И. В. Ребро, В. А. Носенко, Н. Н. Короткова

Прикладная математическая статистика для технических специальностей

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
ВОЛЖСКИЙ ПОЛИТЕХНИЧЕСКИЙ ИНСТИТУТ (ФИЛИАЛ)
ГОСУДАРСТВЕННОГО ОБРАЗОВАТЕЛЬНОГО УЧРЕЖДЕНИЯ ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ «ВОЛГОГРАДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

И.В. Ребро, В.А. Носенко, Н.Н. Короткова

Прикладная математическая статистика для технических специальностей

Учебное пособие

Волгоград 2011

УДК 519.2

Р е ц е н з е н т ы:

доктор физ.-мат. наук, доцент Кульков В.Г.

доктор тех.наук, профессор

Печатается по решению редакционно-издательского совета
Волгоградского государственного технического университета

Ребро, И.В.

Прикладная математическая статистика для технических специиальностей: учеб. пособие / И.В. Ребро, В.А. Носенко, Н.Н. Короткова; ВПИ (филиал) ВолгГТУ. – Волгоград: ИУНЯ ВолгГТУ, 2011. – 149 с.

ISBN

Учебное пособие предназначено для студентов специальностей 120100 «Технология машиностроения» и 200503 «Стандартизация и сертификация» высших технических учебных заведений. Пособие содержит краткие теоретические сведения, решения типовых примеров и задания для самостоятельной работы. Кроме задач алгоритмически-вычислительного характера в пособии рассмотрены задачи иллюстрирующие применение теории в практической деятельности и способствующие более глубокому усвоению темы, развивающие навыки самостоятельной работы и формированию математического мышления.

ISBN

ОГЛАВЛЕНИЕ

Предисловие

В начале прошлого века характеризовали качество выпускаемой продукции, как соответствие продукции требованиям стандартов. В промышленности активно внедрялась система Тейлора, и для поставки качественной продукции было достаточно приемочного контроля.

В настоящее время технический прогресс глобально изменил жизни, приоритеты и мышление человечества. Еще в недалеком прошлом человечество мечтало о компьютерных технологиях, а сегодня они являются необходимым средством почти в каждой деятельности. И что особенно важно процесс совершенствования и изменения технических средств, а также человека как личности, продолжается.

Новые разработки дают не только новые возможности по изготовлению и усовершенствованию продукции, но и направляют нас изменить подходы к жизни и науке.

Так, с развитием производства, ростом объемов выпускаемой продукции, важным техническим и экономическим показателем становится процент брака. Поскольку цена брака переносится на качественную продукцию, то на рынке выигрывают фирмы с минимальным процентом брака. Необходимость учета, анализа, выводов и принятия мер приводит к активному использованию статистических методов. Вводится статистический приемочный контроль, в частности, строятся гистограммы, отражающие распределение объемов выпускаемой продукции по отношению к устанавливаемым браковочным признакам. Математический анализ гистограмм переводит их в дисперсионные кривые, в том числе характеризующие среднее квадратическое отклонение от номинального значения [Ю.А. Богомолов].

По мере совершенствования концепций качества изобретались новые подходы, критерии оценки, методологии и т.п., но в развитии реализовывался накопительный принцип - не отказываться от ранее созданного. Сегодня, статистический метод оценки качества преобразовался в методологию шести сигм. Это, значит, что обеспечивается такая стабильность производства, в которой количество несоответствий возможно на уровне 34 от 10 000 000 (при выпуске десяти миллионов изделий из них бракованных может быть 34)[Ю.А. Богомолов].

Применение статистических методов на современном этапе производства является необходимым многофункциональным инструментом, который характеризует концепцию качества продукции, соответствующего требованиям стандарта при определенной стабильности производственных процессов.

Таким образом, авторами была выдвинута основная задача курса прикладной математической статистики, которая заключается в овладении знаниями общих основ статистической науки, развитии способностей по организации и проведению статистических исследований, анализу полученных данных и прогнозированию дальнейших результатов.

Целью данного учебного пособия является научить применять прикладные методы математической статистики при необходимости в производственной практике. Поэтому изложенный материал содержит общие критерии, оценки и частные критерии, оценки, которые наиболее часто встречаются в производственном процессе. Также, для наглядности и убедительности, изложение материала дополнено демонстрационными техническими задачами.

Предложенное учебное пособие необходимо для студентов изучающих соответствующую дисциплину, а также полезно магистрантам, аспирантам и всем, желающим на современном уровне проводить эксперименты и обрабатывать их результаты. Настоящее учебное пособие направлено на развитие исследовательских способностей в условиях современного технического прогресса.

Глава 1. Статистика как наука

Ядро математической статистики является теория статистики, которая обеспечивает теоретическую и методологическую основу подготовки будущих специалистов. Теоретическую основу статистики составляют понятия и категории, в совокупности которых выражаются основные принципы данной науки. Методологическую основу статистики составляют положения и принципы диалектического метода познания.

1.1. Понятие статистики

В связи с развитием современной техники особую важность приобрели многочисленные вопросы повышения эффективности различного рода устройств. Комплексная автоматизация производственных процессов ставит перед управляющими устройствами исключительно ответственные задачи, которые должны выполняться безупречно на протяжении всего периода работы автоматической линии, автоматизированного цеха или предприятия. Перерыв в работе управляющего устройства может привести не только к ухудшению качества производимой продукции или к полному прекращению производственного процесса, но и к весьма серьезным авариям, выходящим за локальные рамки предприятия.

Для исследования и прогнозирования возникновения проблем, связанных с качеством и надежностью, используются методы статистической обработки данных. Например, если мы изготовили в определенных условиях из одной и той же партии сырья большое число определенного типа изделий и затем собрали статистические данные о длительности их бесперебойной работы, то выясняется: длительность безотказной работы изделия имеет значительный разброс и в отношении каждого определенного изделия нет возможности точно предсказать длительность его службы. В то же время относительно больших партий этих изделий можно делать достаточно определенные выводы о продолжительности их работоспособности или о причинах поломок. Все это помогает определить статистика.

В настоящее время под термином «статистика» понимают следующее.

Статистика – это самостоятельная общественная наука, имеющая свой предмет исследования и свои специфические методы.

Часто слово «статистика» употребляется в качестве эквивалента для словосочетания статистические методы. Статистические методы применяют тогда, когда из большого массива данных требуется выделить полезную информацию.

Основные черты статистической науки:

1. Статистика как наука заключается в исследовании массовых физических явлений и процессов. Задача статистического исследования состоит в получении обобщающих показателей и выявлении закономерностей, которые проявляются лишь в большей массе явлений через преодоление случайности, свойственной единичным элементам. Объектом статистического исследования является статистическая совокупность. Статистическая совокупность – это множество единиц, обладающих массовостью, однородностью, определенной целостностью, взаимозависимостью состояний отдельных единиц и наличием вариации. Единицы статистической совокупности характеризуются общими свойствами, именуемыми признаками.
2. Статистика как наука заключается в том, что она изучает количественную сторону явлений и процессов в конкретных условиях. Количественная характеристика выражается через числа, которые называются статистическими показателями. Статистический показатель отражает результат измерения единиц совокупности и совокупности в целом.
3. Статистика как наука заключается в том, что она характеризует структуру явлений. Структура – это внутреннее строение явлений, в данном случае статистического множества.
4. Статистика как наука заключается в выявлении изменения в пространстве (т.е. в статике, эти изменения выявляются анализом структуры явлений) и изменения во времени (т.е. в динамике).
5. Статистика как наука заключается в выявлении связей необходимых для того, чтобы воздействовать на явления с целью их изменения.

Прикладная статистика нацелена на решение реальных задач. Поэтому в ней возникают новые постановки математических задач анализа статистических данных, развиваются и обосновываются новые методы. Обоснование часто проводится математическими методами, то есть путем доказательства теорем. Большую роль играет методологическая составляющая - как именно ставить задачи, какие предположения принять с целью дальнейшего математического изучения.

1.2. Статистическая методология

Под статистической методологией понимается система приемов, способов и методов, направленных на изучение количественных закономерностей, проявляющихся в структуре, динамике и взаимосвязях явлений.

Статистическое исследование состоит из трех основных стадий:

- статистического наблюдения;

- первичной обработки, сводки и группировки результатов наблюдений;

- анализ полученных сводных материалов.

Прохождение каждой стадии исследования связано с использованием специальных методов, объясняемых содержанием выполняемой работы.

Первый метод – метод массовых наблюдений.

Начальной стадией статистического исследования является статистическое наблюдение, то есть научно организованный сбор сведений об изучаемых процессах и явлений.

Полученные в результате статистического наблюдения данные являются исходным материалом для выполнения последующих этапов статистического исследования. Результатом статистического наблюдения является получение данных, характеризующих каждую единицу наблюдения. Эти результаты необходимо определенным образом обработать с тем, чтобы из статистического «сырья» выявить статистические данные. Такая обработка является следующей после наблюдения стадий статистического исследования и представляет собой сводку исходных данных для получения обобщающих характеристик исследуемого процесса или явления, проводимую с помощью использования метода группировок и таблиц.

Второй метод – метод статистических группировок и таблиц.

Включает в себя вторую стадию статистического исследования, которая представляет собой комплекс последовательных действий по обобщению конкретных единичных фактов, образующих совокупность в целях выявления типичных признаков и закономерностей, присущих изучаемому явлению в целом. Важнейшим специфическим методом на этой стадии является метод группировок. Статистическая сводка включает в себя распределение исходных данных по группам, качественно однородным по одному или нескольким признакам, и получение групповых итогов.

Для правильного выделения качественно однородных групп следует выбирать основные, наиболее существенные для данного явления или процесса признаки. В зависимости от числа и вида признаков, решаемых задач и исходных данных группировки подразделяются на следующие виды: простые и комбинационные; типологические, структурные и аналитические; одномерные и многомерные; первичные и вторичные.

Одним из этапов группировки является построение рядов распределения. Результаты статистической группировки и сводки излагаются в виде статистических таблиц.

Третий метод – метод анализа с помощью обобщающих показателей.

Статистический анализ является заключительной стадией статистического исследования.

Характерным для статистических методов на этой стадии является применение обобщающих показателей: абсолютных, относительных, средних величин и т.п.

1.3. Шкалы измерений

Теория измерений является одной из составных частей прикладной статистики. Она входит в состав статистики объектов нечисловой природы.

Используют следующие типы шкал измерений: номинальная, порядковая (ординальная), интервальная, относительная (шкала отношения). Соответственно имеются четыре типа переменных: номинальная, порядковая (ординальная), интервальная и относительная.

Номинальные переменные используются только для качественной классификации. Это означает, что данные переменные могут быть измерены только в терминах принадлежности к некоторым существенно различным классам, при этом вы не сможете определить количество или упорядочить эти классы. Типичными примерами номинальных переменных являются завод-изготовитель, тип изделия, признак и т. д. Часто номинальные переменные называются категориальными. Близким к ним являются категоризованные переменные, то есть переменные, искусственно превращенные в категориальные.

Порядковые переменные позволяют ранжировать (упорядочить) объекты, если указано, какие из них в большей или меньшей степени обладают качеством, выраженным данной переменной. Однако они не позволяют определить «на сколько больше» или «на сколько меньше» данного качества содержится в переменной. Порядковые переменные иногда также называют ординальными. Само расположение шкал в порядке возрастания их информативности X: номинальная, порядковая, интервальная. Например, можно сказать, что измерения в номинальной шкале предоставляют меньше информации, чем в порядковой шкале, а в порядковой X меньше, чем в интервальной. Однако невозможно придать термину «меньше» точный количественный смысл или сравнить между собой эти различия. При оценке качества продукции и услуг, в так называемой квалиметрии (буквальный перевод: измерение качества) популярны порядковые шкалы. А именно, единица продукции оценивается как годная или не годная. При более тщательном анализе используется шкала с тремя градациями: есть значительные дефекты - присутствуют только незначительные дефекты - нет дефектов. Иногда применяют четыре градации: имеются критические дефекты (делающие невозможным использование) - есть значительные дефекты - присутствуют только незначительные дефекты - нет дефектов. Аналогичный смысл имеет сортность продукции - высший сорт, первый сорт, второй сорт, ...

Интервальные переменные позволяют не только упорядочивать объекты измерения, но и численно выражать и сравнивать различия между ними. Такого рода переменные часто возникают в естественных науках, при снятии показателей с физических приборов, в медицине и т. д. Например, температура, измеренная в градусах Фаренгейта или Цельсия, образует интервальную шкалу. Вы можете не только сказать, что температура 40 градусов выше, чем температура 30 градусов, но и то, что увеличение температуры с 20 до 40 градусов вдвое больше увеличения температуры от 30 до 40 градусов.

Относительные переменные очень похожи на интервальные переменные. В дополнение ко всем свойствам переменных, измеренных в интервальной шкале, их характерной чертой является наличие определенной точки абсолютного нуля, таким образом, для этих переменных являются обоснованными утверждения типа: x в два раза больше, чем y. Например, температура по Кельвину образует шкалу отношения, и вы можете не только утверждать, что температура 200 градусов выше, чем 100 градусов, но и то, что она вдвое выше. Интервальные шкалы (например, шкала Цельсия) не обладают данным свойством шкалы отношения. Однако в большинстве статистических процедур не делается тонкого различия между свойствами интервальных шкал и шкал отношения.

1.4. Простейшие описательные статистики

Квантиль (термин был впервые использован Кендаллом в 1940 г.).

Квантилью порядка р ($0<p<1$) называется такое число x_p, что вероятность события $P(X< x_p)$ равна заданной величине p, то есть в этой точке функция распределения принимает значение p ($F(x_p)= p$) или имеет место «скачок» со значения, меньшего чем p, до значения, большего чем p.

Математическое ожидание (среднее выборочное).

Математическим ожиданием $M[X]$ дискретной случайной величины X называется сумма произведений всех её значений на соответствующие им вероятности: $M[X]=\sum_i x_i\, p_i$.

Математическое ожиданием $M[X]$ непрерывной случайной величины X с плотность распределения $f(x)$ вычисляется по формуле: $M[X]=\int_{-\infty}^{+\infty} x\, f(x)\, dx$.

Свойства:

1. $M[C]=C$, *где* $C=const$.
2. $M[C\cdot X]=C\cdot M[X]$.

$3. M[X+Y]=M[X]+M[Y].$

$4. M[X\cdot Y]=M[X]\cdot M[Y],$ *где* $X,Y-$ *независимые случайные величины.*

Дисперсия.

Дисперсией случайной величины X называется математическое ожидание квадрата ее отклонения от математического ожидания: $D[X]=M\left[(X-M[X])^2\right].$

Замечание. На практике дисперсию удобнее вычислять по формуле $D[X]=M\left[X^2\right]-M[X]^2$, где для дискретной величины X $M\left[X^2\right]=\sum_i x_i^2\, p_i$, для непрерывной $M\left[X^2\right]=\int_{-\infty}^{+\infty} x^2\, f(x)\,dx$.

Свойства:

$1. D[C]=0$, *где* $C=const$

$2. D[C\cdot X]=C^2\cdot D[X]$

$3. D[X\pm Y]=D[X]+D[Y]$

$4. D[X\cdot Y]=D[X]\cdot D[Y]+M^2[Y]\cdot D[X]+M^2[X]\cdot D[Y],$

где $X,Y-$ *независимые случайные величины*

Среднее квадратическое отклонение (стандартное отклонение).

Средним квадратическим отклонением или стандартным отклонением случайной величины X называется квадратный корень из ее дисперсии: $\sigma[X]=\sqrt{D[X]}$.

Свойства:

$1. \sigma[C]=0$, *где* $C=const$

$2. \sigma[C\cdot X]=|C|\cdot\sigma[X]$

$3. \sigma[C+X]=\sigma[X]$

Медиана.

Медиана это квантиль порядка p=0,5. Медиана определяется как серединное значение в ранжированном ряду данных. Это значит, что по обе стороны от неё расположено ровно по половине данных. Применительно к кривой распределения медиана представляет собой такую точку на оси абсцисс, что вертикальная кривая $x=M_e(X)$, проходящая через неё, делит площадь под кривой на две равные части.

Мода.

Мода (термин был впервые введен Пирсоном, 1894) - это наиболее часто встречающееся (наиболее модное) в распределении значение переменной. Если к данным таблицы распределения подобрать теоретическую кривую распределения, то мода равна абсциссе точки, имеющей максимальную для этой кривой ординату.

Замечание.

1. Если распределение имеет несколько мод, то говорят, что оно мультимодально или полимодально (при этом только одно значение удовлетво-

ряет определению моды, и это значение носит название наибольшей моды).

2. На величину среднего выборочного влияют все значения ряда, поэтому появление “выбросов”, то есть данных, значения которых находятся далеко от центра группирования, может существенно исказить величину среднего выборочного, особенно в выборках небольшого объёма. В таких ситуациях, когда речь идёт о выборках из непрерывных распределений, более предпочтительной является медиана, которая нечувствительна к “выбросам”. К “выбросам” нечувствительна и мода, но если число значений в выборке невелико, она непоказательна, особенно когда данные не обладают явно выраженной тенденцией группироваться у центра.

Асимметрия, или коэффициент асимметрии («скошенности»).

Асимметрия это отношение математического ожидания третьей степени от разности случайной величины X и её среднего значения к третьей степени среднего квадратического отклонения случайной величины: $A = \dfrac{M[X - MX]^3}{\sigma_X^3}$.

Замечание.

1. Асимметрия, или коэффициент асимметрии является мерой несимметричности распределения. Если этот коэффициент значительно отличается от 0, распределение является асимметричным.

2. Распределения, скошенные влево, имеют положительную асимметрию, а скошенные вправо-отрицательную. Для симметричных распределений, для которых среднее и медиана совпадают, асимметрия равна нулю.

Эксцесс, или коэффициент эксцесса («островершинности»).

Оценка эксцесса, или выборочный эксцесс – это отношение математического ожидания четвертой степени от разности случайной величины X и её среднего значения к четвёртой степени среднего квадратического отклонения случайной величины: $Ex = \dfrac{M[X - MX]^4}{\sigma_X^4}$.

Замечание.

1. Величина эксцесса для нормальной кривой распределения, играющей в статистике и в теории вероятностей важную роль, равна 3. Заострённость этой кривой принимают за стандарт, и поэтому в качестве показателя эксцесса используют $E_x - 3$. Кривые, более островершинные, чем нормальная, обладают положительным эксцессом, более плосковершинные - отрицательным.

2. Эксцесс может принимать очень большие значения, но он не может быть меньше единицы.

Глава 2. Распределения вероятностей случайных величин

Распределение значений случайной величины по вероятности их появления характеризуется интегральной функцией распределения и плотностью распределения. Для описания функции и плотности распределения пользуются специальными мерами, позволяющими охарактеризовать положение, форму и другие их особенности. Если известен закон распределения вероятностей случайной величины, то можно решать многие задачи, возникающие при статистическом анализе экспериментальных данных.

2.1. Непрерывные распределения

2.1.1. Нормальное распределение

Нормальное распределение дает хорошую модель для реальных явлений, в которых:

1) имеется сильная тенденция данных группироваться вокруг центра;

2) положительные и отрицательные отклонения от центра равновероятны;

3) частота отклонений быстро падает, когда отклонения от центра становятся большими.

Примеры: распределение размеров изделия, средней температуры и т.д.

Свойства

Параметры	a, σ
Плотность распределения	$f(x,a,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{\left(-\frac{(x-a)^2}{2\sigma^2}\right)}$
Функция распределения	$F(x,a,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{\left(-\frac{(t-a)^2}{2\sigma^2}\right)} dt$
Среднее выборочное	$M(x)=a$
Дисперсия	$D(x)=\sigma^2$
Стандартное отклонение	σ

Коэффициент вариации	$V=\sigma/a$
Коэффициент асимметрии	$A=0$
Коэффициент эксцесса	$E=3$
Мода	$M_0=a$
Медиана	$M_e=a$

Замечание. Распределение симметрично относительно точки $x=a$.

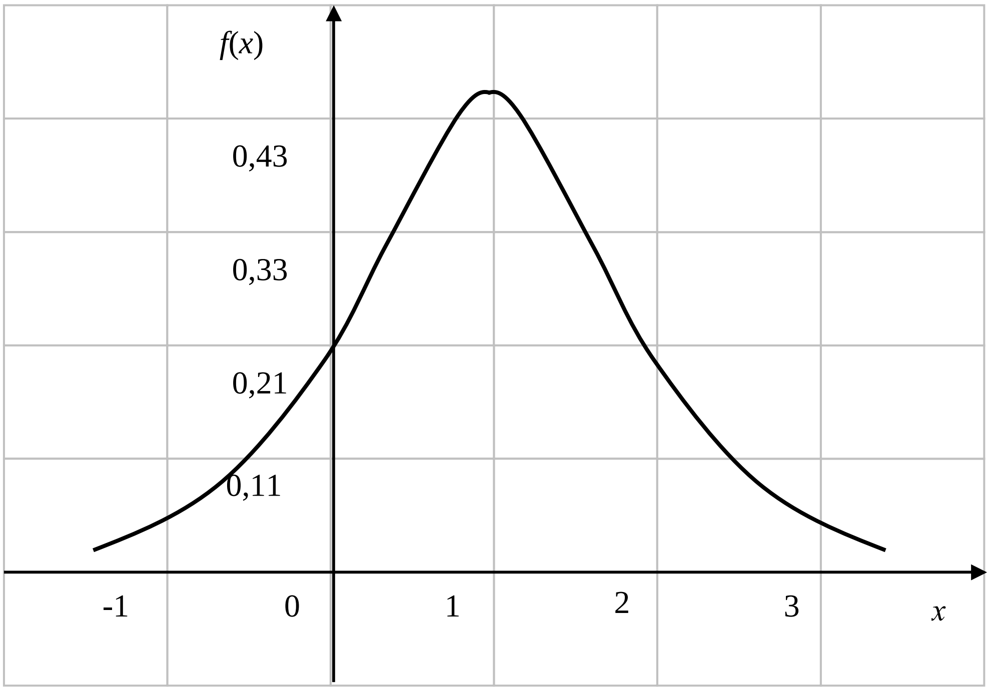


Рисунок 1 – График плотности нормального распределения ($\sigma^2=1,\ a=1$)

Замечание. Визуально график нормальной плотности образует коло-колоообразную кривую.

Для удобства в практических приложениях применяется нормированная случайная величина $z=\dfrac{x-a}{\sigma}$, распределение которой называется стандартным нормальным с нулевым средним и единичной дисперсией.

p-квантиль нормально распределенной случайной величины связана с квантилью u_p^c случайной величины, имеющей стандартное нормальной распределение, соотношением $u_p=a+u_p^c\sigma$.

Замечание. В силу симметричности нормального распределения $u_p^c=-u_{1-p}^c$.

Значения функций и квантилий приведены таблицах, но таблицы всегда ограничены, поэтому приведем некоторые известные аппроксимации:

- аппроксимация 1: $F(z) = 1 - \left(\sqrt{2\pi} e^{z^2/2}\right)^{-1} \sum_{i=1}^{3} a_i \lambda^i$, где

$\lambda = (1 + bz)^{-1}$; $b = 0{,}33267$; $a_1 = 0{,}4361836$; $a_2 = -0{,}1201676$; $a_3 = 0{,}937298$.

Абсолютная погрешность $\le 10^{-5}$.

- аппроксимация 2: $F(z) = 1 - \left(\sqrt{2\pi} e^{z^2/2}\right)^{-1} \sum_{i=1}^{5} a_i \lambda^i$, где

$\lambda = (1 + bz)^{-1}$; $b = 0{,}2316419$; $a_1 = 0{,}31938153$; $a_2 = -0{,}35656378$; $a_3 = 1{,}7814779$;
$a_4 = -1{,}821256$; $a_5 = 1{,}3302744$.

Абсолютная погрешность $\le 10^{-7}$.

