0080	0,34	0,1331	0,66	0,2454	0,98	0,3365
0,03	0,0120	0,35	0,1368	0,67	0,2486	0,99	0,3389
0,04	0,0160	0,36	0,1406	0,68	0,2517	1,00	0,3413
0,05	0,0199	0,37	0,1443	0,69	0,2549	1,01	0,3438
0,06	0,0239	0,38	0,1480	0,70	0,2580	1,02	0,3461
0,07	0,0279	0,39	0,1517	0,71	0,2611	1,03	0,3485
0,08	0,0319	0,40	0,1554	0,72	0,2642	1,04	0,3508
0,09	0,0359	0,41	0,1591	0,73	0,2673	1,05	0,3531
0,10	0,0398	0,42	0,1628	0,74	0,2703	1,06	0,3554
0,11	0,0438	0,43	0,1664	0,75	0,2784	1,07	0,3577
0,12	0,0478	0,44	0,1700	0,76	0,2764	1,08	0,3599
0,13	0,0517	0,45	0,1736	0,77	0,2794	1,09	0,3621
0,14	0,0557	0,46	0,1772	0,78	0,2823	1,10	0,3643
0,15	0,0596	0,47	0,1808	0,79	0,2852	1,11	0,3665
0,16	0,0636	0,48	0,1844	0,80	0,2881	1,12	0,3686
0,17	0,0675	0,49	0,1879	0,81	0,2910	1,13	0,3708
0,18	0,0714	0,50	0,1915	0,82	0,2939	1,14	0,3729
0,19	0,0753	0,51	0,1950	0,83	0,2967	1,15	0,3949
0,20	0,0793	0,52	0,1985	0,84	0,2995	1,16	0,3770
0,21	0,0832	0,53	0,2019	0,85	0,3023	1,17	0,3790
0,22	0,0871	0,54	0,2054	0,86	0,3051	1,18	0,3810
0,23	0,0910	0,55	0,2088	0,87	0,3078	1,19	0,3830
0,24	0,0948	0,56	0,2123	0,88	0,3106	1,20	0,3849
0,25	0,0987	0,57	0,2157	0,89	0,3133	1,21	0,3869
0,26	0,1026	0,58	0,2190	0,90	0,3159	1,22	0,3883
0,27	0,1064	0,59	0,2224	0,91	0,3186	1,23	0,3907
0,28	0,1103	0,60	0,2257	0,92	0,3212	1,24	0,3925
0,29	0,1141	0,61	0,2291	0,93	0,3238	1,25	0,3944
0,30	0,1179	0,62	0,2324	0,94	0,3264		
0,31	0,1217	0,63	0,2357	0,95	0,3289		

Продолжение прил. 1

1	2	3	4	5	6	7	8
1,26	0,3962	1,59	0,4441	1,92	0,4726	2,50	0,4938
1,27	0,3980	1,60	0,4452	1,93	0,4732	2,52	0,4941
1,28	0,3997	1,61	0,4463	1,94	0,438	2,54	0,4945
1,29	0,4015	1,62	0,4474	1,95	0,4744	2,56	0,4948
1,30	0,4032	1,63	0,4484	1,96	0,4750	2,58	0,4951
1,31	0,4049	1,64	0,4495	1,97	0,4756	2,60	0,4953
1,32	0,4066	1,65	0,4505	1,98	0,4761	2,62	0,4956
1,33	0,4082	1,66	0,4515	1,99	0,4767	2,64	0,4959
1,34	0,4099	1,67	0,4525	2,00	0,4772	2,66	0,4961
1,35	0,4115	1,68	0,4535	2,02	0,4783	2,68	0,4963
1,36	0,4131	1,69	0,4545	2,04	0,4793	2,70	0,4965
1,37	0,4147	1,70	0,4554	2,06	0,4803	2,72	0,4967
1,38	0,4162	1,71	0,4564	2,08	0,4812	2,74	0,4969
1,39	0,4177	1,72	0,4573	2,10	0,4821	2,76	0,4971
1,40	0,4192	1,73	0,4582	2,12	0,4830	2,78	0,4973
1,41	0,4207	1,74	0,4591	2,14	0,4838	2,80	0,4974
1,42	0,4222	1,75	0,4599	2,16	0,4846	2,82	0,4976
1,43	0,4236	1,76	0,4608	2,18	0,4854	2,84	0,4977
1,44	0,4251	1,77	0,4616	2,20	0,4861	2,86	0,4979
1,45	0,4265	1,78	0,4625	2,22	0,4868	2,88	0,4980
1,46	0,4279	1,79	0,4633	2,24	0,4875	2,90	0,4981
1,47	0,4292	1,80	0,4641	2,26	0,4881	2,92	0,4982
1,48	0,4306	1,81	0,4649	2,28	0,4887	2,94	0,4984
1,49	0,4319	1,82	0,4656	2,30	0,4893	2,96	0,4985
1,50	0,4332	1,83	0,4664	2,32	0,4898	2,98	0,4986
1,51	0,4345	1,84	0,4671	2,34	0,4904	3,00	0,49865
1,52	0,4357	1,85	0,4678	2,36	0,4909	3,20	0,49931
1,53	0,4370	1,86	0,4686	2,38	0,4913	3,40	0,49966
1,54	0,4382	1,87	0,4693	2,40	0,4918	3,60	0,499841
1,55	0,4394	1,88	0,4699	2,42	0,4922	3,80	0,499928
1,56	0,4406	1,89	0,4706	2,44	0,4927	4,00	0,499968
1,57	0,4418	1,90	0,4713	2,46	0,4931	4,50	0,499997
1,58	0,4429	1,91	0,4719	2,48	0,4934	5,00	0,499999