- аппроксимация 3: $F(z) = 1 - 0{,}852 e^{-\left(\frac{z+1{,}5774}{2{,}0637}\right)^{2{,}34}}$, где $z > 0$,

при $z < 0$ $F(-z) = 1 - F(z)$.

Абсолютная погрешность $\le 0{,}0002$.

- аппроксимация 4: $u_p = -\left[-2\ln\left(\sqrt{2\pi}\, p\right)\right]^{1/2}$, где $p \le 0{,}01$.

- аппроксимация 5: $u_p = t - \dfrac{\sum_{i=0}^{2} c_i t^i}{1 + \sum_{i=1}^{3} d_i t^i}$, где

$t = \left[-2\ln(1-p)\right]^{1/2}$; $c_0 = 2{,}515517$; $c_1 = 0{,}802853$; $c_2 = 0{,}010328$;
$d_1 = 1{,}432788$; $d_2 = 0{,}189269$; $d_3 = 0{,}001308$.

Абсолютная погрешность $\le 0{,}00045$.

- аппроксимация 6: $u_p = 4{,}91\left(p^{0{,}14} - (1-p)^{0{,}14}\right)$.

Относительная погрешность $\le 0{,}3\%$.

Для многих приложений требуется знание математического ожидания i-й порядковой статистики в выборке объема n из стандартного нормального распределения, то есть математического ожидания i-го по номеру члену выборки, упорядоченной по возрастанию. Тогда используется факт: математическое ожидание i-го по величине члена выборки из стандартного нормального распределения $M(z_i) = u_{p'}$, где $p' = \dfrac{i - 3/8}{n + 1/4}$.

Задача 1. Вычислить значение функции распределения вероятностей в точке x=200, если случайная величина распределена нормально со средним a=100 и дисперсией σ^2=2500.

Решение.

Значение нормированной переменной определим по формуле: $z = \frac{x-a}{\sigma} = \frac{200-100}{50} = 2.$ *Таким образом, смысл задачи состоит в нахождении значения* $F(z) = P(z \le 2)$.

Табличное значение F(z)=0,97725 (приложение 1).

Воспользуемся аппроксимацией:

- *аппроксимация 1 -* $\lambda = \frac{1}{1+0{,}332567 \cdot 2} = 0{,}60047798$*;*

$$\left(\sqrt{2\pi}e^{\frac{z^2}{2}}\right)^{-1} = \left(2{,}506628e^2\right)^{-1} = 0{,}05399\text{; } \sum_{i=1}^{3} a_i \lambda^i = 0{,}421529962.$$

Получаем, $F(2) = 1 - 0{,}05399 \cdot 0{,}421529962 = 0{,}97724119$.

Относительная погрешность $\delta = \frac{0{,}97725 - 0{,}97724119}{0{,}97725} = 0{,}0009\%$.

- *аппроксимация 2 -* $\lambda = \frac{1}{1+0{,}2316419 \cdot 2} = 0{,}683394431$*;*

$$\left(\sqrt{2\pi}e^{\frac{z^2}{2}}\right)^{-1} = \left(2{,}506628e^2\right)^{-1} = 0{,}05399\text{; } \sum_{i=1}^{3} a_i \lambda^i = 0{,}421367913.$$

Получаем, $F(2) = 1 - 0{,}05399 \cdot 0{,}421367913 = 0{,}977249939$.

Относительная погрешность $\delta = \frac{0{,}97725 - 0{,}977249939}{0{,}97725} = 0{,}000006\%$.

Задача 2. Для случайной величины, имеющей стандартное нормальное распределение, найти значение, вероятность превышения которого равна 0,05.

Решение.

В задаче требуется найти значение u_p*:* $P(x \le u_p) = 0{,}95$ *или* $P(x > u_p) = 0{,}05$.

По определению u_p *- верхняя 95%-я квантиль случайной величины, табличное значение его равно* $u_{0,95} = 1{,}64485363$ *(приложение 2).*

Для самостоятельного расчета воспользуемся аппроксимациями:

- *аппроксимация 4 – Так как эта аппроксимация применима только для значений* $p \le 0{,}01$, *воспользуемся соотношением* $u_p = -u_{1-p}$ *и будем искать квантиль* $u_{1-p} = u_{0,95} = -u_{0,05}$.

Получаем $u_{0,05} = -\left[-2\ln\left(\sqrt{2\pi}\cdot 0{,}05\right)\right]^{1/2} = -\left[-2\ln(0{,}125331413)\right]^{1/2} = -2{,}038035201$.

Таким образом, $u_{0,95} = 2{,}038035201$.

Относительная погрешность $\delta = 23{,}95\%$ *очень велика.*

- *аппроксимация 5 – Вычислим:* $t = \left[-2\ln(1-0{,}95)\right]^{1/2} = 2{,}447746831$,

$\sum\limits_{i=0}^{2} c_i t^i = 4{,}542577732$, $\sum\limits_{i=1}^{3} d_i t^i = 4{,}66028338$.

Тогда $\dfrac{4{,}542577732}{1+4{,}66028338} = 0{,}80253539$.

Получаем $u_{0,95} = 2{,}447746831 - 0{,}80253539 = 1{,}64521144$.

Относительная погрешность $\delta = 0{,}02\%$.

Задача 3. Найти математические ожидания 3-й, 7-й и 9-й порядковых статистик в выборке объема *n*=15 из стандартного нормального распределения.

Решение.

Вычислим необходимые вероятности:

- *для третьей порядковой статистики* $p_3 = \dfrac{3-3/8}{15+1/4} = 0{,}172131148$;

- *для седьмой порядковой статистики* $p_7 = \dfrac{7-3/8}{15+1/4} = 0{,}434426230$;

- *для девятой порядковой статистики* $p_9 = \dfrac{9-3/8}{15+1/4} = 0{,}565573770$.

Воспользуемся аппроксимацией 6, получаем:

$M(z_3) = 4{,}91\left(0{,}172131148^{0,14} - (1-0{,}172131148)^{0,14}\right) = -0{,}9438796$;

$M(z_7) = 4{,}91\left(0{,}434426230^{0,14} - (1-0{,}434426230)^{0,14}\right) = -0{,}1643847$;

$M(z_9) = 4{,}91\left(0{,}565573770^{0,14} - (1-0{,}565573770)^{0,14}\right) = 0{,}1643847$.

2.1.2. Равномерное распределение.

Равномерное распределение применяется при описании переменных, у которых каждое значение равновероятно, то есть значения переменной равномерно распределены в некоторой области.

Равномерное распределение дает вероятность того, что наблюдение будет лежать в определенном интервале, когда вероятность прямо пропорциональна длине интервала.

Равномерное распределение используется для генерирования случайных чисел.

Свойства

Параметры	$a,\ b$
Плотность распределения	$f(x,a,b)=\begin{cases}(b-a)^{-1}, & a\le x\le b;\\ 0, & x<a \text{ и } x>b.\end{cases}$
Функция распределения	$F(x,a,b)=\begin{cases}0, & x\le a;\\ \dfrac{x-a}{b-a}, & a<x<b;\\ 1, & x\ge b.\end{cases}$
Среднее выборочное	$M(x)=\dfrac{b+a}{2}$
Дисперсия	$D(x)=\dfrac{(b-a)^2}{12}$
Стандартное отклонение	$\sigma=\sqrt{D(x)}$
Коэффициент вариации	$v=\dfrac{1}{\sqrt{3}}\dfrac{b-a}{b+a}$
Коэффициент асимметрии	А=*0*
Коэффициент эксцесса	Е=*1,8*
Мода	*не определена*
Медиана	$M_e=M(x)$

Замечание.

1. Равномерное распределение является частным случаем бета-распределения.

2. Сумма *n* независимых равномерно распределенных случайных величин описывается нормальным распределением уже при $n\ge 5$.

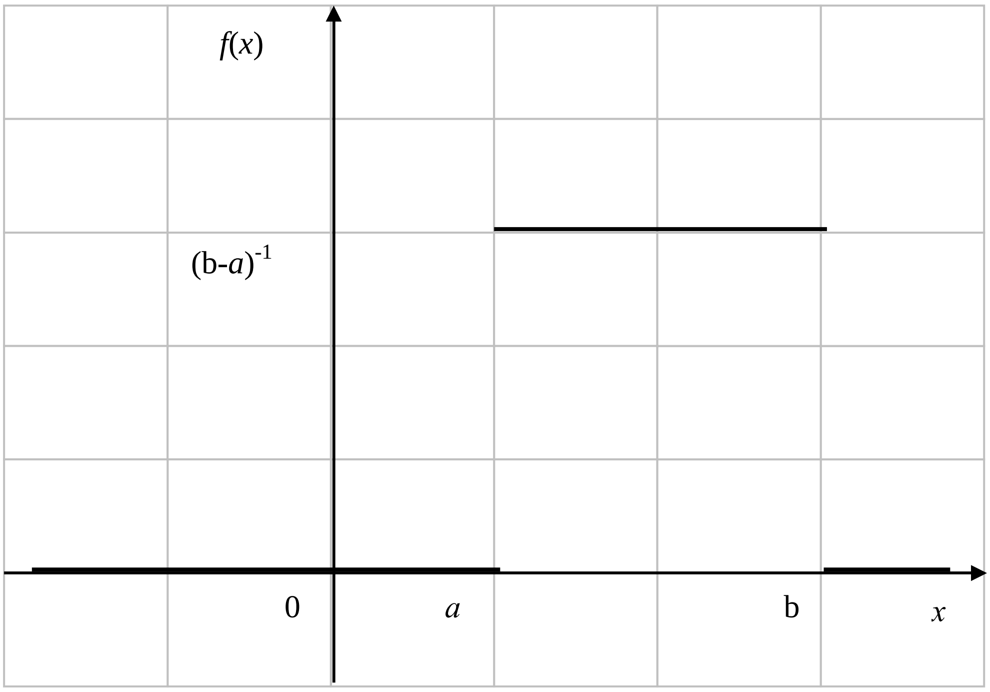


Рисунок 2 – График плотности равномерного распределения

Задача 4. Погрешность измерения прибора распределена равномерно на интервале [5;10]. Найти вероятность того, что погрешность прибора не превышает 7 ед. Вычислить погрешность измерения, вероятность которой равна 0,95. Вычислить вероятность того, что погрешность измерения будет находиться в интервале (6;8) ед.

Решение.

Так как имеем равномерно распределенную случайную величину, то вероятность того, что погрешность не превысит 7 ед., равна $F(7) = \frac{7-5}{10-5} = 0{,}4$.

Для вычисления значения погрешности измерения воспользуемся условием: $F(x) = \frac{x-a}{b-a} = 1 - 0{,}95 = 0{,}05$, *получаем* $x_{0,95} = (10-5) \cdot 0{,}05 + 5 = 5{,}25$.

Вероятность того, что погрешность измерения будет находится в интервале (6;8), находим из условия:

$$P(6 \le x \le 8) = F(8) - F(6) = \frac{8-5}{10-5} - \frac{6-5}{10-5} = 0{,}4.$$

2.1.3. Логарифмически нормальное распределение

Логарифмически нормальное распределение позволяет описать случайные величины, логарифм которых распределен по нормальному закону.

Логарифмически нормальное распределение – модель для процесса, появляющегося в результате большого числа небольших мультипликативных ошибок. Применимо, когда наблюдаемое значение случайной величины составляет случайную долю ранее наблюдаемого значения.

Примеры: распределение размеров кусков породы при ее дроблении, распределение времени безотказной работы транзисторов некоторых типов.

Свойства

Параметры	$a,\ b$
Плотность распределения	$f(x,a,b)=\dfrac{1}{xb\sqrt{2\pi}}e^{\left(-\frac{(\ln(x)-a)^2}{2b^2}\right)}$, $x>0$
Функция распределения	$F(x,a,b)=\dfrac{1}{\sqrt{2\pi}b}\int\limits_0^x e^{\left(-\frac{(\ln(t)-a)^2}{2b^2}\right)}dt$
Среднее выборочное	$M(x)=e^{a+\frac{1}{2}b^2}$
Дисперсия	$D(x)=e^{2a+b^2}\cdot\left(e^{b^2}-1\right)$
Стандартное отклонение	$\sigma=\sqrt{D(x)}$
Коэффициент вариации	$v=\left(e^{b^2}-1\right)^{1/2}$
Коэффициент асимметрии	$A=\left(e^{b^2}-1\right)^{1/2}\left(e^{b^2}+2\right)$
Коэффициент эксцесса	$E=3+\left(e^{b^2}-1\right)\left(e^{3b^2}+3e^{2b^2}+6\right)$
Мода	$M_0=e^{a-b^2}$
Медиана	$M_e=e^a$

Замечание. Логарифмически нормальное распределение иногда ошибочно принимается за экспоненциальное.

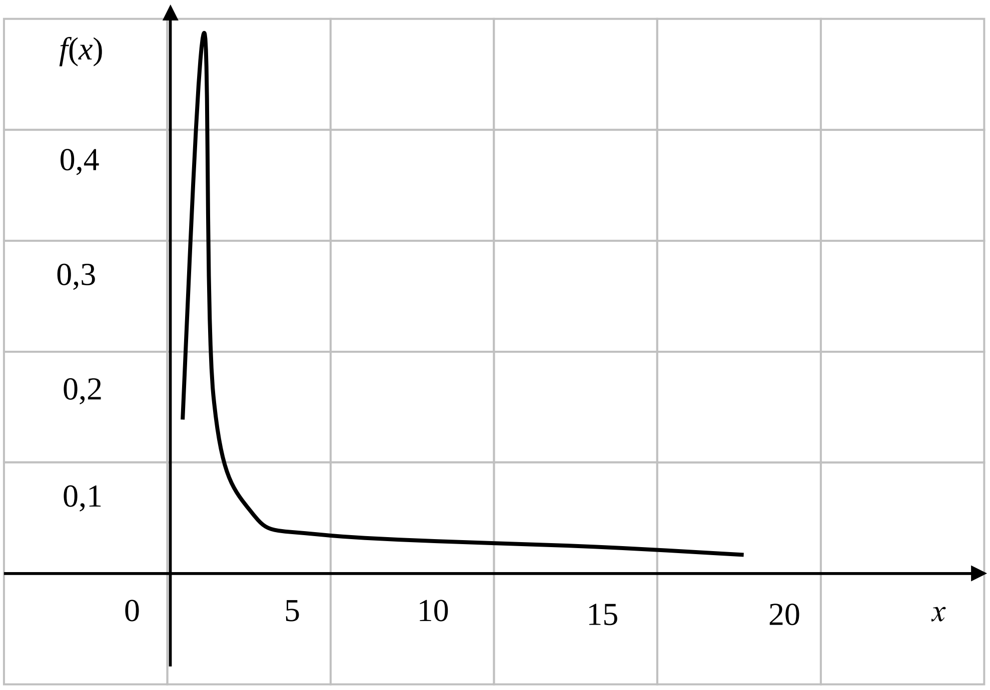


Рисунок 3 – График плотности логнормального распределения ($b^2=1,\ a=0$)

Задача 5. Предположим, что время безотказной работы элемента прибора – случайная величина, подчиняющаяся логарифмически нормальному закону распределения вероятностей с медианой, равной 1000 ч., и модой, равной 400 ч. Вычислить вероятность того, что элемент будет работать меньше 2000 ч.

Решение.

Имеем $M_e = e^a = 1000 \Rightarrow a = \ln(1000) = 6{,}908$,

$M_o = e^{a-b^2} = 400 \Rightarrow a - b^2 = \ln(400) = 5{,}991 \Rightarrow b^2 = 6{,}908 - 5{,}991 = 0{,}917 \Rightarrow b = 0{,}958$

Воспользуемся нормированной случайной величиной $z = \dfrac{\ln(x) - a}{b}$, *где* *x=2000. Получаем,* $z = \dfrac{\ln(2000) - 6{,}908}{0{,}958} = 0{,}723$.

Применим аппроксимацию 3:

$$P(x < 2000) = P(z < 0{,}723) = F(z) = 1 - 0{,}852e^{-\left(\frac{0{,}723+1{,}5774}{2{,}0637}\right)^{2{,}34}} = 0{,}76553.$$

2.1.4. Гамма-распределение

Гамма-распределение является основным распределением математической статистики для случайных величин, ограниченных с одной стороны. Гамма-распределение описывает время, необходимое для появления ровно k независимых событий при условии появления событий с постоянной интенсивностью.

Замечание. Широко используется в теории надежности и в теории массового обслуживания.

Примеры: наработка между несмежными отказами, распределение времени между повторными калибровками прибора, требующего повторной калибровки после пользования им k раз; распределение времени между моментами пополнения запасов; распределение времени безотказной работы системы с резервными компонентами.

Свойства

Параметры	$\alpha,\ \beta$
Плотность распределения	$f(x,\alpha,\beta)=\dfrac{1}{\alpha!\,\beta^{\alpha+1}}x^{\alpha}e^{-\left(\frac{x}{\beta}\right)}$, $x>0,\alpha>1,\beta>0$
Функция распределения	$F(x,\alpha,\beta)=1-\sum\limits_{i=0}^{\alpha}e^{-\frac{x}{\beta}}\left(\dfrac{x}{\beta}\right)^{i}\dfrac{1}{i!}$, $x>0,\alpha>1,\beta>0$
Среднее выборочное	$M(x)=\beta(\alpha+1)$
Дисперсия	$D(x)=\beta^{2}(\alpha+1)$
Стандартное отклонение	$\sigma=\sqrt{D(x)}$
Коэффициент вариации	$v=(\alpha+1)^{-1/2}$
Коэффициент асимметрии	$A=2(\alpha+1)^{-1/2}$
Коэффициент эксцесса	$E=3+6(\alpha+1)^{-1}$
Мода	$M_0=\alpha\beta$
Медиана	*не определена*

Замечание.

1. Гамма-распределение с целочисленным значением параметра называется также распределением Эрланга. При α =0 гамма-распределение переходит в экспоненциальное.
2. В теории надежности интенсивность отказа убывает при α <0, постоянна при α =0 и возрастает при α >0.

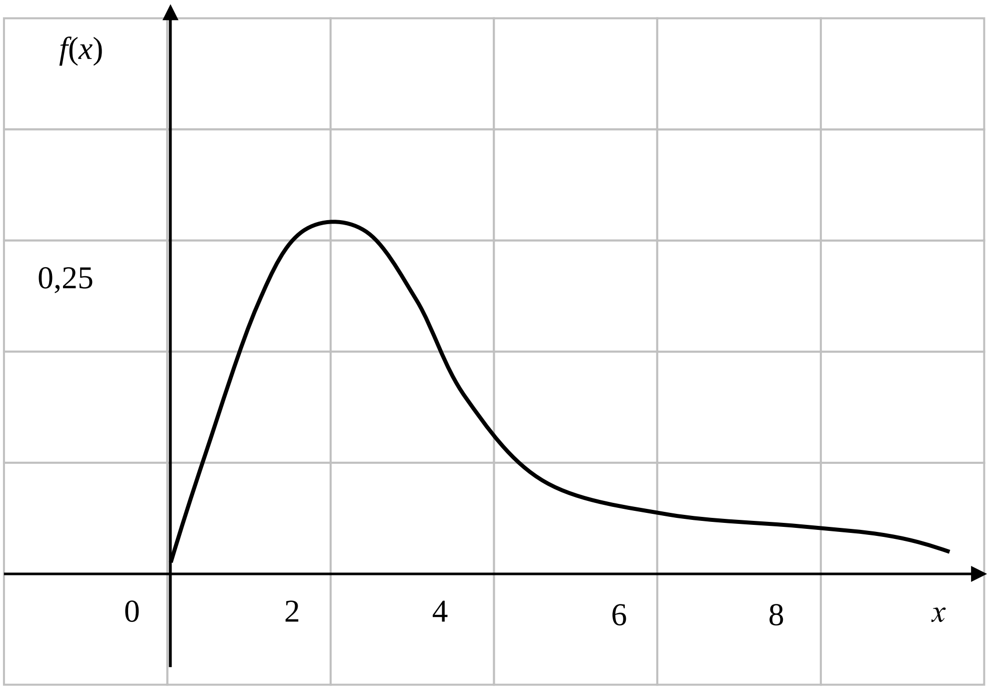


Рисунок 4 – График плотности гамма-распределения

Задача 6. Испытываются четыре прибора, интенсивность отказов которых известна и равна $\lambda = 10^{-5} ч^{-1}$. Вычислить вероятность того, что суммарная наработка приборов не превысит 300000 ч.

Решение.

Имеем гамма-распределение с параметрами $\alpha = 4 - 1 = 3$ *и* $\beta = \dfrac{1}{\lambda} = 10^5$.

Тогда вероятность равна

$$P(x < 300000) = F(300000) = 1 - e^{-3}\left(1 + 3 + 3^2 \cdot \frac{1}{2} + 3^3 \cdot \frac{1}{6}\right) = 0{,}352768.$$

2.1.5. Распределение Вейбулла

Распределение Вейбулла описывает общее распределение времени безотказной работы при самых разнообразных интенсивностях отказов.

Примеры: распределение времени безотказной работы для некоторых конденсаторов, шариковых подшипников, реле и т.д., распределение отказов элементов машин, одновременно имеющих признаки износовых и внезапных отказов.

Свойства

Параметры	$\alpha,\ \beta$
Плотность распределения	$f(x,\alpha,\beta)=\dfrac{\beta}{\alpha^{\beta}}x^{\beta-1}e^{-\left(\frac{x}{\alpha}\right)^{\beta}}$, $x\ge 0,\alpha,\beta>0$
Функция распределения	$F(x,\alpha,\beta)=1-e^{-\left(\frac{x}{\alpha}\right)^{\beta}}$, $x\ge 0,\alpha,\beta>0$
Среднее выборочное	$M(x)=\alpha\Gamma\left(1+\dfrac{1}{\beta}\right)$
Дисперсия	$D(x)=\alpha^{2}\left(\Gamma\left(1+\dfrac{2}{\beta}\right)-\left(\Gamma\left(1+\dfrac{1}{\beta}\right)\right)^{2}\right)$
Стандартное отклонение	$\sigma=\sqrt{D(x)}$
Коэффициент вариации	$v=\left(\dfrac{\Gamma\left(1+\frac{2}{\beta}\right)}{\left(\Gamma\left(1+\frac{1}{\beta}\right)\right)^{2}}-1\right)^{\frac{1}{2}}$
Коэффициент асимметрии	$A=\dfrac{\Gamma\left(1+\frac{3}{\beta}\right)-3\Gamma\left(1+\frac{2}{\beta}\right)\Gamma\left(1+\frac{1}{\beta}\right)+2\left(\Gamma\left(1+\frac{1}{\beta}\right)\right)^{3}}{\left(\Gamma\left(1+\frac{2}{\beta}\right)-\left(\Gamma\left(1+\frac{1}{\beta}\right)\right)^{2}\right)^{2}}$
Коэффициент эксцесса	$E=\dfrac{\Gamma\left(1+\frac{4}{\beta}\right)-4\Gamma\left(1+\frac{3}{\beta}\right)\Gamma\left(1+\frac{1}{\beta}\right)+6\Gamma\left(1+\frac{2}{\beta}\right)\left(\Gamma\left(1+\frac{1}{\beta}\right)\right)^{2}-3\left(\Gamma\left(1+\frac{1}{\beta}\right)\right)^{4}}{\left(\Gamma\left(1+\frac{2}{\beta}\right)-\left(\Gamma\left(1+\frac{1}{\beta}\right)\right)^{2}\right)^{2}}$

Мода	$M_0 = \alpha\left(1-\dfrac{1}{\beta}\right)^{\frac{1}{\beta}}$, $\beta > 1$
Медиана	$M_e = \alpha(\ln 2)^{\frac{1}{\beta}}$

Замечание. При $\beta = 1$ распределение Вейбулла переходит в экспоненциальное распределение с параметром α, при $\beta = 2$ - в распределение Рэлея. Вычисление моментов распределения Вейбулла производится по таблицам гамма-функции.

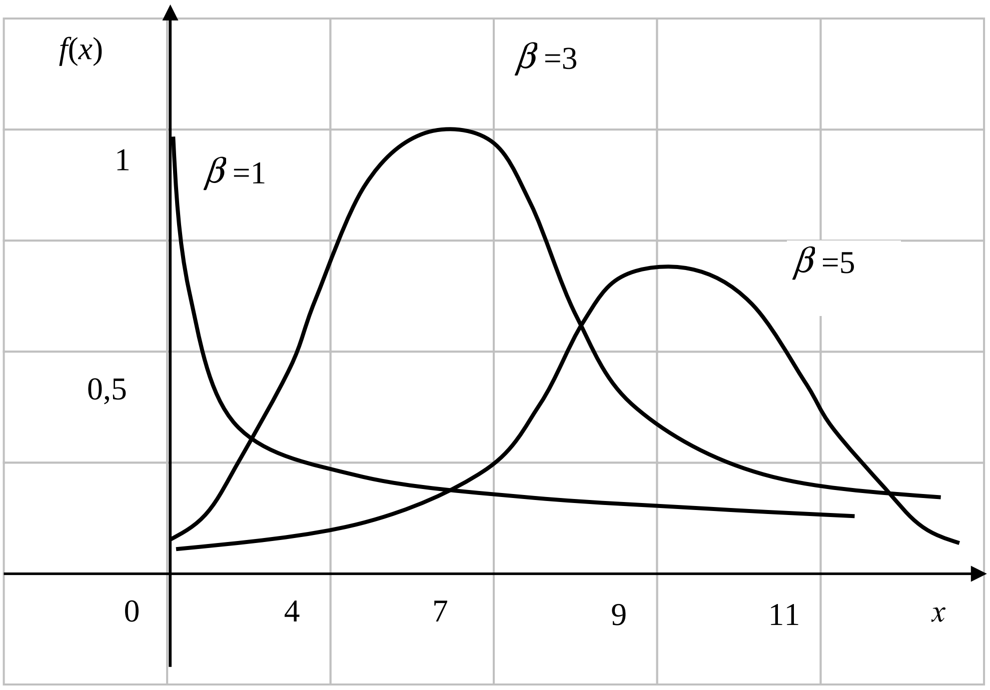


Рисунок 5 – График плотности распределения Вейбулла

Задача 7. Наработка прибора подчиняется распределению Вейбулла с параметрами $\alpha = 2$ и $\beta = 3$. Вычислить моду распределения и вероятность нахождения наработки прибора в интервале (5;6).

Решение.

Найдем моду распределения $M_0 = \alpha\left(1-\dfrac{1}{\beta}\right)^{\frac{1}{\beta}} = 2\left(1-\dfrac{1}{3}\right)^{\frac{1}{3}} = 1{,}74716$. *Тогда*

$$P(5 \le x \le 6) = F(6) - F(5) = 1 - e^{-\left(\frac{6}{2}\right)^3} - 1 + e^{-\left(\frac{5}{2}\right)^3} = e^{-\left(\frac{5}{2}\right)^3} - e^{-\left(\frac{6}{2}\right)^3} = 1{,}637 \cdot 10^{-7}.$$

2.1.6. Распределение Рэлея

Распределение Рэлея используется для описания радикальной ошибки на плоскости, когда ошибки по каждой оси координат независимы и распределены по нормальному закону с одинаковой дисперсией и нулевым математическим ожиданием.

Свойства

Параметры	α
Плотность распределения	$f(x,\alpha)=\frac{x}{\alpha^2}e^{\left(-\frac{x^2}{2\alpha^2}\right)}$, $x\ge 0, \alpha>0$
Функция распределения	$F(x,\alpha)=1-e^{\left(-\frac{x^2}{2\alpha^2}\right)}$, $x\ge 0, \alpha>0$
Среднее выборочное	$M(x)=\sqrt{\frac{\pi}{2}}\alpha\approx 1{,}253\alpha$
Дисперсия	$D(x)=\left(2-\frac{\pi}{2}\right)\alpha=0{,}429\alpha^2$
Стандартное отклонение	$\sigma=\sqrt{D(x)}$
Коэффициент вариации	$v=\left(\frac{4}{4-\pi}\right)^{\frac{1}{2}}=1{,}913$
Коэффициент асимметрии	$A=\frac{2\sqrt{\pi}(\pi-3)}{\sqrt{(4-\pi)^3}}=0{,}631$
Коэффициент эксцесса	$E=\frac{24\pi-16\pi^2-16}{(4-\pi)^2}=0{,}245$
Мода	$M_0=\alpha$
Медиана	$M_e=\alpha(2\ln 2)^{\frac{1}{2}}=1{,}177\alpha$

Замечание. Распределение Рэлея является частным случаем распределения Вейбулла при β=2.

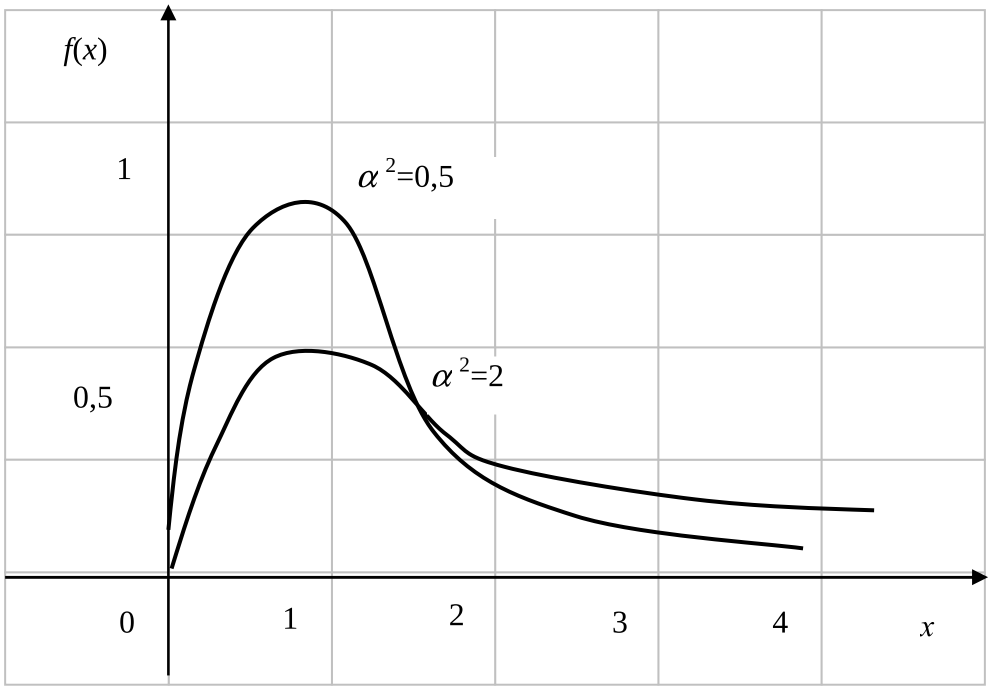


Рисунок 6 – График плотности распределения Рэлея

Задача 8. Случайная величина распределена по закону Рэлея с параметром $\alpha = 4$. Вычислить вероятность того, что случайная величина не превысит значения x=3, и вычислить 95%-ю квантиль распределения.
Решение.

Имеем $P(x<3) = 1 - e^{-\frac{3^2}{2\cdot 4^2}} = 0{,}24516$.

Для вычисления 95%-й квантили ($x_{0,95}$) используем равенство

$0{,}95 = 1 - e^{-\frac{x_{0,95}^2}{2\cdot 16}} \Rightarrow \ln(0{,}05) = -\frac{x_{0,95}^2}{32} \Rightarrow x_{0,95}^2 = 95{,}86$, *то есть с вероятностью 0,95 случайная величина не превысит значение 9,971.*

2.1.7. Экспоненциальное распределение

Экспоненциальное распределение описывает распределение времени между независимыми событиями, появляющимися с постоянной интенсивностью; распределение времени безотказной работы при постоянной интенсивности отказов.

Замечание. Это наиболее часто встречающееся распределение в теории надежности и в теории массового обслуживания.

Примеры: распределение времени между моментами поступления частиц в счетчик; распределение электронных элементов систем управления станками; распределение тонких стержневых инструментов, выходящие из строя в результате поломок; распределение внезапных отказов, когда износом можно пренебречь; распределение времени безотказной работы сложных нерезервированных систем и времени использования некоторых компонент, в частности, когда они подвергаются начальной приработке, а профилактическое обслуживание позволяет заменить детали до полного износа.

Свойства

Параметры	a
Плотность распределения	$f(x,a)=\frac{1}{a}e^{\left(-\frac{x}{a}\right)}$, $x\ge 0$
Функция распределения	$F(x,a)=1-e^{\left(-\frac{x}{a}\right)}$, $x\ge 0$
Среднее выборочное	$M(x)=a$
Дисперсия	$D(x)=a^2$
Стандартное отклонение	$\sigma=\sqrt{D(x)}$
Коэффициент вариации	$v=1$
Коэффициент асимметрии	$A=2$
Коэффициент эксцесса	$E=9$
Мода	$M_0=0$
Медиана	$M_e=a\ln(2)=0{,}6931$

Замечание.

1. Экспоненциальное распределение является частным случаем гамма-распределения и распределения Вейбулла.

2. Интенсивность отказа вычисляется по формуле: $\lambda=\frac{1}{a}$.

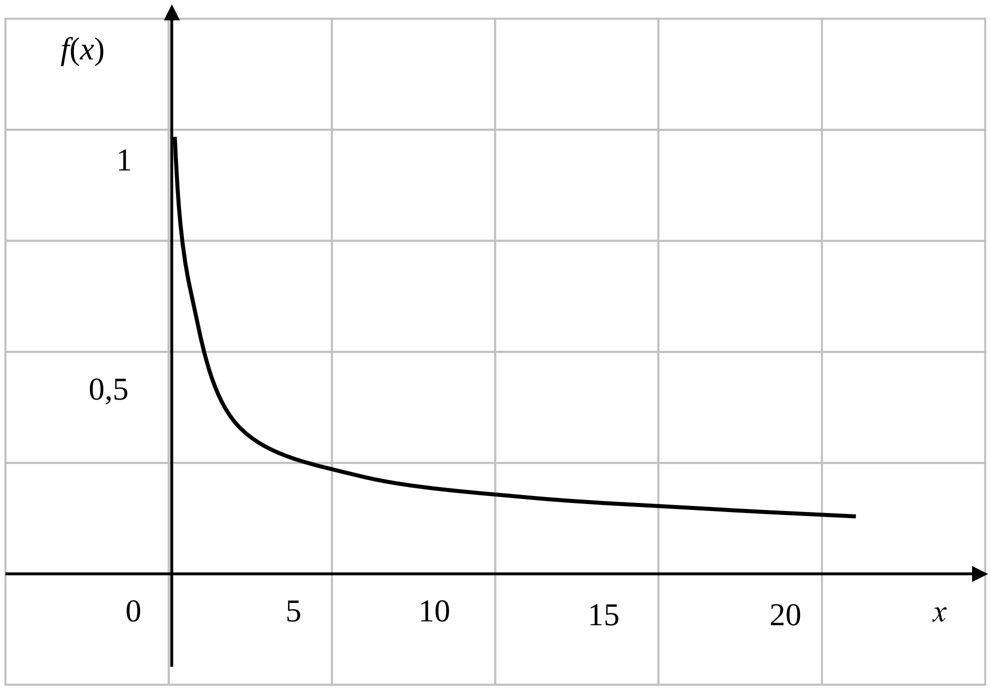


Рисунок 7 – График плотности экспоненциального распределения ($a=1$)

Задача 9. Наработка на отказ прибора распределена экспоненциально с интенсивностью отказов $\lambda = 10^{-5}\ ч^{-1}$. Вычислить вероятность того, что наработка на отказ превысит 1000 ч. Найти вероятность того, что наработка на отказ будет находиться в интервале от 1200 до 1500 ч. Вычислить значение наработки, вероятность превышения которой 0,8. Определить, как изменится наработка прибора при уменьшении интенсивности отказов до $\lambda = 0{,}5 \cdot 10^{-5}\ ч^{-1}$.

Решение.

1) Вероятность того, что наработка на отказ превысит 1000 ч., равна:

$$P(x > 1000) = 1 - F(1000) = 1 - \left(1 - e^{-1000 \cdot 10^{-5}}\right) = e^{-10^{-2}} = 0{,}99.$$

2) Вероятность того, что наработка будет находиться в интервале от 1200 ч. до 1500 ч., определим по формуле:

$$P(1200 < x < 1500) = F(1500) - F(1200) = 1 - e^{-1500 \cdot 10^{-5}} - 1 + e^{-1200 \cdot 10^{-5}} = e^{-0{,}012} - e^{-0{,}015} =$$
$$= 0{,}988071712 - 0{,}985111939 = 0{,}0029598.$$

3) Наработку y, вероятность превышения которой равна 0,8, находим из соотношения $P(x > y) = 1 - F(y) = 0{,}8$. *Отсюда имеем* $1 - 1 + e^{-y \cdot 10^{-5}} = 0{,}8 \Rightarrow$

$$\Rightarrow y \cdot 10^{-5} = 0{,}22314355 \Rightarrow y = 2{,}23 \cdot 10^{4}\ ч.$$

4) При снижении интенсивности отказов до $\lambda = 0{,}5 \cdot 10^{-5}\ ч^{-1}$ *получим* $y \cdot 0{,}5 \cdot 10^{-5} = 0{,}22314355 \Rightarrow y = 4{,}46 \cdot 10^{4}\ ч.$

2.1.8. Бета-распределение

Бета-распределение является основным распределением математической статистики для случайных величин, ограниченных с обеих сторон.

Замечание. Через бета-распределение могут быть выражены практически все применяемые распределения вероятностей.

Примеры: распределение доли совокупности, заключенной между наименьшим и наибольшим значениями выборки; распределение суточного производства на предприятии; распределение времени, ставшегося до завершения работы.

Свойства

Параметры	$\alpha,\ \beta$
Плотность распределения	$f(x,\alpha,\beta)=\dfrac{(\alpha+\beta+1)!}{\alpha!\beta!}x^{\alpha}(1-x)^{\beta}$, $0<x<1, \alpha,\beta\ge -1$
Функция распределения	$F(x,\alpha,\beta)=\dfrac{(\alpha+\beta+1)!}{\alpha!\beta!}\int\limits_0^x x^{\alpha}(1-x)^{\beta}dx$, $0<x<1, \alpha,\beta\ge -1$
Среднее выборочное	$M(x)=\dfrac{\alpha+1}{\alpha+\beta+2}$
Дисперсия	$D(x)=\dfrac{(\alpha+1)(\beta+1)}{(\alpha+\beta+2)^2(\alpha+\beta+3)}$
Стандартное отклонение	$\sigma=\sqrt{D(x)}$
Коэффициент вариации	$v=\left(\dfrac{\beta+1}{(\alpha+1)(\alpha+\beta+3)}\right)^{\frac{1}{2}}$
Коэффициент асимметрии	$A=\dfrac{2(\beta-\alpha)}{\alpha+\beta+4}\left(\dfrac{\alpha+\beta+3}{(\alpha+1)(\beta+1)}\right)^{\frac{1}{2}}$
Коэффициент эксцесса	$E=\dfrac{3(\alpha+\beta+2)(\alpha+\beta+3)}{(\alpha+\beta+4)(\alpha+1)(\beta+1)}\left(\dfrac{(\alpha+2)(-\alpha+2\beta+1)}{\alpha+\beta+5}+\dfrac{(\alpha+1)(\alpha-\beta)}{\alpha+\beta+2}\right)$
Мода	$M_0=\dfrac{\alpha}{\alpha+\beta}$
Медиана	*не определена*

Замечание. Частным случаем бета-распределения является равномерное распределение.

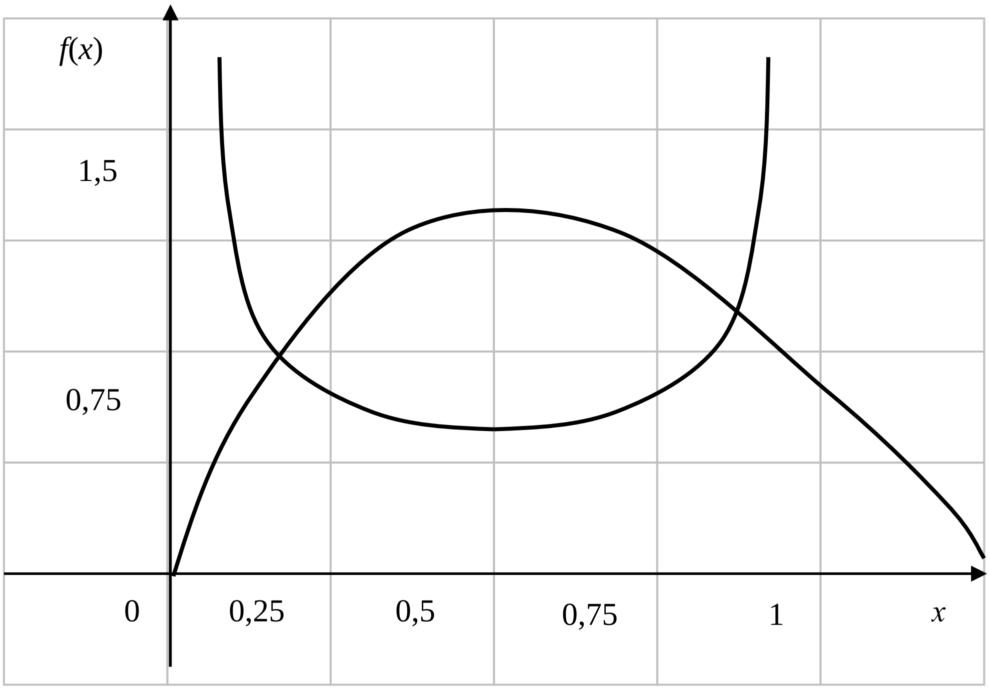


Рисунок 8 – График плотности бета-распределения

2.1.9. Распределение χ^2 (распределение Пирсона)

С помощью распределения χ^2 проверяются гипотезы относительно значений дисперсий, проверяется согласие экспериментальных данных с теоретическими законами распределения. Распределение χ^2 широко применяется в непараметрической статистике, являясь предельным для многих выборочных статистик.

Свойства

Параметры	f
Плотность распределения	$f(\chi^2, f) = \frac{1}{\sqrt{2}\Gamma\left(\frac{f}{2}\right)}\left(\chi^2\right)^{\frac{f-2}{2}} e^{-\frac{\chi^2}{2}}, \ \chi^2 \geq 0$
Функция распределения	$F_f(x) = P\left(\chi^2(f), x\right) = \frac{1}{\sqrt{2}\Gamma\left(\frac{f}{2}\right)}\int\limits_0^x y^{\frac{f}{2}-1} e^{-\frac{y}{2}} dy, \ x > 0$

Среднее выборочное	$M(\chi^2(f)) = f$
Дисперсия	$D(\chi^2(f)) = 2f$
Стандартное отклонение	$\sigma = \sqrt{D(\chi^2(f))}$
Коэффициент вариации	$v = \left(\frac{2}{f}\right)^{\frac{1}{2}}$
Коэффициент асимметрии	$A = 2\left(\frac{2}{f}\right)^{\frac{1}{2}}$
Коэффициент эксцесса	$E = 3 + \frac{12}{f}$
Мода	$M_0 = f - 2,\ f \geq 2$
Медиана	$M_e \approx f - \frac{2}{3}$

Замечание. В распределение χ^2 параметр f – число степеней свободы, определенный количеством независимых случайных величин, сумма квадратов которых составляет χ^2 - распределение. Распределение χ^2 определяется из таблиц.

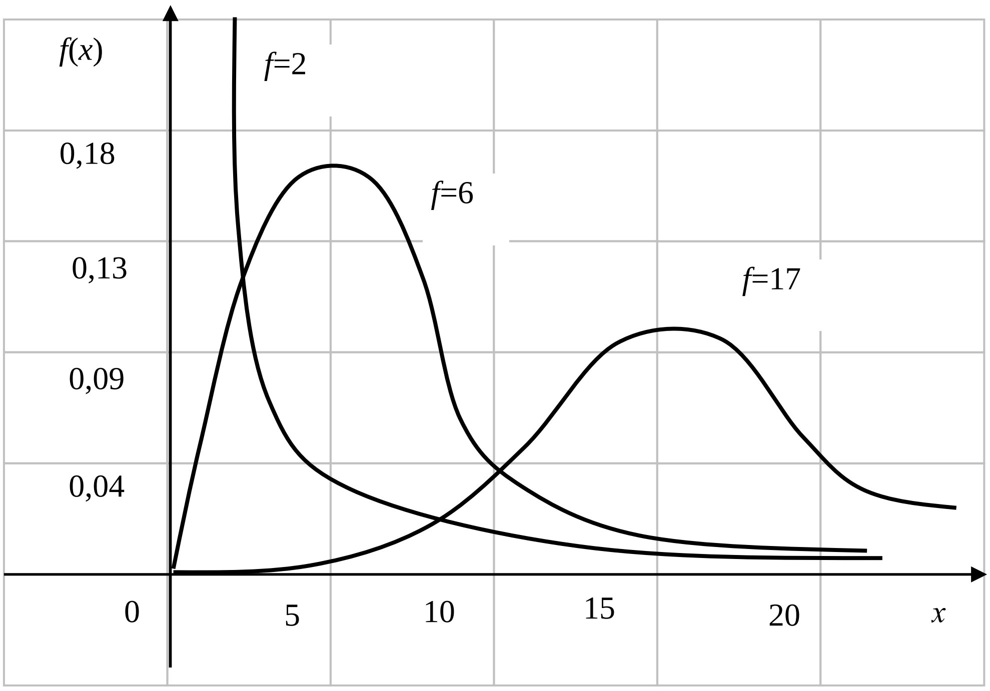


Рисунок 9 – График плотности χ^2 - распределения

Введем аппроксимации квантилей χ^2 - распределения:

- аппроксимация 7 . При f>200 $\chi_p^2(f) = f + \sqrt{2f} \cdot u_p$.

- аппроксимация 8 . При f>100 $\chi_p^2(f) = \frac{1}{2}(u_p + \sqrt{2f-1})^2$.

- аппроксимация 9 .

$$\chi_p^2(f) = f\left(1 - \frac{2}{9f} + \frac{4x^4+16x^2-28}{1215f^2} + \frac{8x^6+720x^4-3216x^2+2904}{229635f^3} + \right.$$

$$\left. + \left(\frac{2}{f}\right)^{\frac{1}{2}} \left(\frac{\frac{x}{3} + \left(-x^3+3x\right)}{162f} - \frac{3x^5+40x^3+45x}{5832f^2} + \frac{301x^7-1519x^5-3269x^3-79349x}{7873200f^3} \right)\right).$$

2.1.10. Распределение Стьюдента (t-распределение)

Распределение Стьюдента широко применяется в задачах обработки экспериментальных данных (например, при построении доверительных интервалов и проверке гипотез относительно среднего при неизвестной дисперсии). С помощью распределения Стьюдента описываются распределения коэффициентов корреляции и регрессии.

Свойства

Параметры	$t,\ f$
Плотность распределения	$f(t,f) = \dfrac{\Gamma\left(\frac{f+1}{2}\right)}{\sqrt{\pi f}\,\Gamma\left(\frac{f}{2}\right)}\left(1+\dfrac{t^2}{f}\right)^{-\frac{f+1}{2}}$
Функция распределения	$F(t;f) = P\big(t(f) < t\big) = \dfrac{1}{\sqrt{\pi f}}\int\limits_{-\infty}^{t}\left(1+\dfrac{y^2}{f}\right)^{-\frac{f+1}{2}} dy$
Среднее выборочное	$M(t) = 0$
Дисперсия	$D(t) = \dfrac{f}{f-2},\ f > 2$
Стандартное отклонение	$\sigma = \sqrt{D(t)}$

Коэффициент вариации	$v = 0$
Коэффициент асимметрии	$A = 0$
Коэффициент эксцесса	$E = \frac{3(f-2)}{f-4}, \ f > 4$
Мода	$M_0 = 0$
Медиана	$M_e = 0$

Замечание. При $f \to \infty$ t-распределение совпадает со стандартным нормальным (хорошая аппроксимация достигается уже при $f > 30$).