ПРИЛОЖЕНИЕ 2

Таблица значений $t_\gamma = t(\gamma, n)$

γ / n	0,95	0,99	0,999	γ / n	0,95	0,99	0,999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,720	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,991	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

ПРИЛОЖЕНИЕ 3

Таблица значений $q = q(\gamma, n)$

γ / n	0,95	0,99	0,999	γ / n	0,95	0,99	0,999
5	1,37	2,67	5,64	20	0,37	0,58	0,88
6	1,09	2,01	3,88	25	0,32	0,49	0,73
7	0,92	1,62	2,98	30	0,28	0,43	0,63
8	0,80	1,38	2,42	35	0,26	0,38	0,56
9	0,71	1,20	2,06	40	0,24	0,35	0,50
10	0,65	1,08	1,80	45	0,22	0,32	0,46
11	0,59	0,98	1,60	50	0,21	0,30	0,43
12	0,55	0,90	1,45	60	0,188	0,269	0,38
13	0,52	0,83	1,33	70	0,174	0,245	0,34
14	0,48	0,78	1,23	80	0,161	0,226	0,31
15	0,46	0,73	1,15	90	0,151	0,211	0,29
16	0,44	0,70	1,07	100	0,143	0,198	0,27
17	0,42	0,66	1,01	150	0,115	0,160	0,211
18	0,40	0,63	0,96	200	0,099	0,136	0,185
19	0,39	0,60	0,92	250	0,089	0,120	0,162

ПРИЛОЖЕНИЕ 4

Критические точки распределения χ^2

Число степеней свободы r	Уровень значимости α					
	0,01	0,025	0,05	0,95	0,975	0,89
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26,1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	7,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63
20	37,6	34,2	31,4	10,9	9,59	8,26
21	38,9	35,5	32,7	11,6	10,3	8,90
22	40,3	36,8	33,9	12,3	11,0	9,54
23	41,6	38,1	35,2	13,1	11,7	10,2
24	43,0	39,4	36,4	13,8	12,4	10,9
25	44,3	40,6	37,7	14,6	13,1	11,5
26	45,6	41,9	38,9	15,4	13,8	12,2
27	47,0	43,2	40,1	16,2	14,6	12,9
28	48,3	44,5	41,3	16,9	15,3	13,6
29	49,6	45,7	42,6	17,7	16,0	14,3
30	50,9	47,0	43,8	18,5	16,8	15,0

СПИСОК ЛИТЕРАТУРЫ

Основная

1. Гмурман, В.Е. Теория вероятностей и математическая статистика / В.Е.Гмурман. – М.: Высшее образование : Юрайт, 2009. - 479с.

2. Гмурман, В.Е. Руководство к решению задач по теории вероятностей и математической статистике / В.Е.Гмурман. –М.: Высшее образование, 2009.- 404 с.

3. Данко, П.Е. Высшая математика в упражнениях и задачах /П.Е.Данко, А.Г.Попов, Т.Я.Кожевникова. - Ч.1, 2. –М.: ОНИКС 21век; Мир и Образование, 2006.-416с.

4. Феллер, В. Введение в теорию вероятностей и ее приложения/ В.Феллер. – Т.1,2. -М.: Мир, 1984.

Дополнительная

5. Вентцель, Е.С. Теория вероятностей и ее инженерные приложения / Е.С.Вентцель, Л.А. Овчаров.– М.: Высшая школа, 2007.-479с.

6. Кремер, Н.Ш. Теория вероятностей и математическая статистика / Н.Ш Кремер. – М.: ЮНИТИ- ДАНА, 2002.-543с.

Учебные пособия кафедры

7. Учебно -методическое пособие по математической статистике/ Т.В.Умергалина, З.Р Хакимова. – Уфа: УГНТУ, 2004.

8. Практикум по математической статистике / Т.В.Умергалина, З.Р. Хакимова. – Уфа: УГНТУ, 2003.

9. Расчетные задания по математической статистике / Т.В.Умергалина, З.Р. Хакимова. – Уфа: УГНТУ, 2007.

10. Функции нескольких переменных: учебно-методическое пособие / Р.Г.Гимаев, Т.В.Умергалина.- Уфа: УГНТУ, 2005 (Раздел «Метод наименьших квадратов»).

Учебное издание

Бахтизин Рамиль Назифович, Галиуллин Марат Мидхатович, Галиакбарова Эмилия Вильевна, Гимаев Роберт Гильмутдинович, Зарипов Раиль Муталлапович, Исламгулова Галия Файзеевна, Ковалева Элла Александровна, Лазарев Владимир Анатольевич, Майский Равиль Анварович, Мухаметзянов Ирек Зирягович, Нагаева Зиля Мунировна, Сахарова Лидия Александровна, Сокова Инна Александровна, Сулейманов Игорь Нугуманович, Умергалина Татьяна Васильевна, Фаткуллин Николай Юрьевич, Хайбуллин Радик Яруллович, Хакимова Зульфия Разифовна, Чернятьева Рита Раисовна, Юлдыбаев Лев Хадиевич, Шамшович Валентина Федоровна, Шварева Елена Николаевна, Якупов Вагизьян Минигалиевич, Янчушка Анна Павловна, Абзалимов Рамиль Рафикович, Акмадиева Танслу Рифхатовна, Аносова Елизавета Петровна, Ахтямов Наиль Тагирович, Байрамгулова Римма Сагитовна

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

2-е издание, исправленное и дополненное