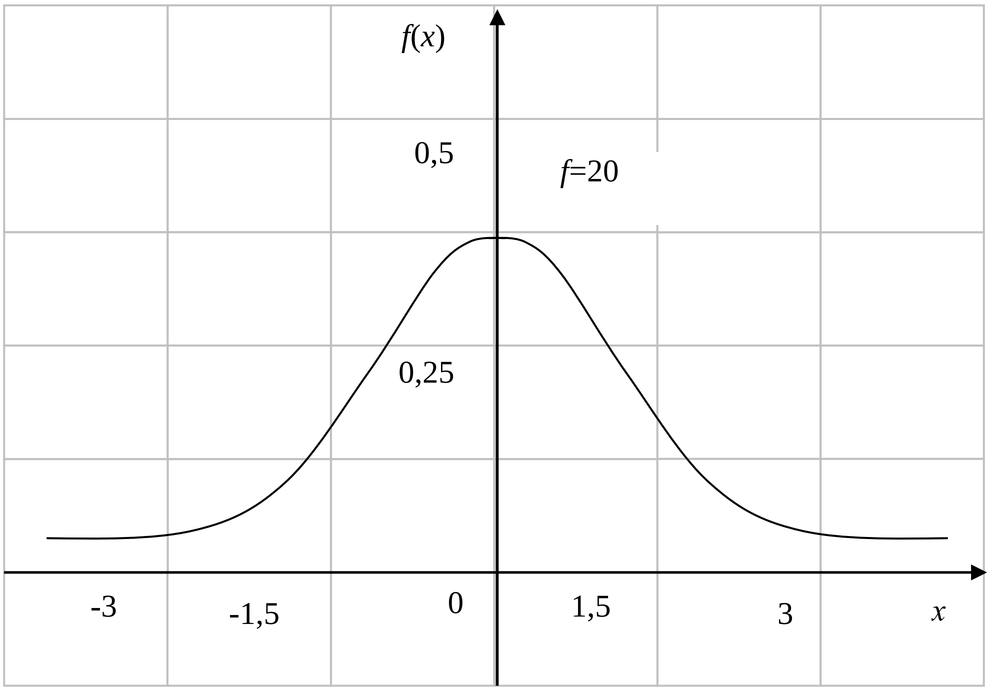


Рисунок 10 – График плотности распределения Стьюдента

2.1.11. Распределение Фишера (F-распределение)

F-распределение широко применяется при обработке данных (при сравнении дисперсий, анализе корреляций). С помощью F-распределения можно вычислить некоторые дискретные распределения, например, биномиальное.

Свойства

Параметры	$t,\ f_1,\ f_2$
Плотность распределения	$f(t; f_1, f_2) = \dfrac{\Gamma\left(\dfrac{f1+f_2}{2}\right)}{\Gamma\left(\dfrac{f_1}{2}\right)\Gamma\left(\dfrac{f_2}{2}\right)}\left(\dfrac{f_1}{f_2}\right)^{\frac{f_1}{2}} \dfrac{x^{\frac{f_1}{2}-1}}{\left(1+\dfrac{f_1}{f_2}x\right)^{\frac{f_1+f_2}{2}}},\ x \ge 0$
Функция распределения	$F(x) = P(F < x) = \dfrac{\Gamma\left(\dfrac{f_1+f_2}{2}\right)}{\Gamma\left(\dfrac{f_1}{2}\right)\Gamma\left(\dfrac{f_2}{2}\right)} f_1^{\frac{f_1}{2}} f_2^{\frac{f_2}{2}} \int\limits_0^x y^{\frac{f_1}{2}-1}(f_2+f_1)^{\frac{f_1+f_2}{2}}dy,$ $x \ge 0$
Среднее выборочное	$M(F(f_1, f_2)) = \dfrac{f_2}{f_2-2},\ f_2 > 2$
Дисперсия	$D(F(f_1, f_2)) = \dfrac{2f_2^2(f_1+f_2-2)}{f_1(f_2-2)^2(f_2-4)},\ f_2 > 4$
Стандартное отклонение	$\sigma = \sqrt{D(F(f_1, f_2))}$
Коэффициент вариации	$v = \left(\dfrac{2(f_1+f_2-2)}{f_1(f_2-4)}\right)^{\frac{1}{2}},\ f_2 > 4$
Коэффициент асимметрии	$A = \dfrac{2f_1+f_2-2}{f_2-6}\left(\dfrac{8(f_2-4)}{f_1+f_2-2}\right)^{\frac{1}{2}},\ f_2 > 6$
Коэффициент эксцесса	*не определен*
Мода	$M_0 = \dfrac{f_2(f_1-2)}{f_1(f_2+2)}$
Медиана	*не определена*

***Замечание*.**

1. Если $f_1 = 1$ *и* $f_2 = \infty$ или $f_1 = \infty$ *и* $f_2 = 1$, то F - распределение совпадает с нормальным распределением.

2. Если $f_2 \to \infty$, то F - распределение совпадает с χ^2 - распределением при f_1 степенях свободы.

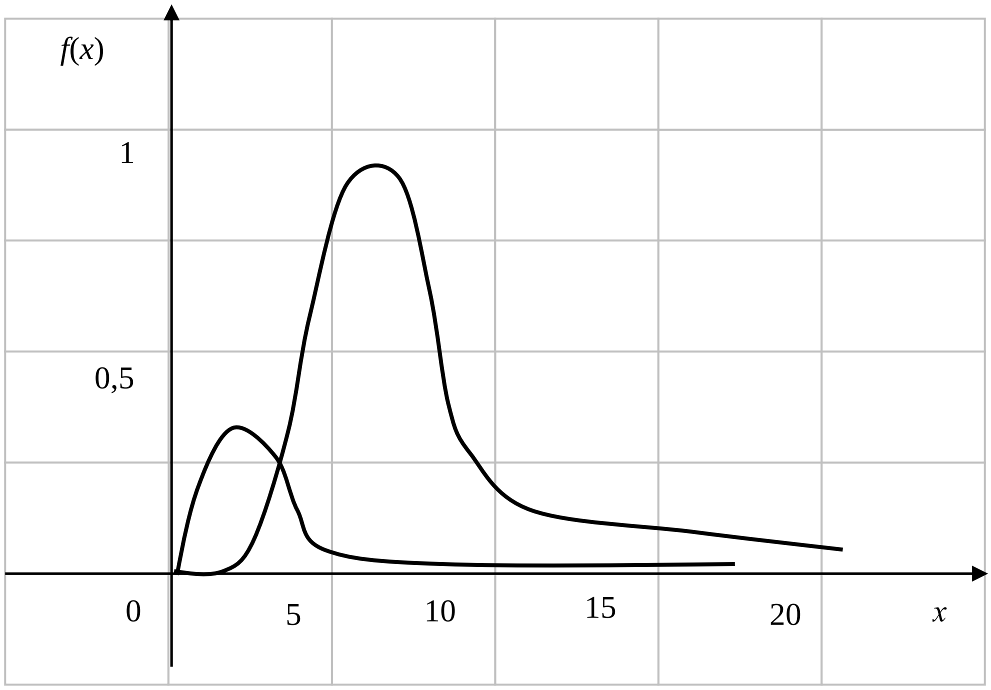


Рисунок 11 – График плотности распределения Фишера

2.1.12. Распределение Парето

Применяется в основном в описательной статистике, так, например, показывает, какие из небольшого количества проблем имеют наибольшее внимание. Иногда используется как простейшая математическая модель изменения интенсивности отказов приборов на этапе приработки.

Свойства

Параметры	c
Плотность распределения	$f(x,c) = cx^{-(c+1)}$, $1 \le x \le \infty$
Функция распределения	$F(x,c) = 1 - x^{-c}$, $1 \le x \le \infty$
Среднее выборочное	$M(x) = \frac{c}{c-1}$, $c > 1$
Дисперсия	$D(x) = \frac{c}{c-2} - \left(\frac{c}{c-1}\right)^2$, $c > 2$
Стандартное отклонение	$\sigma = \sqrt{D(x)}$

Коэффициент вариации	$v = \left(\frac{c-1}{c}\right)^{\frac{1}{2}}, \; c > 2$
Коэффициент асимметрии	$A = \frac{2(c+1)}{c-3}\sqrt{\frac{c-2}{c}}, \; c > 3$
Коэффициент эксцесса	$E = \frac{6(c^3 + c^2 - 6c - 2)}{c(c-3)(c-4)}, \; c > 4$
Мода	*не определена*
Медиана	*не определена*

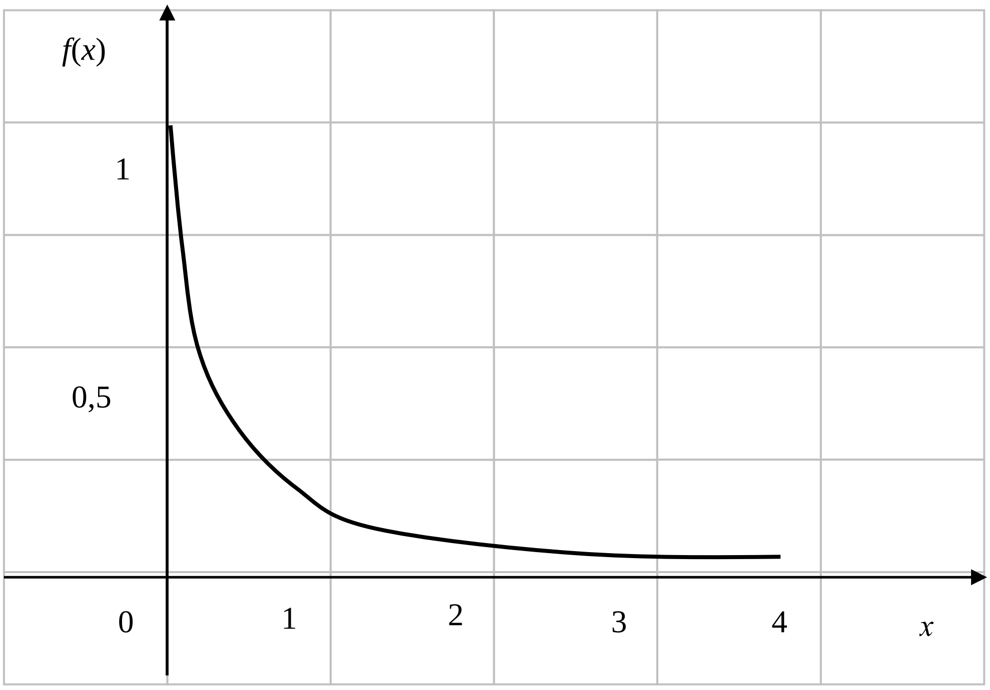


Рисунок 12 – График плотности распределения Парето

Пример. Исследуется проблема появления брака при выпуске деталей. С учётом того, что потери от брака одной детали каждого вида примерно одинаковы, в качестве единицы измерения выбирается число дефектных деталей каждого вида. После заполнения контрольных листков получаем данные, представленные в таблице:

№ детали	1	2	3	4	5	6	Прочие
Число дефектных деталей	255	101	59	39	26	15	11

По полученным данным составляется таблица для проверки данных.

номер детали	число дефек-тивных деталей	накопительная сумма деталей	процент деталей	накопительный процент
1	255	255	50,39526	50,39525692
2	101	356	19,96047	70,35573123
3	59	415	11,66008	82,01581028
4	39	454	7,70751	89,72332016
5	26	480	5,13834	94,86166008
6	15	495	2,964427	97,82608696
прочие	11	506	2,173913	100
итого	506			

Результаты расчетов для наглядности отображают через диаграмму (рисунок 13).

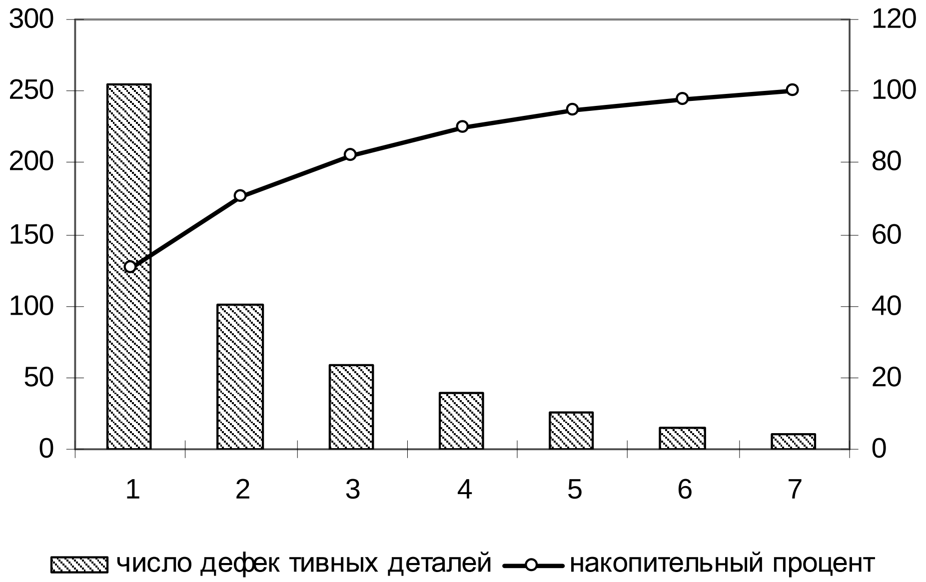


Рисунок 13 – Диаграмма Парето по числу дефектных деталей

Из диаграммы, получат вывод: к группе А можно отнести детали 1 и 2 (70% от брака), к группе В – детали 3,4,5, к группе С – детали 6 и прочие. Для выяснения наиболее важных дефектов целесообразно построить диаграммы Парето по явления дефектности в деталях 1 и 2.

Построим диаграмму для детали 1. В качестве единицы измерения выбирается сумму потерь от брака, млн. руб. После исследования явлений дефектности получили данные, представлены в таблице:

Дефект	Сумма потерь, млн. руб.
Наружный диаметр занижен	8,3
На режущей кромке резца налипы	6,9
Зависание	1,9
Шаг резьбы завышен	1,5
Осталась чернота	0,9
Скос кромки увеличен	0,6
Пропуск операции	0,4
Прочие	0,2

Построенная по этим данным диаграмма Парето имеет вид:

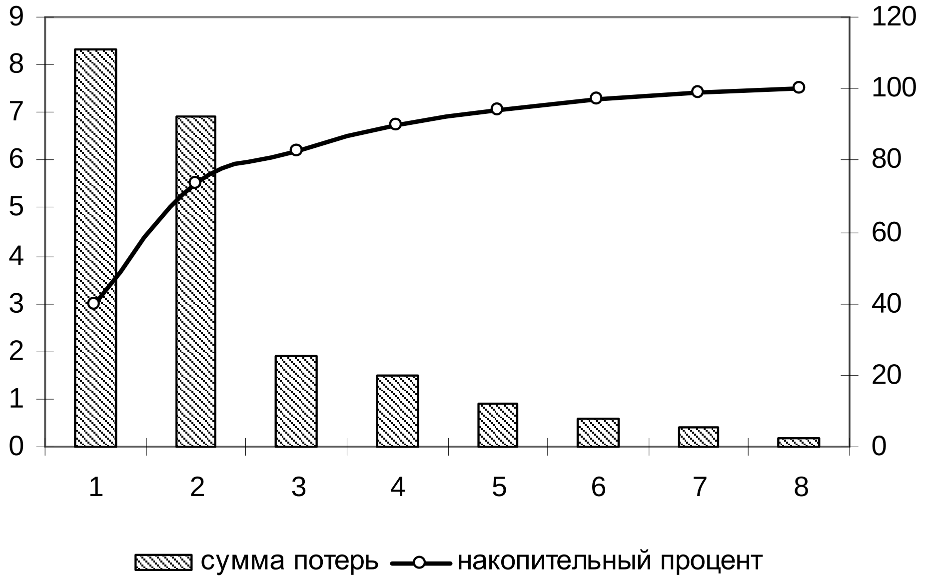


Рисунок 14 – Диаграмма Парето по дефектам детали 1

Из диаграммы получаем вывод, к группе А можно отнести занижение наружного диаметра и налипы на режущей кромке резца (73% от суммы потерь), к группе В – зависание, завышение шага резьбы, остаточную черноту, к группе С – увеличение скоса кромки, пропуск операции и прочие.
Для выяснения наиболее важных причин потерь целесообразно построить диаграммы Парето по причинам занижения наружного диаметра и налипов на режущей кромке резца.

2.2. Дискретные распределения

2.2.1. Биномиальное распределение (распределение Бернулли)

Биномиальному распределению подчиняются испытания невосстанавливаемых изделий на фиксированную наработку число отказов. Широко применяется при выборочном приемном контроле качества продукции и статистическом предупредительном контроле технологических процессов в производстве.

Свойства

Параметры	$n,\ p$
Плотность распределения	$f(x,n,p)=C_n^x p^x(1-p)^{n-x}$, $0\le p\le 1$
Функция распределения	$F(x,n,p)=\sum_{i=0}^{x} C_n^i p^i(1-p)^{n-i}$, $0\le p\le 1$
Среднее выборочное	$M(x)=np$
Дисперсия	$D(x)=np(1-p)$
Стандартное отклонение	$\sigma=\sqrt{D(x)}$
Коэффициент вариации	$v=\left(\frac{1-p}{np}\right)^{\frac{1}{2}}$
Коэффициент асимметрии	$A=(1-2p)(np(1-p))^{-\frac{1}{2}}$
Коэффициент эксцесса	$E=3-\frac{6}{n}+\frac{1}{np(1-p)}$
Мода	$M_0=(n+1)p$
Медиана	$M_e=np-1$, *или* $M_e=np$, *или* $M_e=np+1$

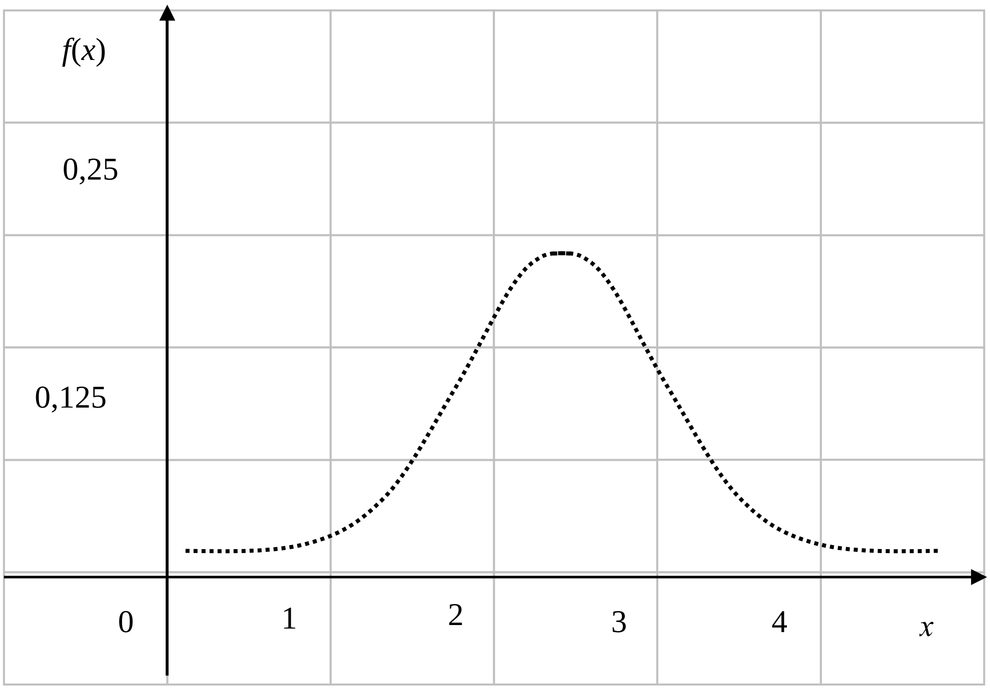


Рисунок 15 – График плотности биномиального распределения

Задача 10. Вероятность появления дефектного изделия в производстве равна 0,1. Вычислить вероятность появления в партии из 60 изделий не более 10 дефектных.

Решение.

Имеем $p = 0{,}1;\ n = 60; x = 10\ \Rightarrow np = 0{,}1 \cdot 60 = 6$. *Получаем*

$$P(x \le 10) = F(10;0{,}1;60) = \sum_{i=0}^{10} C_{60}^{i} \cdot 0{,}1^{i} \cdot (1-0{,}1)^{60-i} = C_{60}^{0} \cdot 0{,}9^{60} + C_{60}^{1} \cdot 0{,}1 \cdot 0{,}9^{59} + C_{60}^{2} \cdot 0{,}1^{2} \cdot 0{,}9^{58} +$$

$$+ C_{60}^{3} \cdot 0{,}1^{3} \cdot 0{,}9^{57} + C_{60}^{4} \cdot 0{,}1^{4} \cdot 0{,}9^{56} + C_{60}^{5} \cdot 0{,}1^{5} \cdot 0{,}9^{55} + C_{60}^{6} \cdot 0{,}1^{6} \cdot 0{,}9^{54} + C_{60}^{7} \cdot 0{,}1^{7} \cdot 0{,}9^{53} +$$

$$+ C_{60}^{8} \cdot 0{,}1^{8} \cdot 0{,}9^{52} + C_{60}^{9} \cdot 0{,}1^{9} \cdot 0{,}9^{51} + C_{60}^{10} \cdot 0{,}1^{10} \cdot 0{,}9^{50} = 0{,}9657.$$

2.2.2. Распределение Пуассона

Распределение Пуассона широко применяется в теории массового обслуживания. Распределению Пуассона подчинено число отказов невосстанавливаемых электронных приборов в течение периода приработки.

Свойства

Параметры	λ
Плотность распределения	$f(x,\lambda)=\dfrac{\lambda^x e^{-\lambda}}{x!}$, $\lambda>0$
Функция распределения	$F(x,\lambda)=\sum\limits_{i=0}^{x}\dfrac{\lambda^i e^{-\lambda}}{i!}$
Среднее выборочное	$M(x)=\lambda$
Дисперсия	$D(x)=\lambda$
Стандартное отклонение	$\sigma=\sqrt{D(x)}$
Коэффициент вариации	$v=\lambda^{-\frac{1}{2}}$
Коэффициент асимметрии	$A=\lambda^{-\frac{1}{2}}$
Коэффициент эксцесса	$E=3+\dfrac{1}{\lambda}$
Мода	$M_0=[\lambda]$
Медиана	*не определена*

Замечание. Распределение Пуассона является предельной формой биномиального распределения, то есть $\lim\limits_{\substack{n\to\infty \\ np\to\lambda}}\sum\limits_{i=0}^{x}C_n^i p^i(1-p)^{n-i}=\dfrac{\lambda^x}{x!}e^{-\lambda}$.

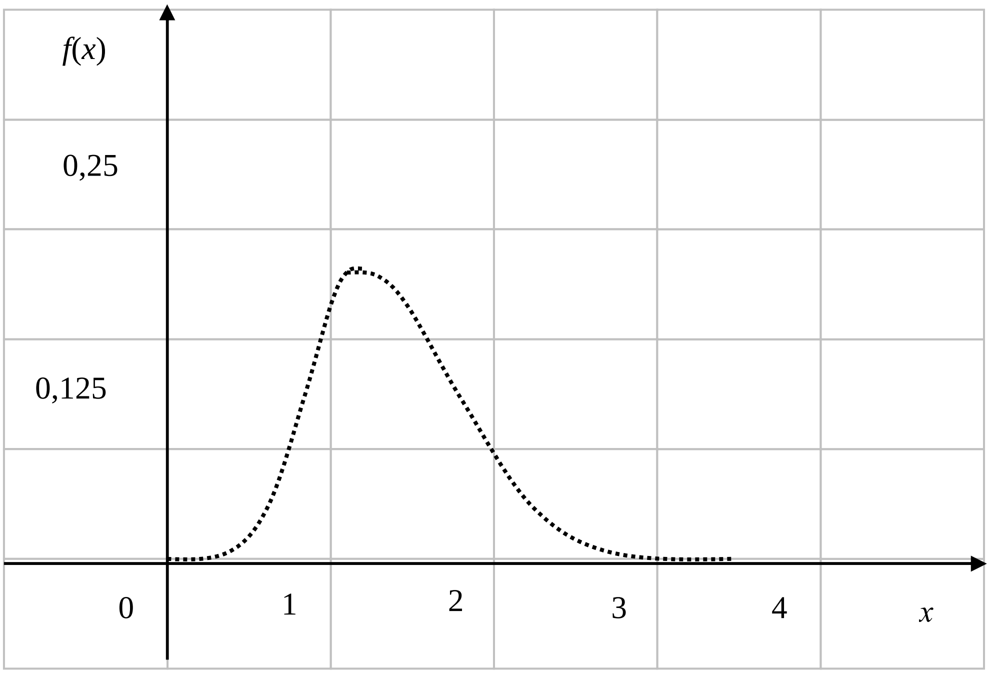


Рисунок 16 – График плотности распределения Пуассона

2.2.3. Отрицательное биномиальное распределение

Распределение используется при планировании запуска изделий в производство для получения требуемого количества годных изделий при известном проценте выхода годных, при планировании объема испытаний до получения заданного числа отказов.

Свойства

Параметры	$m,\ p$
Плотность распределения	$f(x,m,p)=C_{x+m-1}^{x}p^{m}(1-p)^{x}$, $0\le p\le 1$
Функция распределения	$F(x,m,p)=\sum_{i=1}^{m}C_{m+i-1}^{i}p^{m}(1-p)^{i}$, $0\le p\le 1$
Среднее выборочное	$M(x)=\dfrac{m(1-p)}{p}$
Дисперсия	$D(x)=\dfrac{m(1-p)}{p^{2}}$
Стандартное отклонение	$\sigma=\sqrt{D(x)}$
Коэффициент вариации	$v=\dfrac{1}{\sqrt{m(1-p)}}$
Коэффициент асимметрии	$A=(2-p)(m(1-p))^{-\frac{1}{2}}$
Коэффициент эксцесса	$E=3+\dfrac{6}{m}+\dfrac{p^{2}}{m(1-p)}$
Мода	$M_{0}=\dfrac{(m-1)(1-p)}{p}$, $m>1$
Медиана	*не определена*

Замечание. В отличие от биномиального, при отрицательном биномиальном распределении множество возможных значений случайной величины не ограничено сверху.

Задача 11. Вероятность получения дефектного изделия равна 0,1. Вычислить вероятность того, что будут произведены 50 годных изделий до появления 10-го дефектного изделия. Вычислить вероятность того, что до

появления 2-го дефектного изделия будут произведены не менее 5 годных изделий.
Решение.
Имеем $m=10,\ p=0{,}1,\ x=50$. *Получаем вероятность того, что потребуется произвести 50 изделий до появления 10-го дефектного изделия:*
$P(50;10;0{,}1)=C_{50+10-1}^{10}\cdot 0{,}1^{10}\cdot 0{,}9^{50}=0{,}03238$.
Вероятность того, что до появления 2-го дефектного изделия будет произведено не более 5 годовых изделий: $P(x<5)=\sum_{i=1}^{5}C_{2+i-1}^{i}\cdot 0{,}1^{2}\cdot 0{,}9^{i}=0{,}13969$.

2.2.4. Геометрическое распределение (распределение Фарри)

С помощью геометрического распределения можно определить объем выборки, необходимой для получения одного отказа при заданной вероятности отказа одного прибора.
Свойства

Параметры	p
Плотность распределения	$f(x,p)=p(1-p)^{x-1}$, $0\le p\le 1$
Функция распределения	$F(x,p)=\sum_{i=1}^{x}p(1-p)^{i-1}=1-(1-p)^{x}$, $0\le p\le 1$
Среднее выборочное	$M(x)=\frac{1}{p}$
Дисперсия	$D(x)=\frac{1-p}{p^2}$
Стандартное отклонение	$\sigma=\sqrt{D(x)}$
Коэффициент вариации	$v=\sqrt{1-p}$
Коэффициент асимметрии	$A=\frac{2-p}{\sqrt{1-p}}$
Коэффициент эксцесса	$E=9+\frac{p^2}{1-p}$
Мода	$M_0=0$

Медиана	*не определена*

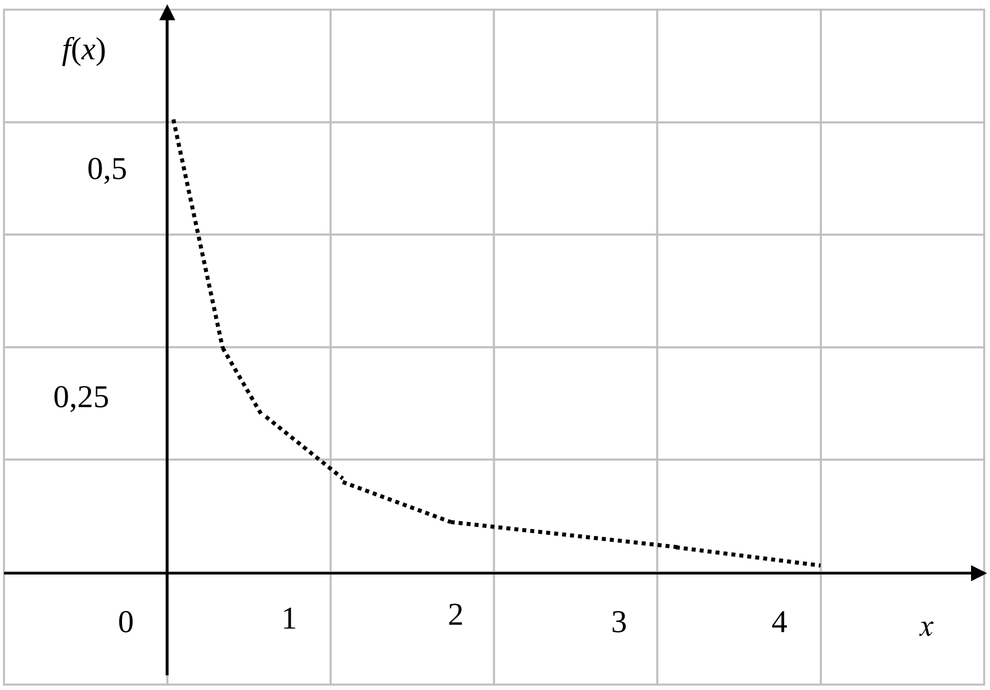


Рисунок 17 – График плотности геометрического распределения

Задача 12. Вероятность безотказной работы изделия равна 0,95. Вычислить вероятность того, что для получения одного отказа необходимо испытать выборку изделий из 10 приборов. Вычислить вероятность того, что для получения первого отказа понадобиться испытать не более 5 приборов.

Решение.

Вероятность того, что для получения одного отказа необходимо испытать 10 приборов, равна $P(10;0{,}05) = 0{,}05 \cdot 0{,}95^{10-1} = 0{,}031$.

Вероятность того, что для получения первого отказа понадобиться испытать не более 5 приборов, равна $P(x \le 5) = F(5;0{,}05) = 1 - (1 - 0{,}05)^5 = 0{,}2262$.

2.2.5. Гипергеометрическое распределение

Гипергеометрическое распределение широко применяется в задачах выборочного контроля качества продукции.

Свойства

Параметры	$N,\ n,\ D$
Плотность распределения	$f(x,N,n,D)=\dfrac{C_D^x C_{N-D}^{n-x}}{C_N^n}$
Функция распределения	$F(x,N,n,D)=\sum\limits_{i=0}^{x}\dfrac{C_D^i C_{N-D}^{n-i}}{C_N^n}$
Среднее выборочное	$M(x)=\dfrac{nD}{N}$
Дисперсия	$D(x)=\dfrac{nD}{N}\left(1-\dfrac{D}{N}\right)\left(\dfrac{N-n}{N-1}\right)$
Стандартное отклонение	$\sigma=\sqrt{D(x)}$
Коэффициент вариации	$v=\left(\dfrac{N-D}{nD}\cdot\dfrac{N-n}{N-1}\right)^{\frac{1}{2}}$
Коэффициент асимметрии	$A=(N-2D)(nD(N-D))^{-\frac{1}{2}}\left(\dfrac{N-1}{N-n}\right)^{\frac{1}{2}}\dfrac{N-2n}{N-2}$
Коэффициент эксцесса	$E=\dfrac{N^2(N-1)}{(N-2)(N-3)n(N-n)}\left(\dfrac{N(N+1)-6N(N-n)}{D(N-D)}+\dfrac{3n(N-n)(N+6)}{N^2}-6\right)$
Мода	$M_0=\dfrac{(D+1)(n+1)}{N+2}$
Медиана	*не определена*

Задача 13. Из партии, состоящей из 1000 изделий, 30 из которых дефектные, взята выборка объёмом 50 изделий. Построить график плотности распределения вероятностей, используя гипергеометрическое распределение.

Решение.

Введем обозначения: N=1000, D=50, n=30. Составим таблицу расчетных данных:

M	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*≤30*
f(m)	*0,21*	*0,34*	*0,26*	*0,13*	*0,044*	*0,011*	*0,002*	*0,0003*	*≈0*

Построим график плотности гипергеометрического распределения:

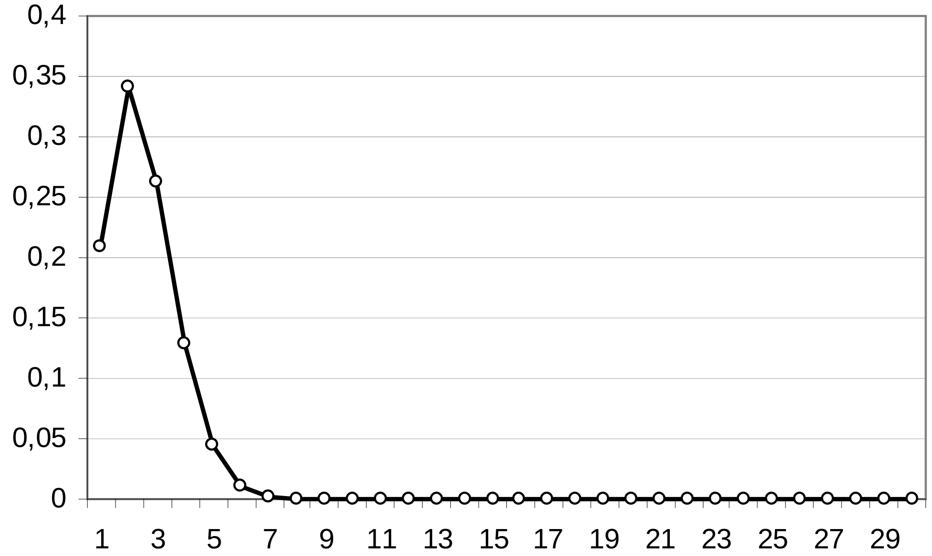


Рисунок 18 – График плотности гипергеометрического распределения

Глава 3. Методы анализа законов распределения вероятностей случайных величин

Зная закон распределения вероятностей наблюдаемой случайной величины, исследователь может решить многие практические задачи, связанные с планированием производства, обеспечением качества продукции, оценкой эффективности и стабильности производства.

3.1. Общие критерии согласия

3.1.1. Критерий согласия χ^2 (Пирсона)

Для применения критерия полученные данные группируются по интервалам частот и сравниваются с ожидаемым числом наблюдений для принятого распределения. На основе этого сравнения вычисляется критерий, который приближенно следует распределению χ^2 только в том случае, если модель выбрана правильно, если модель выбрана неправильно, то значения критерия превысит значения случайной величины, распределенной по закону χ^2.

Проверка с помощью критерия χ^2 проходит следующим образом:

Этап 1. Диапазон изменения экспериментальных данных разбивается на k интервалов, где количество интервалов вычисляется по формуле: $k \approx 1 + \log_2(n)$. Затем, подсчитывается число наблюдений в каждом интервале.

Этап 2. По полученным эмпирическим данным строится гистограмма (полигон) частот или относительных частот. На основании сравнения полигона относительных частот с графиками плотностей теоретических распределений выдвигается основная гипотеза H_0.

Этап 3. Подсчитывается статистика:

1 способ - $\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}$, где n_i - количество значений случайной величины, попавших в i-й интервал; $n = \sum_{i=1}^{k} n_i$ - объем выборки; $p_i = F(x_{i+1}) - F(x_i)$ - теоретическая вероятность попадания случайной величины в i-й интервал; $F(x)$ - гипотетический теоретический закон распределения вероятностей случайной величины.

2 способ - $\chi^2 = n\sum_{i=1}^{k}\frac{(w_i - p_i)^2}{p_i}$, где w_i - относительные частоты, заданные статистической таблицей, а p_i - вероятности, полученные по некоторому теоретическому закону распределения.

Этап 4. Рассматривается разность $r = k - l$, называемая числом степеней свободы, где k – число разрядов статистической таблицы, а l – число условий, налагаемых на частоты $w_1, w_2, ..., w_k$.

Замечание. Для нормального закона распределения l=3, так как используются следующие условия: $\sum_{i=1}^{k} w_i = 1$, $\sum_{i=1}^{k} w_i x_i = M(x)$, $\sum_{i=1}^{k}(x_i - M(x))^2 w_i = D(x)$.

По таблице приложения 6 определяется $\chi^2_{\alpha,r}$ (α - уровень значимости). Если $\chi^2 \le \chi^2_{\alpha,r}$, то гипотезу H_0 принимают; если $\chi^2 > \chi^2_{\alpha,r}$, то гипотезу H_0 отвергают.

Этап 5. Используя таблицу приложения 3, по значениям χ^2 и r определяют величину p, характеризующую вероятность согласованности теоретического и эмпирического распределения. Если $p < 0{,}1$, то можно сделать вывод, что гипотеза о принятом теоретическом распределении не противоречит опытным данным.

Замечание. В.И. Романовским предложен следующий критерий согласия: если величина $R = \frac{|\chi^2 - r|}{\sqrt{2r}} \ge 3$, то расхождение теоретических и эмпирических частот надо считать неслучайным; если же $R = \frac{|\chi^2 - r|}{\sqrt{2r}} < 3$, то это расхождение можно считать случайным.

Основным преимуществом критерия является его гибкость. Его можно легко использовать для проверки допущения о любом распределении, даже не зная значений параметров распределения. Основными его недостатками являются нечувствительность к обнаружению адекватной модели, когда число наблюдений невелико, и зачастую необходимость группирования данных по произвольным интервалам, что может оказать влияние на результат проверки.

Задача 14. Проведено исследование работы автоматической линии. Наработка между последовательными отказами автоматической линии приведена в таблице:

Интервал	(4,1; 4,2]	(4,2; 4,3]	(4,3; 4,4]	(4,4; 4,5]	(4,5; 4,6]	(4,6; 4,7]	(4,7; 4,8]	(4,8; 4,9]	(4,9; 5]
n_x	1	2	3	4	5	8	8	9	10
Интервал	(5; 5,1]	(5,1; 5,2]	(5,2; 5,3]	(5,3; 5,4]	(5,4; 5,5]	(5,5; 5,6]	(5,6; 5,7]	(5,7; 5,8]	(5,8; 5,9]
n_x	10	9	9	7	5	4	3	2	1

Определить модель распределения и подтвердить гипотезу критерием χ^2 и Романовского, при этом уровень значимости $\alpha = 0{,}1$.

Решение.

Этап 1. Имеем объем выборки $n = 100$. *Определим середины каждого интервала и вычислим относительные частоты* w_i*:*

k	*интервал*	n_x	x_i	w_i
1	(4,1;4,2]	1	4,15	0,01
2	(4,2;4,3]	2	4,25	0,02
3	(4,3;4,4]	3	4,35	0,03
4	(4,4;4,5]	4	4,45	0,04
5	(4,5;4,6]	5	4,55	0,05
6	(4,6;4,7]	8	4,65	0,08
7	(4,7;4,8]	8	4,75	0,08
8	(4,8;4,9]	9	4,85	0,09
9	(4,9;5]	10	4,95	0,1
10	(5;5,1]	10	5,05	0,1
11	(5,1;5,2]	9	5,15	0,09
12	(5,2;5,3]	9	5,25	0,09
13	(5,3;5,4]	7	5,35	0,07
14	(5,4;5,5]	5	5,45	0,05
15	(5,5;5,6]	4	5,55	0,04
16	(5,6;5,7]	3	5,65	0,03
17	(5,7;5,8]	2	5,75	0,02
18	(5,8;5,9]	1	5,85	0,01
сумма		*100*		*1*

Этап 2. Построим полигон частот.

Рисунок 19 – График полигона частот

Сравнивая полученный график с графиками теоретических распределений случайных величин выдвигаем гипотезу H_0: данное распределение является нормальным распределением.

Этап 3. Проверим верность выдвинутой гипотезы по критерию χ^2. Для это выполним предварительные вычисления:

- *математическое ожидание –* $a = M(x) = \sum_{i=1}^{k} x_i w_i = 4{,}999$;

- *дисперсия -* $D(x) = \sum_{i=1}^{k} x_i^2 w_i - M^2(x) = 0{,}140299$;

-*среднее квадратическое отклонение -* $\sigma(x) = \sqrt{D(x)} = 0{,}374565$.

Вычислим теоретическую плотность, пользуясь свойством нормального распределения, результат приведем в таблице:

k	*интервал*	n_x	x_i	w_i	$f(x_i)$	$hf(x_i)=p_i$	$(w_i - p_i)^2$	$(w_i - p_i)^2/p_i$
1	(4,1;4,2]	1	4,15	0,01	0,081614	0,008161	3,38037E-06	0,000414
2	(4,2;4,3]	2	4,25	0,02	0,144243	0,014424	3,10881E-05	0,002155
3	(4,3;4,4]	3	4,35	0,03	0,237394	0,023739	3,91946E-05	0,001651
4	(4,4;4,5]	4	4,45	0,04	0,363823	0,036382	1,30875E-05	0,00036
5	(4,5;4,6]	5	4,55	0,05	0,519225	0,051923	3,69601E-06	7,12E-05
6	(4,6;4,7]	8	4,65	0,08	0,690026	0,069003	0,000120942	0,001753
7	(4,7;4,8]	8	4,75	0,08	0,853927	0,085393	2,90817E-05	0,000341
8	(4,8;4,9]	9	4,85	0,09	0,984059	0,098406	7,06596E-05	0,000718
9	(4,9;5]	10	4,95	0,1	1,056007	0,105601	3,13674E-05	0,000297
10	(5;5,1]	10	5,05	0,1	1,055254	0,105525	3,05303E-05	0,000289
11	(5,1;5,2]	9	5,15	0,09	0,981957	0,098196	6,717E-05	0,000684
12	(5,2;5,3]	9	5,25	0,09	0,85089	0,085089	2,41183E-05	0,000283
13	(5,3;5,4]	7	5,35	0,07	0,686592	0,068659	1,79768E-06	2,62E-05

14	(5,4;5,5]	5	5,45	0,05	0,515905	0,05159	2,52966E-06	4,9E-05
15	(5,5;5,6]	4	5,55	0,04	0,360982	0,036098	1,52241E-05	0,000422
16	(5,6;5,7]	3	5,65	0,03	0,235205	0,02352	4,1984E-05	0,001785
17	(5,7;5,8]	2	5,75	0,02	0,142709	0,014271	3,28222E-05	0,0023
18	(5,8;5,9]	1	5,85	0,01	0,080631	0,008063	3,75148E-06	0,000465
Сумма		*100*		*1*		*0,984045*		*0,014064*

Получаем $\chi^2 = 100 \cdot 0{,}014064 \approx 1{,}4$.
Сравним графики полигона относительных частот и плотности нормального распределения:

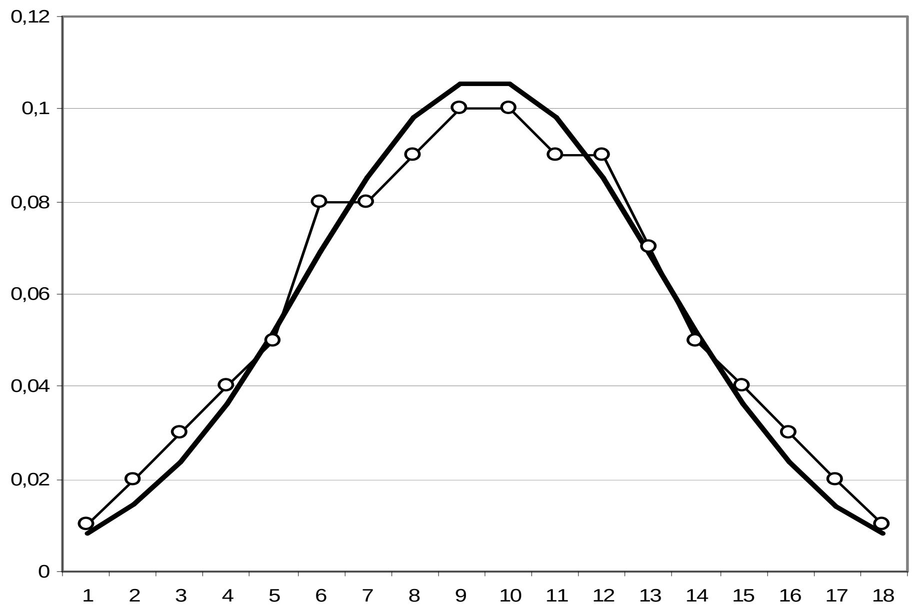


Рисунок 20 – Графики полигона относительных частот и плотности нормального распределения

Этап 4. Из таблицы приложения 8, при уровне значимости $\alpha = 0{,}1$ *и степени свободы* $r = 18 - 3 = 15$, *имеем* $\chi^2_{0,1;15} = 8{,}54$. *Таким образом, так как* $\chi^2 = 1{,}4 < \chi^2_{0,1;15} = 8{,}54$ *гипотеза* H_0 *принимается, то есть данное распределение является нормальным распределением.*
Этап 5. Используя таблицу приложения 5, по значениям χ^2 *и r определяем величину p: если* $\chi^2 \approx 1$, *то p=1; если* $\chi^2 \approx 2$, *то p=1. Получаем, что гипотеза о нормальном распределении не противоречит опытным данным. Воспользуемся критерием Романовского:* $R = \frac{|\chi^2 - r|}{\sqrt{2r}} = \frac{|1{,}4 - 15|}{\sqrt{30}} \approx 2{,}483 < 3$.
Следовательно, этот критерий также подтверждает, что распределение является нормальным распределением.

Задача 15. Полученные данные, анализа отношения длины к ширине зерна после шлифования, занесены в таблицу статистического распределения:

x_i	0,8	1,2	1,6	2	2,4	2,8	3,2	3,6	4	4,4	4,8	5,2
w_i	0,1	0,15	0,19	0,17	0,13	0,09	0,07	0,05	0,03	0,01	0,009	0,001

Объем выборки $n = 50$. Определить модель распределения и подтвердить гипотезу критерием χ^2.

Решение.

Этап 2. Построим полигон относительных частот.

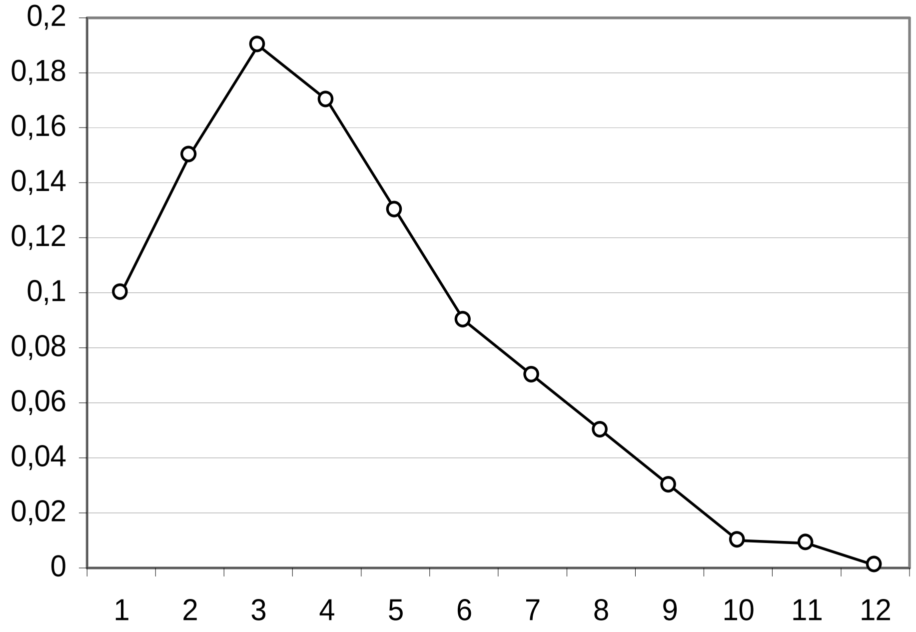


Рисунок 21 – График полигона относительных частот

Сравнивая полученный график с графиками теоретических распределений случайных величин выдвигаем гипотезу H_0:

1 предположение – данное распределение является логнормальным распределением

2 предположение – данное распределение является распределением Рэлея (частный случай распределения Вейбулла).

Рассмотрим 1 предположение.

Этап 3. Проверим верность выдвинутой гипотезы по критерию χ^2.

Для это выполним предварительные вычисления:

- *математическое ожидание* – $M(x) = \sum_{i=1}^{k} x_i w_i = 2{,}0844$;

- *мода* - $M_0 = 1{,}6$.

Тогда согласно свойствам имеем систему:

$$\begin{cases} \ln(2{,}08) = a + \dfrac{1}{2} b^2, \\ \ln(1{,}6) = a - b^2 \end{cases}$$

$$\Rightarrow\ b^2 = \frac{2}{3}\left(\ln(2{,}08) - \ln(1{,}6)\right) = 0{,}1749$$

$$и\ \ a = \ln(2{,}08) - \frac{1}{2} b^2 = 0{,}6449\,.$$

Вычислим теоретическую плотность, пользуясь свойством логнормального распределения, результат приведем в таблице:

k	x_i	w_i	$f(x_i)$	$hf(x_i)=p_i$	$w_i - p_i$	$(w_i - p_i)^2$	$(w_i - p_i)^2/p_i$
1	0,8	0,1	0,13836	0,05534	0,0446525	0,0019938	0,0360241
2	1,2	0,15	0,43128	0,17251	-0,0225146	0,00050691	0,0029383
3	1,6	0,19	0,54640	0,21856	-0,0285608	0,0008157	0,0037322
4	2	0,17	0,47390	0,18956	-0,0195636	0,0003827	0,0020190
5	2,4	0,13	0,34151	0,13660	-0,0066060	4,364E-05	0,0003194
6	2,8	0,09	0,22321	0,08928	0,0007152	5,115E-07	5,729E-06
7	3,2	0,07	0,13836	0,05534	0,0146525	0,0002146	0,0038790
8	3,6	0,05	0,083387	0,03335	0,0166448	0,000277	0,0083060
9	4	0,03	0,04956	0,01982	0,0101739	0,0001035	0,0052208
10	4,4	0,01	0,02931	0,01172	-0,0017251	2,976E-06	0,0002538
11	4,8	0,009	0,01734	0,00693	0,0020631	4,256E-06	0,0006135
12	5,2	0,001	0,01029	0,00411	-0,0031197	9,732E-06	0,0023624
Сумма		1		0,993188			0,065674805

Получаем $\chi^2 = 50 \cdot 0{,}065674805 \approx 3{,}2837$.

Сравним графики полигона относительных частот и плотности логнормального распределения:

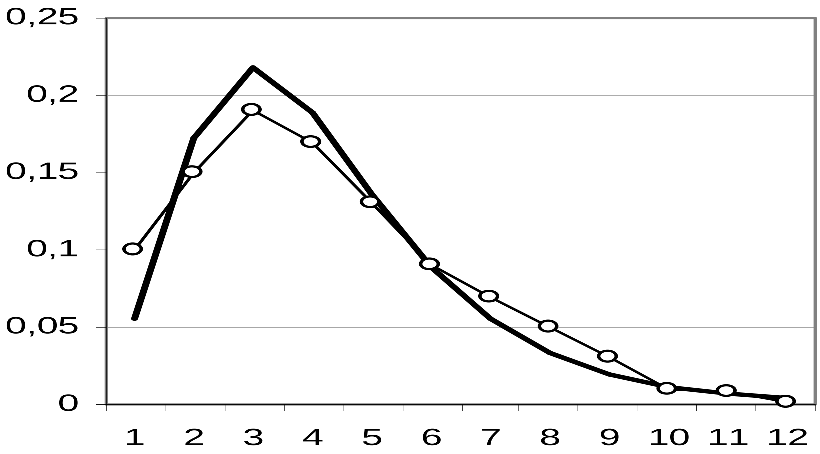


Рисунок 22 – Графики полигона относительных частот и плотности логнормального распределения

Этап 4. Из таблицы приложения 8, при уровне значимости $\alpha = 0{,}1$ *и числу степеней свободы* $r = 12 - 2 = 10$, *имеем* $\chi^2_{0,1;10} = 4{,}86$.

Таким образом, так как $\chi^2 = 3{,}2837 < \chi^2_{0,1;10} = 4{,}86$ *гипотеза H_0 принимается, то есть данное распределение является логнормальным распределением.*

Рассмотрим 2 предположение.

Этап 3. Проверим верность выдвинутой гипотезы по критерию χ^2.

Для это выполним предварительные вычисления:

математическое ожидание, с одной стороны, равно $M(x) = \sum_{i=1}^{k} x_i w_i = 2{,}0844$,

а, с другой стороны, согласно свойству распределения Рэлея, равно $M(x) = \sqrt{\frac{\pi}{2}}\alpha$. *Тогда получаем* $\alpha = \sqrt{\frac{2}{\pi}} M(x) = \sqrt{\frac{2}{\pi}} 2{,}0844 \approx 1{,}66$

Вычислим теоретическую плотность, пользуясь свойством распределения Рэлея, результат приведем в таблице:

k	x_i	w_i	$f(x_i)$	$Hf(x_i)=p_i$	w_i-p_i	$(w_i-p_i)^2$	$(w_i-p_i)^2/p_i$
1	0,8	0,1	0,2575183	0,1030073	-0,003007315	9,04394E-06	8,77991E-05
2	1,2	0,15	0,3342913	0,1337165	0,01628349	0,000265152	0,001982942
3	1,6	0,19	0,3640657	0,1456263	0,044373718	0,001969027	0,013521095
4	2	0,17	0,3508295	0,1403318	0,029668215	0,000880203	0,0062723
5	2,4	0,13	0,3063189	0,1225276	0,007472421	5,58371E-05	0,00045571
6	2,8	0,09	0,2454186	0,0981674	-0,008167443	6,67071E-05	0,000679524
7	3,2	0,07	0,1817927	0,0727171	-0,0027171	7,38263E-06	0,000101525
8	3,6	0,05	0,1251113	0,0500445	-4,45143E-05	1,98152E-09	3,95952E-08
9	4	0,03	0,0802623	0,0321049	-0,002104925	4,43071E-06	0,000138007
10	4,4	0,01	0,0481119	0,0192448	-0,009244752	8,54654E-05	0,004440974
11	4,8	0,009	0,0269947	0,0107979	-0,001797896	3,23243E-06	0,000299357
12	5,2	0,001	0,0141961	0,0056784	-0,00467845	2,18879E-05	0,003854554
сумма		1		0,9339646			0,031833827

Получаем $\chi^2 = 50 \cdot 0{,}03183 \approx 1{,}59$.

Сравним графики полигона относительных частот и плотности распределения Рэлея:

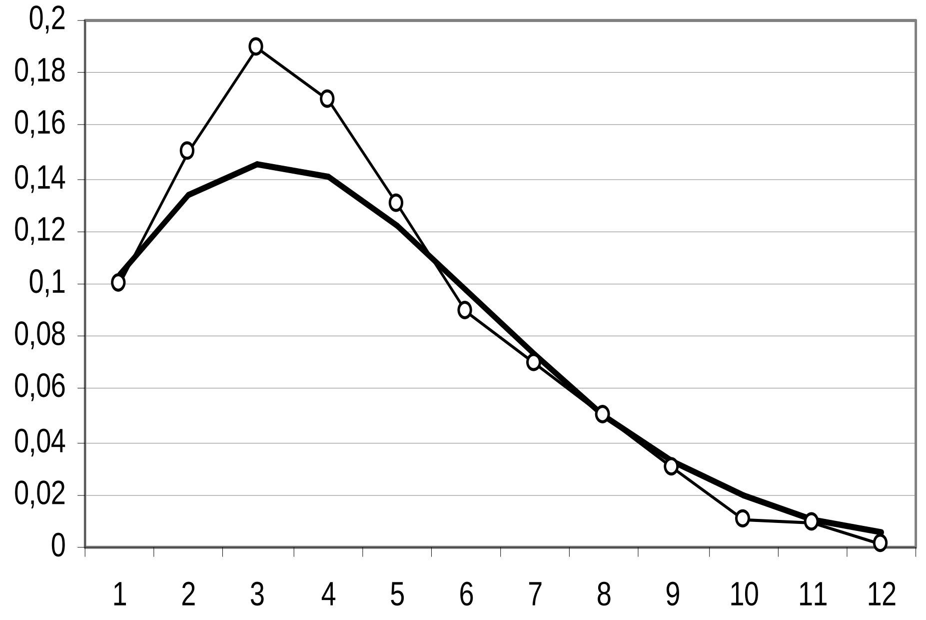


Рисунок 23 – Графики полигона относительных частот и плотности распределения Рэлея

Этап 4. Из таблицы приложения 8, при уровне значимости $\alpha = 0{,}1$ *и числу степеней свободы* $r = 12 - 2 = 10$, *имеем* $\chi^2_{0,1;10} = 4{,}86$. *Таким образом, так как* $\chi^2 = 1{,}59 < \chi^2_{0,1;10} = 4{,}86$ *гипотеза* H_0 *принимается, то есть данное распределение является и распределением Рэлея.*

Этап 5. Используя таблицу приложения 5, по значениям χ^2 *и r определяем величину p: если* $\chi^2 \approx 4$, *то p=0,9473; если* $\chi^2 \approx 5$, *то p=0,8913. Получаем, что гипотеза о распределении Рэлея не противоречит опытным данным.*

Воспользуемся критерием Романовского: $R = \dfrac{|\chi^2 - r|}{\sqrt{2r}} = \dfrac{|1{,}59 - 10|}{\sqrt{20}} \approx 1{,}88 < 3$.

Следовательно, этот критерий также подтверждает, что распределение является распределением Рэлея.

3.1.2. Критерий согласия Колмогорова

Критерий Колмогорова уместно применять в тех случаях, когда нужно проверить, подчиняется ли наблюдаемая случайная величина некоторому закону распределения, известному с точностью до параметров.

Основная идея критериев Колмогорова состоит в измерении меры расхождения между эмпирической функцией распределения и теоретической функцией распределения. В качестве меры расхождения рассматривается максимальное значение модуля разности между статистической функцией распределения $F^*(x)$ и соответствующей теоретической (интегральной) функцией распределения $F(x)$.

Интегральная функция распределения определяется соотношениями:

$$F(x)=\begin{cases}0, & при\ x<x_1,\\ \sum_{j=1}^{k} p_j, при\ x_k\le x\le x_{k+1} & (k=\overline{1,l-1}),\\ 1, & при\ x\ge x_l.\end{cases}$$

где $p_j=hf(x_j)\ (j=\overline{1,l})$, а $f(x)$ - плотность распределения случайной величины X.

Сначала находим величину: $\lambda = D\sqrt{n}$, где $D=\max|F^*(x)-F(x)|$, а n - объем выборки.

Затем из равенства $P(\lambda)=1-\sum_{j=-\infty}^{\infty}(-1)^j e^{-2j^2\lambda^2}$ определяют вероятность того, что за счет чисто случайных причин максимальное расхождение между $F^*(x)$ и $F(x)$ окажется не меньше, чем фактически наблюдаемое.

Если вероятность $P(\lambda)$ мала (меньше 0,05), то гипотезу следует отвергнуть, как неправдоподобную; при сравнительно больших значениях $P(\lambda)$ гипотезу можно считать совместимой с опытными данными.

Замечание. $P(\lambda)$ определяется из таблицы приложения 4.

Задача 16. Пользуясь критерием Колмогорова, установить, согласуется ли с нормальным распределением статистическое распределение.

Интервал	(0;2]	(2;4]	(4;6]	(6;8]	(8;10]	(10;12]	(12;14]	(14;16]	(16;18]	(!8;20]
n_x	10	29	51	58	102	90	81	39	30	10

Решение.

Имеем объем выборки $n=500$. *Определим середины в каждом интервале и вычислим относительные частоты* w_i*:*

K	*интервал*	n_x	x_i	w_i
1	(0;2]	10	1	0,02
2	(2;4]	29	3	0,06
3	(4;6]	51	5	0,1
4	(6;8]	58	7	0,12
5	(8;10]	102	9	0,2
6	(10;12]	90	11	0,18
7	(12;14]	81	13	0,16
8	(14;16]	39	15	0,08
9	(16;18]	30	17	0,06
10	(18;20]	10	19	0,02

сумма		500		1

Для применения критерия Колмогорова выполним предварительные вычисления:

- математическое ожидание – $a = M(x) = \sum_{i=1}^{k} x_i w_i = 10$*;*

- дисперсия - $D(x) = \sum_{i=1}^{k} x_i^2 w_i - M^2(x) = 16{,}52$*;*

-среднее квадратическое отклонение - $\sigma(x) = \sqrt{D(x)} = 4{,}06$.

Вычислим статистическую F(x) и теоретическую F(x) функции распределения, пользуясь свойством нормального распределения. Результат приведем в таблице:*

k	*интервал*	n_x	x_i	w_i	$f(x_i)$	$hf(x_i)$	$F^*(x_i)$	$F(x_i)$	$F^*(x_i)-F(x_i)$
1	(0;2]	10	1	0,02	0,008457	0,016913	0,02	0,016913	0,003087
2	(2;4]	29	3	0,06	0,022275	0,044551	0,08	0,061464	0,018536
3	(4;6]	51	5	0,1	0,046057	0,092113	0,18	0,153577	0,026423
4	(6;8]	58	7	0,12	0,074749	0,149498	0,3	0,303075	-0,00308
5	(8;10]	102	9	0,2	0,095227	0,190454	0,5	0,493529	0,006471
6	(10;12]	90	11	0,18	0,095227	0,190454	0,68	0,683983	-0,00398
7	(12;14]	81	13	0,16	0,074749	0,149498	0,84	0,833481	0,006519
8	(14;16]	39	15	0,08	0,046057	0,092113	0,92	0,925594	-0,00559
9	(16;18]	30	17	0,06	0,022275	0,044551	0,98	0,970145	0,009855
10	(18;20]	10	19	0,02	0,008457	0,016913	1	0,987058	0,012942
сумма		*500*		*1*		*0,987058*			

Из таблицы видно, что почти все значения относительных частот близки к соответствующим значениям вероятностей, найденных с помощью плотности нормального распределения.

Вычислим $D = \max|F^*(x) - F(x)| \approx 0{,}026$ *и* $\lambda = 0{,}026\sqrt{500} \approx 0{,}58$. *Тогда из таблицы приложения 4 имеем P(0,58)=0,90, следовательно, гипотеза о нормальном распределении является совместимой с опытными данными.*

Таким образом, можем утверждать, что верхняя граница абсолютной ошибки приближенного равенства $F^*(x) \approx F(x)$ *будет менее 0,026 для любого значения x. Следовательно, опытные данные представляют нормальное распределение.*

3.2. Частные критерии согласия

3.2.1. Критерий Шапиро-Уилка – критерий проверки экспоненциальности распределения

Пусть имеется выборка $x_1 \le x_2 \le \ldots \le x_n$, причем начальная точка распределения неизвестна, то есть рассматривается плотность вероятностей: $f(x,a) = \frac{1}{a} e^{\left(-\frac{x}{a}\right)}$. Статистика критерия Шапиро-Уилка имеет вид: $W_E = \frac{n(\bar{x} - x_1)^2}{(n-1)\sum_{i=1}^{n}(x_i - \bar{x})^2}$ или при $n \to \infty$ $(n > 50)$ $W_E = \frac{(\bar{x} - x_1)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$.

Гипотеза экспоненциальности наблюдаемого распределения не отклоняется с достоверностью α, если $W_1(\alpha) \le W_E \le W_2(\alpha)$, где $W_1(\alpha)$ и $W_2(\alpha)$ - критические значения, приведенные в таблице 1.

Возможны случаи:

1. Если выборка цензурирована (то есть отсутствуют r_1 наименьших и r_2 наибольших членов выборки). Тогда статистика критерия имеет вид: $$W_1 = \frac{\left(\sum_{i=2}^{n-r_1-r_2} T_{r_1+i}\right)^2}{(n-r_1-r_2)\sum_{i,j=2}^{n-r_1-r_2} a_{ij}^{(n-r_1-r_2)} T_{r_{1+i}} T_{r_{1+j}}},$$
   где $T_i = (n-i+1)(x_i - x_{i-1})$, $a_{ij}^{(n)} = \frac{i-1}{n-j-1}$,
   $(n \ge i \ge j \ge 2,\ a_{ij}^{(n)} = a_{ji}^{(n)},\ i, j = \overline{2,n})$.

   Критические значения статистики W_1 находится из таблицы 1 с заменой n на $n - r_1 - r_2$, то есть гипотеза экспоненциальности не отклоняется, если $W_{1(n-r_1-r_2)}(\alpha) \le W_1 \le W_{2(n-r_1-r_2)}(\alpha)$.

Таблица 1.

Критические значения
критерия экспоненциальности W_E Шапиро-Уилка

n	$\alpha = 0{,}90$		$\alpha = 0{,}95$		n	$\alpha = 0{,}90$		$\alpha = 0{,}95$	
	$W_1(\alpha)$	$W_2(\alpha)$	$W_1(\alpha)$	$W_2(\alpha)$		$W_1(\alpha)$	$W_2(\alpha)$	$W_1(\alpha)$	$W_2(\alpha)$
7	0,071	0,358	0,062	0,404	22	0,026	0,084	0,023	0,094
8	0,062	0,301	0,054	0,342	23	0,025	0,078	0,022	0,087

Продолжение таблицы 1.

9	0,058	0,261	0,050	0,301	24	0,024	0,074	0,021	0,082
10	0,056	0,231	0,049	0,261	25	0,023	0.070	0,021	0,078
11	0,052	0,208	0,046	0,234	26	0,022	0,066	0,020	0,073
12	0,050	0,191	0,044	0,215	27	0,022	0,063	0,020	0,070
13	0,046	0,173	0,040	0,195	28	0,021	0,061	0,019	0,067
14	0,043	0,159	0,038	0,178	29	0,021	0,058	0,019	0,064
15	0,040	0,145	0,036	0,163	30	0,020	0,054	0,018	0,060
16	0,038	0,134	0,034	0,150	31	0,019	0,052	0,017	0,057
17	0,034	0,120	0,030	0,135	32	0,019	0,050	0,017	0,055
18	0,031	0,109	0,028	0,123	33	0,018	0,048	0,017	0,053
19	0,029	0,102	0,026	0,114	34	0,018	0,047	0,017	0,051
20	0,028	0,095	0,025	0,106	35	0,018	0,045	0,016	0,049
21	0,027	0,091	0,024	0,101					

Тогда статистика имеет вид:

$$W_{E_0} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\left(\sum_{i=1}^{n} x_i\right)^2} \quad \text{или} \quad \tilde{W}_{E_0} = \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n\left((n+1)\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right)}$$

(причем критические значения совпадут с $W_1(\alpha)$ и $W_2(\alpha)$(в таблице n заменить на n+1)).

При этом гипотеза экспоненциальности не отклоняется, если $\tilde{W}_1(\alpha) \le W_{E_0} \le \tilde{W}_2(\alpha)$, $\tilde{W}_1(\alpha)$ и $\tilde{W}_2(\alpha)$ - критические значения, приведенные в таблице 2.

Таблица 2.

Критические значения
критерия экспоненциальности W_{E_0} Шапиро-Уилка

n	$\alpha = 0{,}90$		$\alpha = 0{,}95$		n	$\alpha = 0{,}90$		$\alpha = 0{,}95$	
	$\tilde{W}_1(\alpha)$	$\tilde{W}_2(\alpha)$	$\tilde{W}_1(\alpha)$	$\tilde{W}_2(\alpha)$		$\tilde{W}_1(\alpha)$	$\tilde{W}_2(\alpha)$	$\tilde{W}_1(\alpha)$	$\tilde{W}_2(\alpha)$
7	0,033	0,225	0,025	0,260	22	0,022	0,069	0,020	0,080
8	0,032	0,200	0,025	0,230	23	0,021	0,065	0,019	0,075
9	0,031	0,177	0,025	0,205	24	0,021	0,062	0,019	0,069
10	0,030	0,159	0,025	0,184	25	0,020	0.058	0,018	0,065
11	0,030	0,145	0,025	0,166	26	0,020	0,056	0,018	0,062
12	0,029	0,134	0,025	0,153	27	0,020	0,054	0,017	0,058
13	0,028	0,124	0,025	0,140	28	0,019	0,052	0,017	0,056
14	0,027	0,115	0,024	0,128	29	0,019	0,050	0,016	0,054
15	0,026	0,106	0,024	0,119	30	0,019	0,048	0,016	0,053

Продолжение таблицы 2.

16	0,025	0,098	0,023	0,113	31	0,018	0,047	0,016	0,051
17	0,024	0,093	0,023	0,107	32	0,018	0,045	0,015	0,050
18	0,024	0,087	0,022	0,101	33	0,018	0,044	0,015	0,048
19	0,023	0,083	0,022	0,096	34	0,017	0,043	0,014	0,046
20	0,023	0,077	0,021	0,090	35	0,017	0,041	0,014	0,045
21	0,022	0,074	0,020	0,085					

2. Если выборка цензурирована справа, то статистика заменяется на

$$W_2 = \frac{\left(\sum_{i=2}^{n-r_2+1} T_{i-1}\right)^2}{(n-r_2)\sum_{i=2}^{n-r_2+1}\sum_{j=2}^{n-r_2+1} a_{ij}^{(n-r_2+1)} T_{r_{i-1}} T_{r_{j-1}}}.$$ Ее критические значения совпадают с критическими значениями статистики W_{E_0}.

Задача 17. Проверить гипотезу экспоненциальности для ряда наблюдений: x_i: 1, 2, 4, 5, 9, 11, 18, 21, 29, 35. При условии, что первое и два последних наблюдений цензурированы, $\alpha = 0{,}95$.

Решение.

Имеем 1 случай $n = 10$, $r_1 = 1$ *и* $r_2 = 2$ $(n - r_1 - r_2 = 7)$. *Тогда вычислим:*

$T_3 = (10-3+1)(x_3 - x_2) = 16$, $T_4 = (10-4+1)(x_4 - x_3) = 7$, $T_5 = (10-5+1)(x_5 - x_4) = 24$,

$T_6 = (10-6+1)(x_6 - x_4) = 10$, $T_7 = (10-7+1)(x_7 - x_6) = 28$, $T_8 = (10-8+1)(x_8 - x_7) = 9$,

$$a_{ij}^{(n-r_1-r_2)} = a_{ij}^{(7)} = \frac{i-1}{7-j-1} = \frac{i-1}{8-j}:$$

$$a_{22}^{(7)} = \frac{1}{6};\ a_{32}^{(7)} = \frac{2}{6};\ a_{42}^{(7)} = \frac{3}{6};\ a_{52}^{(7)} = \frac{4}{6};\ a_{62}^{(7)} = \frac{5}{6};\ a_{72}^{(7)} = \frac{6}{6};\ a_{33}^{(7)} = \frac{2}{5};$$

$$a_{43}^{(7)} = \frac{3}{5};\ a_{53}^{(7)} = \frac{4}{5};\ a_{63}^{(7)} = \frac{5}{5};\ a_{73}^{(7)} = \frac{6}{5};\ a_{44}^{(7)} = \frac{3}{4};\ a_{54}^{(7)} = \frac{4}{4};\ a_{64}^{(7)} = \frac{5}{4};$$

$$a_{74}^{(7)} = \frac{6}{4};\ a_{55}^{(7)} = \frac{4}{3};\ a_{65}^{(7)} = \frac{5}{3};\ a_{75}^{(7)} = \frac{6}{3};\ a_{66}^{(7)} = \frac{5}{2};\ a_{76}^{(7)} = \frac{6}{2};\ a_{77}^{(7)} = \frac{6}{1}.$$

Получаем:

$$7\left(\frac{1}{6}T_3T_3 + \frac{2}{6}T_3T_4 + \frac{3}{6}T_3T_5 + \ldots + \frac{6}{3}T_6T_7 + \frac{5}{2}T_7T_7 + \frac{6}{2}T_7T_8\right) =$$

$$= 7\left(\frac{16\cdot 16}{6} + \frac{16\cdot 7}{3} + \ldots + \frac{5\cdot 28\cdot 28}{2} + 3\cdot 28\cdot 9\right) = 40056$$

$$(T_3 + T_4 + \ldots + T_7 + T_8)^2 = (16 + 7 + \ldots + 28 + 9)^2 = 8836.$$

Следовательно, $W_E = \frac{8836}{40056} = 0{,}22$. *Из таблицы 1, для* $n - r_1 - r_2 = 7$ *и* $\alpha = 0{,}95$, *имеем* $W_1(0{,}95) = 0{,}062$ *и* $W_2(0{,}95) = 0{,}404$.

Тогда $W_1(0{,}95)=0{,}062<W_E=0{,}22<W_2(0{,}95)=0{,}404$, *то есть гипотеза экспоненциальности распределения вероятностей случайных величин не отклоняется.*

3.2.2. Критерий наибольшего интервала – критерий проверки экспоненциальности распределения

Статистика критерия имеет вид: $\eta_n=\dfrac{\max\limits_{1\le i\le n}(x_i-x_{i-1})}{\sum\limits_{i=1}^{n}x_i}$.

Статистика η_n совпадает со статистикой Кохрана для проверки однородности нескольких дисперсий (при числе степеней свободы $r=2$). Критические значения $\eta_n(\alpha)$ приведены в таблице 3.

Таблица 3

Критические значения статистики $\eta_n(\alpha)$

n	Доверительная вероятность α		n	Доверительная вероятность α	
	0,95	0,99		0,95	0,99
2	0,975	0,995	12	0,392	0,475
3	0,871	0,942	15	0,335	0,407
4	0,768	0,864	20	0,270	0,330
5	0,684	0,788	24	0,235	0,287
6	0,616	0,722	30	0,198	0,241
7	0,561	0,664	40	0,158	0,191
8	0,561	0,664	60	0,113	0,137
9	0,477	0,573	120	0,063	0,076
10	0,445	0,536			

Задача 18. Проверить гипотезу экспоненциальности распределения критерием наибольшего интервала для ряда наблюдений

x_i: 1, 2, 4, 5, 9, 11, 18, 21, 29, 35 при α=0,95.

Решение.

Имеем $\sum\limits_{i=1}^{10}x_i=135,\ \max(x_i-x_{i-1})=8\Rightarrow\eta_n=\dfrac{8}{135}=0{,}0592$.

По таблице 3 $\eta_n(0{,}95)=0{,}445$.

Тогда, так как $\eta_n=0{,}0592<\eta_n(0{,}95)=0{,}445$, *гипотеза экспененциальности не отклоняется.*

3.2.3. Критерий Манн-Фертига-Шуера – критерий проверки распределения Вейбулла

Распределение Вейбулла является обобщающим для экспоненциального распределения. Поэтому распределение Вейбулла рассматривают как альтернативное при проверки экспоненциальности распределения.

Замечание. Критерий был предложен авторами применительно к задаче испытаний изделий на долговечность.

Если $t_1, t_2, ..., t_r$ - первые r порядковых статистик наработки на отказ при испытаниях выборки изделий объема $n \geq r$, то статистика критерия имеет вид: $K = \dfrac{\sum\limits_{i=\left[\frac{r}{2}+1\right]}^{r-1} a_i^{-1}(x_{i+1} - x_i)}{\sum\limits_{i=1}^{r} a_i^{-1}(x_{i+1} - x_i)}$, где $x_i = \ln(t_i)$, $\left[\dfrac{r}{2}\right]$- наибольше целое число $\leq \dfrac{r}{2}$, a_i - коэффициенты, приведенные в таблице приложения 7.

Гипотеза согласия эмпирического распределения с двухпараметрическим распределением Вейбулла отклоняется, если $K > K_\alpha(r, n)$, где $K_\alpha(r, n)$ - критическое значение статистики для доверительной вероятности α (при известных результатах отказов r изделий из n), приведенное в таблице приложения 10.

Задача 19. Проверить гипотезу согласия для ряда наблюдений t_i: 1, 2, 4, 5, 9, 11, 18, 21, 29, 35 с двухпараметрическим распределением Вейбулла при доверительной вероятности α =0,95.

Решение.

Для ряда значений x_i имеем ряд:

$\ln(t_i)$: 0; 0,693; 1,386; 1,609; 2,197; 2,398; 2,890; 3,044; 3,367; 3,555.

С учетом коэффициентов a_i из таблицы приложения 7 для n=r=10 имеем

$$\sum_{i=6}^{9} \frac{\ln(t_{i+1}) - \ln(t_i)}{a_i} = \frac{2,890 - 2,398}{0,269493} + \frac{3,044 - 2,890}{0,271645} + \frac{3,367 - 3,044}{0,300869} + \frac{3,555 - 3,367}{0,405316} = 3,9299959$$

$$\sum_{i=1}^{9} \frac{\ln(t_{i+1}) - \ln(t_i)}{a_i} = \frac{0,693 - 0}{0,1053606} + \frac{1,386 - 0,693}{0,559013} + \frac{1,609 - 1,386}{0,399100} + \frac{2,197 - 1,609}{0,324470} + \frac{2,398 - 2,197}{0,286163} +$$
$$+ 3,9299959 = 8,901$$

$$K = \frac{3,9299959}{8,901} = 0,441$$

Из таблицы приложения 10 для n=r=10 имеем $K_{0,95}(10;10) = 0,69$.

Так как $K = 0,441 < K_{0,95}(10;10) = 0,69$, *нулевая гипотеза не отклоняется. Она включает в себя и гипотезу экспоненциального распределения, так как*

экспоненциальное распределение является частным случаем распределения Вейбулла.
Рассмотрим случай цензурированных испытаний при r=7<n=10. Для этого случая имеем $\sum_{i=4}^{6}\frac{\ln(t_{i+1})-\ln(t_i)}{a_i}=4{,}3402333$; $\sum_{i=1}^{6}\frac{\ln(t_{i+1})-\ln(t_i)}{a_i}=6{,}886416$;
$K=\frac{4{,}3402333}{6{,}886416}=0{,}630026$.
Из таблицы приложения 10 для r=7 и n=10 находим $K_{0,95}(7;10)=0{,}81$.
Так как $K=0{,}630026<K_{0,95}(7;10)=0{,}81$, *нулевая гипотеза не отклоняется и в этой ситуации.*

3.2.4. Критерий Шермана – критерий проверки равномерного распределения

Статистика критерия Шермана для проверки равномерности распределения имеет вид: $\omega_{n=1}=\frac{1}{2}\sum_{i=1}^{n+1}\left|U_i-U_{i-1}-\frac{1}{n+1}\right|,\ U_0=1,U_n=1$.

Критические значения статистики Шермана приведены в таблице 4.

Таблица 4

Критические значения $\omega_n(\alpha)$ статистики Шермана

n	α			n	α		
	0,9	0,95	0,99		0,9	0,95	0,99
1	0,450	0,475	0,495	11	0,442	0,469	0,521
2	0,484	0,537	0,609	12	0,440	0,466	0,516
3	0,467	0,518	0,614	13	0,438	0,463	0,511
4	0,468	0,509	0,589	14	0,436	0,460	0,506
5	0,462	0,502	0,574	15	0,434	0,458	0,502
6	0,458	0,494	0,562	16	0,433	0,455	0,498
7	0,454	0,488	0,551	17	0,431	0,453	0,495
8	0,451	0,482	0,542	18	0,430	0,451	0,491
9	0,448	0,477	0,534	19	0,429	0,449	0,489
10	0,445	0,473	0,527	20	0,427	0,448	0,486

При n>20 критерий Шермана удовлетворительно аппроксимируется нормальным распределением со средним $M(\omega_n)$ и дисперсией $D(\omega_n)$, где
$M(\omega_n)=\left(\frac{n}{n+1}\right)^{n+1}$; $D(\omega_n)=\frac{2n^{n+2}+n(n-1)^{n+2}}{(n+2)(n+1)^{n+2}}-\left(\frac{n}{n+1}\right)^{2n+2}$.

Следовательно, случайная величина $\tilde{\omega}_n = \dfrac{\omega_n - M(\omega_n)}{\sqrt{D(\omega_n)}}$ имеет стандартное нормальное распределение. Если $\omega_n > \omega_n(\alpha)$, где $\omega_n(\alpha)$ - статистика F-распределения Фишера, то с доверительной вероятностью α гипотеза отклоняется.

Задача 20. Имеет ряд наблюдений над случайной величиной:

U_i : 0,047; 0,05; 0,15; 0,18; 0,29; 0,48; 0,52; 0,61; 0,72; 0,91. Необходимо проверить гипотезу равномерности распределения случайной величины U_i на интервале (0; 1) с помощью критерия Шермана при доверительной вероятности α =0,95.

Решение.

Получаем

$$\omega_n = \frac{1}{2}\left(\left|0{,}047 - \frac{1}{11}\right| + \left|0{,}5 - 0{,}047 - \frac{1}{11}\right| + \ldots + \left|0{,}72 - 0{,}61 - \frac{1}{11}\right| + \left|0{,}91 - 0{,}72 - \frac{1}{11}\right|\right) = 0{,}245.$$

Из таблицы 4 для n=10 имеем $\omega_{10}(0{,}95) = 0{,}473$.

Так как $\omega_n = 0{,}245 < \omega_{10}(0{,}95) = 0{,}473$, *гипотеза равномерности не отклоняется.*

Рассмотрим нормальную аппроксимацию ω_n*-распределения.*

Имеем $M(\omega_n) = \left(\dfrac{n}{n+1}\right)^{n+1} = \left(\dfrac{10}{11}\right)^{11} = 0{,}35049$*;*

$$D(\omega_n) = \frac{2n^{n+2} + n(n-1)^{n+2}}{(n+2)(n+1)^{n+2}} - \left(\frac{n}{n+1}\right)^{2n+2} = \frac{2 \cdot 10^{12} + 10 \cdot 9^{12}}{12 \cdot 11^{12}} - \left(\frac{10}{11}\right)^{22} = 0{,}005251.$$

Вычисляем $\tilde{\omega}_n = \dfrac{\omega_n - M(\omega_n)}{\sqrt{D(\omega_n)}} = \dfrac{0{,}245 - 0{,}35049}{\sqrt{0{,}005251}} = -1{,}457$.

Так как $|\tilde{\omega}_n| = 1{,}457 < \omega_{0,95} = 1{,}645$ *(95%-я квантиль стандартного нормального распределения), гипотеза равномерности распределения не отклоняется.*

Глава 4. Сравнение параметров распределения

Проверка научной гипотезы применительно к потребностям исследователя в ходе наблюдений зависит от специфики наблюдаемых процессов и потребностей практики.

Например, часто встречается техническая задача – проверка соответствия параметров разработанного изделия предъявляемым требованиям, в математико-статистической формулировке она звучит так: «необходимо проверить гипотезу о том, что параметр a распределения случайной величины x превосходит заданную величину a_0».

4.1. Сравнение двух средних значений

Имеются две выборки независимых случайных величин $x_1, x_2, ..., x_n$ и $y_1, y_2, ..., y_m$. Необходимо на основе выборочных данных установить наличие значимой разницы в средних значениях двух совокупностей, из которых извлечены выборки, то есть проверить нулевую гипотезу H$_0$: $a_1 = a_2$ против альтернатив H_1: $a_1 \neq a_2$, H_1': $a_1 < a_2$ и H_1'': $a_1 > a_2$.

4.1.1. Сравнение при известных дисперсиях σ_1^2 и σ_2^2

Статистика критерия проверки нулевой гипотезы имеет вид

$$z = (\bar{x} - \bar{y})\left(\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right)^{-\frac{1}{2}}, \text{ где } \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i;\ \bar{y} = \frac{1}{m}\sum_{i=1}^{m} y_i.$$

При справедливости нулевой гипотезы z-статистика имеет нормальное распределение. Гипотеза H$_0$: $a_1 = a_2$ предпочитается альтернативе H$_1$: $a_1 \neq a_2$ с достоверной вероятностью α, если $|z| < u_{(1+\alpha)/2}$; альтернативе H_1': $a_1 < a_2$, если $z > u_{1-\alpha}$ и альтернативе H_1'': $a_1 > a_2$, если $z < u_\alpha$ (u_γ - γ-квантиль стандартного нормального распределения).

Задача 21. Имеются две выборки случайных величин:

$(n = 10)$ x_i : 1,2; 2,1; 3,2; 3,6; 3,8; 4,4; 6,1; 7,1; 9; 10,2;

$(m = 8)$ y_i : 2,4; 2,8; 4,1; 4,4; 6,8; 7,2; 8,9.

Известны дисперсии $\sigma_x^2 = 8{,}7$ и $\sigma_y^2 = 6{,}1$. Проверить гипотезу равенства средних значений H$_0$: $a_1 = a_2$ в двух выборках при доверительной вероятности α =0,95 против альтернативы H$_1$: $a_1 \neq a_2$.
Решение.

Вычислим оценки $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 5{,}07$ *и* $\overline{y} = \frac{1}{m}\sum_{i=1}^{m} y_i = 4{,}575$, *тогда статистика нулевой гипотезу равна* $z = (5{,}07 - 4{,}575)\left(\frac{8{,}7}{10} + \frac{6{,}1}{8}\right)^{-\frac{1}{2}} = 0{,}387$.

Для α *=0,95 имеем, согласно таблицы приложения 3, значение* $u_{\frac{1+\alpha}{2}} = u_{0{,}975} = 1{,}96$.

Так как $|z| = 0{,}387 < u_{0{,}975} = 1{,}96$, *нулевая гипотеза равенства средних не отклоняется.*

4.1.2. Модифицированный критерий Стьюдента

Статистика критерия $T = 2\frac{\overline{x} - \overline{y}}{\omega_1 + \omega_2}$, где $\omega_1 = x_{\max} - x_{\min}$ и $\omega_2 = y_{\max} - y_{\min}$ - размахи сравниваемых выборок.

Критические значения T_α статистики для выборок одинакового объема n=m приведены в таблице 5.

Таблица 5

Критические значения модифицированного критерия Стьюдента
для сравнения средних по двум выборкам равного объема

n	Доверительная вероятность α				n	Доверительная вероятность α			
	0,95	0,975	0,99	0,995		0,95	0,975	0,99	0,995
2	2,322	3,427	5,553	7,916	12	0,214	0,260	0,315	0,355
3	0,974	1,272	1,715	2,093	13	0,201	0,243	0,294	0,331
4	0,644	0,813	1,047	1,237	14	0,189	0,228	0,276	0,311
5	0,493	0,613	0,772	0,896	15	0,179	0,216	0,261	0,293
6	0,405	0,499	0,621	0,714	16	0,170	0,205	0,247	0,278
7	0,347	0,426	0,525	0,600	17	0,162	0,195	0,236	0,264
8	0,306	0,373	0,459	0,521	18	0,155	0,187	0,225	0,252
9	0,275	0,334	0,409	0,464	19	0,149	0,179	0,216	0,242
10	0,250	0,304	0,371	0,419	20	0,143	0,172	0,207	0,232

Если $|T| < T_{\frac{1+\alpha}{2}}$, то гипотеза H$_0$: $a_1 = a_2$ предпочитается альтернативе H$_1$: $a_1 \neq a_2$; если $T > T_\alpha$ - альтернативе H_1': $a_1 < a_2$; если $T < T_\alpha$ - альтернативе H_1'': $a_1 > a_2$ (α - доверительная вероятность).

Замечание. При $n \le 10$ критерий не уступает по эффективности обычному критерию Стьюдента. При $n \ge 20$ критерием пользоваться не рекомендуется.

Задача 22. Имеются две выборки объема n=m=10:

x_{i1} : 2, 4, 6, 7, 9, 12, 14, 16, 19, 24; x_{i2} : 9, 14, 19, 21, 25, 29, 35, 41, 46, 50.

Необходимо проверить гипотезу равенства средних модифицированным критерием Стьюдента при α =0,95 против альтернативы H$_1$: $a_1 \neq a_2$.

Решение.

Имеем $\overline{x_1} = 11{,}3;\ \overline{x_2} = 28{,}9;\ \omega_1 = 24 - 2 = 22;\ \omega_2 = 50 - 9 = 41.$

Тогда $T = 2\dfrac{11{,}3 - 28{,}9}{22 + 41} = -0{,}559$.

Из таблицы 5 для $\dfrac{1+\alpha}{2} = \dfrac{1+0{,}95}{2} = 0{,}975$ *находим критическое значение T-статистики для n=10:* $T_{0{,}975} = 0{,}304$.

Так как $|T| = 0{,}559 > T_{0{,}975} = 0{,}304$, *нулевая гипотеза равенства средних отклоняется.*

4.1.3. Критерий Кохрана-Кокса (сравнение при неизвестных неравных дисперсиях)

Статистика критерия $t_K = \dfrac{1}{s}(\overline{x} - \overline{y})$, где $s^2 = \dfrac{s_1^2}{n} + \dfrac{s_2^2}{m}$.

Критические значения статистики вычисляются по формуле

$t'_\alpha = \dfrac{v_1 t_\alpha(f_1) + v_2 t_\alpha(f_2)}{v_1 + v_2}$, где $v_1 = \dfrac{s_1^2}{n}$ и $v_2 = \dfrac{s_2^2}{m}$; $t_\alpha(f)$ - α-квантиль распределения Стьюдента с f степенями свободы (в частности $f_1 = n-1,\ f_2 = m-1$).

Задача 23. Имеются две выборки данных:

$(n = 10)\ x_i$: 2, 4, 6, 7, 9, 12, 14, 16, 19, 24;

$(m = 9)\ y_i$: 9, 14, 19, 21, 25, 29, 35, 41, 46.

Необходимо проверить гипотезу равенства средних при достоверности α =0,95.

Решение.

Имеем $\overline{x} = 11{,}33;\ \overline{y} = 26{,}55;\ s_1^2 = \dfrac{1}{9}\sum_{i=1}^{10}(x_i - \overline{x})^2 = 49{,}12;\ s_2^2 = \dfrac{1}{8}\sum_{i=1}^{9}(y_i - \overline{y})^2 = 152{,}52$;

$v_1 = \frac{s_1^2}{n} = 4{,}912;\ v_2 = \frac{s_2^2}{m} = 16{,}947;\ s^2 = v_1 + v_2 = 21{,}8595;\ s = 4{,}675.$

По таблице приложения 2 значение $t_{0,95}(f_1 = n-1 = 9) = 1{,}833$ *и* $t'_{0,95}(f_2 = m-1 = 8) = 1{,}86$.

Вычисляем статистику критерия $t_K = \frac{26{,}55 - 11{,}33}{4{,}675} = 3{,}255$.

Вычислим критическое значение $t'_{0,95} = \frac{4{,}912 \cdot 1{,}833 + 16{,}947 \cdot 1{,}86}{4{,}912 + 16{,}947} = 1{,}853$.

Так как $t_K = 3{,}255 > t_{0,95} = 1{,}853$, *нулевая гипотеза равенства средних отклоняется.*

4.1.4. Сравнение нескольких средних. Дисперсионный критерий

Имеются k выборок равного объема n из нормально распределенной совокупности $x_{11}, x_{12}, ..., x_{1n_1}; x_{21}, ..., x_{2n_2}; ...; x_{k1}, ..., x_{kn_k}$. Проверке подлежит нулевая гипотеза о статистической неразличимости средних $H_0 : a_1 = a_i$ против альтернативы $H_1 : |a_{i+1} - a_i| > 0\ (i = \overline{1,k})$.

Статистика критерия имеет вид $F = \frac{kn(n-1)}{k-1} \frac{\sum_{j+1}^{k} (\overline{x_j} - \overline{x})^2}{\sum_{j=1}^{k} \sum_{i=1}^{n} (x_{ij} - \overline{x_j})^2}$,

где $\overline{x_j} = \frac{1}{n} \sum_{i=1}^{n} x_{ij}$, $\overline{x} = \frac{1}{kn} \sum_{j=1}^{k} \sum_{i=1}^{n} x_{ij}$.

При справедливости нулевой гипотезы статистика критерия имеет распределение Фишера с $f_1 = k-1$ и $f_2 = k(n-1)$ степенями свободы.

При $F > F_\alpha(k-1; k(n-1))$ нулевая гипотеза отклоняется. Здесь $F_\alpha(f_1, f_2)$ - α-квантиль F-распределения, ее значения могут быть найдены по соответствующим таблицам приложения 9.

Замечание. Применение этого критерия при отсутствии нормальности исходных распределений величин x_{ij} не рекомендуется, так как он становится в этом случае неустойчивым.

Устойчивость к отклонениям от нормального распределения повышается, если использовать модифицированные степени свободы для F-

критерия: $f_1 = d(k-1)$ и $f_2 = dk(n-1)$, где $d = 1 + \frac{kn+1}{kn-1}\frac{c}{kn-c}$; $c = \frac{\lambda}{v^2}$; $v = \frac{s_2}{kn-1}$;
$\lambda = \frac{kn(kn+1)s_4 - 3(kn-1)s_2^2}{(kn-1)(kn-2)(kn-3)}$; $s_m = \sum_{j=1}^{k}\sum_{i=1}^{n}(x_{ij} - \overline{x})^m$.

Задача 24. В результате испытаний пяти выборок приборов объемом n=8 каждая, изготовленных разными заводами, получены следующие значения долговечности приборов (ч):

x_{1i}:	11	14	18	21	30	32	40	45
x_{2i}:	15	19	21	22	26	38	41	52
x_{3i}:	8	11	14	19	31	32	44	58
x_{4i}:	4	5	9	18	24	31	45	61
x_{5i}:	24	26	32	48	54	62	66	70.

Проверить гипотезу равенства средних при доверительной вероятности α=0,95.

Решение.

1 способ.

Вычислим $\overline{x_1} = 26{,}375$; $\overline{x_2} = 29{,}25$; $\overline{x_3} = 27{,}125$; $\overline{x_4} = 24{,}625$; $\overline{x_5} = 47{,}75$;

$$\overline{x} = \frac{1}{5\cdot 8}\sum_{j=1}^{5}\sum_{i=1}^{8}x_{ij} = 31{,}025;\ \sum_{j=1}^{5}(\overline{x_j} - \overline{x})^2 = 363{,}66875;\ \sum_{j=1}^{5}\sum_{i=1}^{8}(x_{ij} - \overline{x_j})^2 = 9611{,}63.$$

Тогда $F = \frac{5\cdot 8\cdot 7}{5-1}\frac{363{,}66875}{9611{,}63} = 2{,}645$.

Для α *=0,95,* $f_1 = 5-1 = 4$ *и* $f_2 = 5(8-1) = 35$ *имеем из таблицы* $F_{0,95}(4;35) = 2{,}65$.

Так как $F = 2{,}645 < F_{0,95}(4;35) = 2{,}65$, *нулевая гипотеза принимается.*

2 способ.

Рассмотрим более устойчивый критерий. Для этого вычислим

$$s_2 = \sum_{j=1}^{5}\sum_{i=1}^{8}(x_{ij} - \overline{x_j})^2 = 9611{,}63;\ s_4 = \sum_{j=1}^{5}\sum_{i=1}^{8}(x_{ij} - \overline{x_j})^4 = 5137436{,}514;$$

$$\lambda = \frac{5\cdot 8\cdot 41\cdot 5137436{,}514 - 3\cdot 39\cdot 9611{,}63^2}{39\cdot 38\cdot 37} = -43465{,}698;$$

$$v = \frac{9611{,}63}{39} = 246{,}45;\ c = -\frac{43465{,}698}{246{,}45^2} = -0{,}7156;$$

$$d = 1 + \frac{41}{39}\cdot\frac{-0{,}7156}{40 + 0{,}7156} = 0{,}9816;$$

$$f_1 = 0{,}9816(5-1) = 3{,}92;\ f_2 = 0{,}9816\cdot 5\cdot 7 = 34{,}3.$$

Так как число степеней свободы снижается, то это позволяет отклонить гипотезу.

4.2. Сравнение двух дисперсий

Для двух выборок нормально распределённых случайных величин $x_1, x_2, ..., x_n$ и $y_1, y_2, ..., y_m$ необходимо проверить гипотезу равенства дисперсий σ_1^2 и σ_2^2, опираясь на их выборочные оценки s_1^2 и s_2^2.

4.2.1. Критерий Фишера

Если выборочными оценками максимального правдоподобия дисперсий являются $s_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ и $s_2^2 = \frac{1}{m-1}\sum_{j=1}^{m}(y_j - \bar{y})^2$, то статистика критерия Фишера записывается как

$$F = \frac{s_1^2}{s_2^2}.$$

При справедливости нулевой гипотезы $H_0 : \sigma_1^2 = \sigma_2^2$ статистика критерия имеет распределение Фишера с $f_1 = n-1$ и $f_2 = m-1$ степенями свободы, где n и m - объёмы сравниваемых выборок.

Если $F > F_{\frac{1+\alpha}{2}}(n-1, m-1)$ и $F < F_{\frac{1-\alpha}{2}}(n-1, m-1)$, то нулевая гипотеза отклоняется в пользу альтернативной $H_1 : \sigma_1^2 \neq \sigma_2^2$.

Если $F > F_{\alpha}(n-1, m-1)$, то нулевая гипотеза отклоняется в пользу альтернативы $H_1^{/} : \sigma_1^2 > \sigma_2^2$ ($\alpha -$ доверительная вероятность). В числителе всегда должна быть большая по величине из двух сравниваемых дисперсий.

Критерий Фишера очень чувствителен к отклонениям от нормального распределения x_i и y_j. Его устойчивость к отклонениям от нормальности может быть повышена соответствующей корректировкой степеней свободы. Вместо f_1, f_2 в этом случае используются степени свободы $f_1^{/} = d \cdot f_1$ $f_2^{/} = b \cdot f_2$, где

$$d = \left\{1 + \frac{1}{2}\left(\frac{n+m-4}{n+m-(b-3)}\right)(b-3)\right\}^{-1}, \quad b = (m+n)\frac{\sum_{i=1}^{n}(x_i - \bar{x})^4 + \sum_{j=1}^{m}(y_j - \bar{y})^4}{\left(\sum_{i=1}^{n}(x_i - \bar{x})^2 + \sum_{j=1}^{m}(y_j - \bar{y})^2\right)^2}.$$

В дальнейшем процедура проверки нулевой гипотезы не отличается от обычного *F*- критерия. Критические значения *F*-статистики приведены в таблице приложения 9.

Задача 25. Имеются две выборки нормально распределённых случайных величин (n=m=10)

x_i	2,1	3,1	4,8	6,1	7,4	8,5	10,1	12,1	14,0	15,6
y_j	4,6	6,1	8,2	9,8	9,9	10,4	13,1	14,5	16,1	19,1

Необходимо проверить гипотезу равенства дисперсий $H_0: \sigma_1^2 = \sigma_2^2$ против альтернативы $H_1: \sigma_1^2 \neq \sigma_2^2$ при доверительной вероятности α =0,95.

Решение.

Получаем $\bar{x} = 8{,}38$; $\bar{y} = 11{,}18$;

$$s_1^2 = \frac{1}{10-1}\sum_{i=1}^{10}(x_i - \bar{x})^2 = 20{,}75; \quad s_2^2 = \frac{1}{10-1}\sum_{j=1}^{10}(y_j - \bar{y})^2 = 20{,}419.$$

Далее $F = \frac{s_1^2}{s_2^2} = \frac{20{,}757}{20{,}419} = 1{,}016$. *Из таблицы приложения 9 находим*

$$F_{\frac{1+\alpha}{2}}(n-1, m-1) = F_{\frac{1+0{,}95}{2}}(10-1, 10-1) = F_{0{,}975}(9{,}9) = 4{,}03.$$

Так как $F = 1{,}016 < F_{0{,}975}(9{,}9) = 4{,}03$, *нулевая гипотеза принимается.*

Рассмотрим теперь критерий с откорректированными степенями свободы.

Имеем

$$\sum_{i=1}^{10}(x_i - \bar{x})^2 = 186{,}813; \sum_{j=1}^{10}(y_j - \bar{y})^2 = 183{,}771; \sum_{i=1}^{10}(x_i - \bar{x})^4 = 6439{,}996;$$

$$\sum_{j=1}^{10}(y_j - \bar{y})^4 = 7281{,}725;$$

$$b = (10+10)\frac{6439{,}996 + 7281{,}725}{(186{,}813 + 183{,}771)^2} = 1{,}99832;$$

$$d = \left\{1 + \frac{1}{2}\left(\frac{10+10-4}{10+10-(1{,}99832-3)}\right)(1{,}99832-3)\right\}^{-1} = 1{,}802.$$

Окончательно имеем $f_1^{/} = d \cdot f_1 = 1{,}802 \cdot 9 = 16{,}22$, $f_2^{/} = b \cdot f_2 = 16{,}22$.

Из таблиц для дробных степеней свободы $\frac{1+\alpha}{2} = 0{,}975$ *имеем*

$$F_{0{,}975}(16{,}22; 16{,}22) \approx 2{,}65.$$

Так как $F = 1{,}016 < F_{0{,}975}(16{,}22; 16{,}22) = 2{,}65$, *нулевая гипотеза принимается.*

4.2.2. Критерий Романовского

Статистика критерия $R = \frac{|Q-1|}{\sigma_Q}$, $Q = \frac{n-3}{m-1}\cdot\frac{s_1^2}{s_2^2}$, $\sigma_Q = \sqrt{\frac{2(m+n-4)}{(n-1)(m-5)}}$.

Если $R \geq 3$, то нулевая гипотеза равенства дисперсий отклоняется с достоверностью не менее 0, 89.

Задача 26. Имеются две выборки нормально распределённых случайных величин (n=m=10):

x_i	2,1	3,1	4,8	6,1	7,4	8,5	10,1	12,1	14,0	15,6
y_j	4,6	6,1	8,2	9,8	9,9	10,4	13,1	14,5	16,1	19,1

Проверить гипотезу равенства дисперсий критерием Романовского.
Решение.

$Q = \dfrac{10-3}{10-1}\cdot\dfrac{20{,}757}{20{,}419} = 0{,}7906$, $\sigma_Q = \sqrt{\dfrac{2(10+10-4)}{(10-1)(10-5)}} = 0{,}84327$, $R = \dfrac{|0{,}7906-1|}{0{,}84327} = 0{,}248$.

Так как $R < 3$, нулевая гипотеза равенства дисперсий принимается.

4.2.3. Критерий отношения размахов

Статистика критерии имеет вид $F^* = \dfrac{\omega_n}{\omega_m}$,

где $\omega_n = x_{\max} - x_{\min}$, $\omega_m = y_{\max} - y_{\min}$ - размахи сравниваемых выборок.

Если $F^* > F_\alpha{}^*$, где $F_\alpha{}^*$ - критическое значение статистики, то нулевая гипотеза отклоняется с вероятностью α. Критические значения $F_\alpha{}^*(n,m)$ приведены в таблице 6.

При $n \le 15$ мощность критерия отношения размахов практически не отличается от мощности критерия Фишера (в числителе всегда должно быть набольшее из двух значений размаха).

Таблица 6

Критические значения $F_\alpha{}^*(n,m)$ отношения размахов
(α-доверительная вероятность)

N	m	α		n	m	α		n	m	α	
		0,95	0,99			0,95	0,99			0,95	0,99
3	2	3,90	7,37	8	6	1,85	2,57	11	10	1,77	2,30
	3	4,37	9,99		7	1,94	2,67		11	1,81	2,34
	4	5,14	11,71		8	2,01	2,76		12	1,84	2,39
	5	5,71	12,98		9	2,08	2,84		13	1,88	2,42
4	3	2,03	3,72		10	2,13	2,91		14	1,91	2,46
	4	2,66	4,79		11	2,18	2,98		15	1,93	2,49
	5	3,07	5,49		12	2,23	2,56	12	10	1,69	2,18
	6	3,38	6,01	9	5	1,63	2,23		11	1,73	2,22
5	3	1,60	2,66		6	1,72	2,34		12	1,77	2,26
	4	2,06	3,32		7	1,80	2,43		13	1,80	2,30

	5	2,35	3,76		8	1,87	2,14		14	1,82	2,33
	6	2,57	4,08		9	1,93	2,20		15	1,85	2,36
	7	2,74	4,36		10	1,98	2,26	13	10	1,63	2,08
	8	3,00	4,73		11	2,02	2,31		11	1,67	2,12
6	4	1,99	2,98		12	2,07	2,75		12	1,70	2,16
	5	2,16	3,22		13	2,10	2,80		13	1,73	2,19
	6	2,30	3,41		14	2,14	2,84		14	1,76	2,22
	7	2,41	3,56		15	2,17	2,88		15	1,78	2,25
	8	2,51	3,70	10	6	1,63	2,17	14	10	1,58	2,00
	9	2,59	3,81		7	1,70	2,26		11	1,61	2,04
	10	2,67	3,92		8	1,76	2,33		12	1,64	2,07
7	5	1,92	2,74		9	1,81	2,39		13	1,67	2,10
	6	2,03	2,89		10	1,86	2,45		14	1,70	2,13
	7	2,13	3,02		11	1,90	2,50		15	1,72	2,16
	8	2,21	3,13		12	1,94	2,54	15	10	1,53	1,93
	9	2,28	3,22		13	1,98	2,59		11	1,57	1,97
	10	2,35	3,30		14	2,01	2,62		12	1,60	2,00
	11	2,40	3,38		15	2,04	2,66		13	1,62	2,03
	12	2,46	3,44	11	8	1,67	2,19		14	1,65	2,06
8	5	1,75	2,44		9	1,72	2,25		15	1,67	2,08

Задача 27. Имеются две выборки нормально распределённых случайных величин (n=m=10):

x_i	2,1	3,1	4,8	6,1	7,4	8,5	10,1	12,1	14,0	15,6
y_j	4,6	6,1	8,2	9,8	9,9	10,4	13,1	14,5	16,1	19,1

Проверить гипотезу равенства дисперсий критерием отношения размахов при доверительной вероятности $\alpha = 0{,}95$.

Решение.

Имеем $\omega_n = 15{,}6 - 2{,}1 = 13{,}5$, $\omega_m = 19{,}1 - 4{,}6 = 14{,}5$.

Далее $F^* = \dfrac{14{,}5}{13{,}5} = 1{,}074$.

Из таблицы 6 для n=m=10 и $\alpha = 0{,}95$ *находим* $F_{0,95}^*(10{,}10) = 1{,}86$.

Так как $F^* = 1{,}074 < F_\alpha^*(10{,}10) = 1{,}86$, *нулевая гипотеза равенства дисперсий принимается*

4.2.4. Критерий “стьюдентизированного” размаха

Статистика критерия: $q = \dfrac{\omega_n}{s_m}$,

где $\omega_n = x_{\max} - x_{\min}$, $s_m = \left(\dfrac{1}{m-1} \sum_{j=1}^{m} (y_j - \bar{y})^2 \right)^{\frac{1}{2}}$.

Нулевая гипотеза равенства дисперсий отклоняется, если $q > q_\alpha$, где q_α - критическое значение статистики, приведённое в таблице приложения 6, следует помнить, что здесь $f=m$-1.

Задача 28. Имеются две выборки нормально распределённых случайных величин ($n=m=10$):

x_i	2,1	3,1	4,8	6,1	7,4	8,5	10,1	12,1	14,0	15,6
y_j	4,6	6,1	8,2	9,8	9,9	10,4	13,1	14,5	16,1	19,1

Проверить гипотезу равенства дисперсий критерием “стьюдентизированного” размаха при $\alpha = 0{,}95$.

Решение.

Имеем $\omega_n = 15{,}6 - 2{,}1 = 13{,}5$, $\bar{y} = 11{,}18$, $s_m = \left(\dfrac{1}{10-1} \sum_{j=1}^{10} (y_j - \bar{y})^2 \right)^{\frac{1}{2}} = 4{,}519$.

Тогда $q = \dfrac{13{,}5}{4{,}519} = 2{,}99$.

Из таблицы приложения 6 для n=10, f=10-1=9 и $\alpha = 0{,}95$ *имеем* $q_{0,95} \approx 5{,}5$.

Так как $q = 2{,}99 < q_{0,95} = 5{,}5$, *нулевая гипотеза равенства дисперсий принимается.*

4.2.5. Сравнение нескольких дисперсий. Критерий Кохрана

Пусть $s_1^2, s_2^2, ..., s_k^2$ - взаимно независимые выборочные оценки дисперсий $\sigma_1^2, \sigma_2^2, ..., \sigma_k^2$ по выборкам объема $n_1, n_2, ..., n_k$. Проверяется нулевая гипотеза $H_0 : \sigma_1^2 = \sigma_{i+1}^2$ против альтернативы $H_1 : \sigma_1^2 \neq \sigma_{i+1}^2$ $(i = \overline{1, k-1})$.

Для случая выборок равных объемов Кохран предложил критерий, основанный на статистике $g = \dfrac{\max\limits_{1 \le i \le k} s_i^2}{\sum\limits_{i=1}^{k} s_i^2}$.

Если $g > g_\alpha(k,n)$, то нулевая гипотеза отклоняется. Значения приведены в таблице 7.

Таблица 7

Критические значения $g_\alpha(k,n)$ статистики Кохрана
для доверительной вероятности α =0,95

K	*n*					
	2	3	4	5	6	7
2	0,999	0,995	0,979	0,959	0,937	0,917
3	0,993	0,942	0,883	0,833	0,793	0,761
4	0,967	0,864	0,781	0,721	0,676	0,641
5	0,928	0,788	0,696	0,633	0,587	0,553
6	0,883	0,722	0,626	0,563	0,519	0,487
7	0,838	0,664	0,568	0,508	0,466	0,435
8	0,794	0,615	0,521	0,463	0,423	0,393
9	0,754	0,573	0,481	0,425	0,387	0,359
10	0,707	0,536	0,447	0,393	0,357	0,331
12	0,653	0,475	0,392	0,343	0,310	0,286
15	0,548	0,407	0,332	0,288	0,259	0,239
20	0,480	0,330	0,265	0,229	0,205	0,188
30	0,364	0,241	0,191	0,163	0,145	0,133
K	*n*					
	8	9	10	11	17	37
2	0,899	0,882	0,867	0,854	0,795	0,707
3	0,733	0,711	0,691	0,673	0,606	0,515
4	0,613	0,590	0,570	0,554	0,488	0,406
5	0,526	0,504	0,485	0,470	0,409	0,335
6	0,461	0,440	0,423	0,408	0,353	0,286
7	0,410	0,391	0,375	0,362	0,310	0,249
8	0,370	0,352	0,337	0,325	0,278	0,221
9	0,338	0,321	0,307	0,295	0,251	0,199
10	0,311	0,294	0,281	0,270	0,230	0,181
12	0,268	0,253	0,242	0,232	0,196	0,153
15	0,223	0,210	0,200	0,192	0,161	0,125
20	0,175	0,165	0,157	0,150	0,125	0,096
30	0,123	0,116	0,110	0,105	0,087	0,066

Задача 29. Имеются четыре выборки объема n=5 каждая

x_{i1}	3	4	5	6	7
x_{i3}	9	11	15	20	28

x_{i2}	2	8	9	11	15
x_{i4}	4	6	8	10	16

Проверить нулевую гипотезу равенства дисперсий критерием Кохрана при доверительной вероятности α =0,95.

Решение.

Имеем $\overline{x_1} = 5;\ \overline{x_2} = 9;\ \overline{x_3} = 16{,}6;\ \overline{x_4} = 8{,}8;$

$$s_1^2 = \frac{1}{4}\sum_{i=1}^{5}(x_{i1} - \overline{x_1})^2 = 2{,}5 \Rightarrow s_1 = 1{,}581;$$

$$s_2^2 = \frac{1}{4}\sum_{i=1}^{5}(x_{i2} - \overline{x_2})^2 = 22{,}5 \Rightarrow s_2 = 4{,}473;$$

$$s_3^2 = \frac{1}{4}\sum_{i=1}^{5}(x_{i3} - \overline{x_3})^2 = 58{,}3 \Rightarrow s_3 = 7{,}635;$$

$$s_4^2 = \frac{1}{4}\sum_{i=1}^{5}(x_{i4} - \overline{x_4})^2 = 21{,}2 \Rightarrow s_4 = 4{,}604;$$

$$\sum_{i=1}^{4} s_i^2 = 2{,}5 + 22{,}5 + 58{,}3 + 21{,}2 = 104{,}5 \Rightarrow \max s_i^2 = 58{,}3.$$

Тогда $g = \dfrac{58{,}3}{104{,}5} = 0{,}558$.

Из таблицы 7 для n=5, k=4 *и* $\alpha = 0{,}95$ *получаем* $g_{0,95}(4;5) = 0{,}721$.

Так как $g = 0{,}558 < g_{0,95}(4;5) = 0{,}721$, *нулевая гипотеза принимается.*

Глава 5. Методы исследования связей между случайными величинами

Математическая теория планирования активного эксперимента применительно к изучению механизма наблюдаемого процесса и построению его статистической модели базируется на методах дисперсионного, корреляционного и регрессионного анализа.

Методами дисперсионного анализа устанавливается наличие влияния заданного фактора на изучаемый процесс, отображаемый наблюдаемой статистической совокупностью выборочных данных. Методы корреляционного анализа позволяют оценить силу такой связи. Методами регрессионного анализа можно выбрать конкретную математическую модель и оценить адекватность отражения ею установленной взаимосвязи случайных величин.

5.1. Дисперсионный анализ

На практике нас часто интересует вопрос о том, в какой мере существенно влияние того или иного фактора или комбинации таких факторов на рассматриваемый признак. Так, например, при выполнении на автоматической линии некоторой операции обработки параллельно на нескольких станках, важно для правильного построения последующей обработки знать, в какой мере однотипными являются средние размере деталей, полученные на параллельно работающих станках. Нередко приходится производить измерения какой-либо физической величины параллельно на нескольких приборах несколькими операторами.

5.1.1. Однофакторный дисперсионный анализ

Предположим, что анализируется влияние фактора A, изучаемого на k уровнях ($A_1, A_2, ..., A_k$). На каждом уровне A_i проведены n наблюдений ($x_{i1}, x_{i2}, ..., x_{in}$). Следовательно, на всех k уровнях фактора A произведены kn наблюдений.

Для удобства восприятия данные представляются в виде таблицы:

Номер наблюдения	Уровни фактора A				
	A_1 A_2	$\dots$	A_i	$\dots$	A_k
1	x_{11} x_{21}	$\dots$	x_{i1}	$\dots$	x_{k1}
2	x_{12} x_{22}	$\dots$	x_{i2}	$\dots$	x_{k2}
.	. .	$\dots$	.	$\dots$	.
j	x_{1j} x_{2j}	$\dots$	x_{ij}	$\dots$	x_{kj}
.	. .	$\dots$	.	$\dots$	.
N	x_{1n} x_{2n}	$\dots$	x_{in}	$\dots$	x_{kn}
$\sum$	X_1 X_2	$\dots$	X_i	$\dots$	X_k

Дисперсионный анализ для однофакторных таблиц проводиться в следующей последовательности. Вычисляются суммы:

$$Q_1 = \sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij}^2 ; \quad Q_2 = \frac{1}{n}\sum_{i=1}^{k} X_i^2 ; \; Q_3 = \frac{1}{kn}\left(\sum_{i=1}^{k} X_i\right)^2 .$$

Далее вычисляем: $S_0^2 = \dfrac{Q_1 - Q_2}{k(n-1)}$; $\quad S_A^2 = \dfrac{Q_2 - Q_3}{k-1}$.

Сравниваем S_A^2 и S_0^2 устанавливаем наличие влияния фактора А. Если $\dfrac{S_A^2}{S_0^2} > F_\alpha[k-1; k(n-1)]$, то влияние фактора А признается значимым. В ином случае всю выборку наблюдений можно считать однородной с общей дисперсией $S^2 = \dfrac{Q_1 - Q_3}{kn-1}$.

Задача 30. Провести дисперсионный анализ данных, представленных таблицей, при доверительной вероятности α =0,95.

I	Уровни фактора A_i				
	A_1	A_2	A_3	A_4	A_5
1	3,2	2,6	2,9	3,6	3,0
2	3,1	3,1	2,6	3,4	3,4
3	3,1	2,7	3,0	3,2	3,2
4	2,8	2,9	3,1	3,3	3,5
5	3,3	2,7	3,0	3,5	2,9
6	3,0	2,8	2,8	3,3	3,1
$\sum$	18,5	16,8	17,4	20,3	19,1

Решение.

Найдем: $Q_1 = \sum_{i=1}^{5}\sum_{j=1}^{6} x_{ij}^2 = 284{,}8$; $\quad Q_2 = \dfrac{1}{6}\sum_{i=1}^{5} X_i^2 = 284{,}025$; $Q_3 = \dfrac{1}{30}\left(\sum_{i=1}^{5} X_i\right)^2 = 282{,}747$.

Тогда $S_0^2 = \dfrac{284{,}87 - 284{,}025}{5(6-1)} = 0{,}0338$; $\quad S_A^2 = \dfrac{284{,}025 - 282{,}747}{5-1} = 0{,}319$

$\Rightarrow \dfrac{S_A^2}{S_0^2} = \dfrac{0{,}319}{0{,}0338} = 9{,}45$.

Из таблицы приложения 9 для $f_1 = k-1 = 4$ *и* $f_2 = k(n-1) = 25$ *находим* $F_{0,95}(4;25) = 2{,}8$. *Так как* $\dfrac{S_A^2}{S_0^2} = 9{,}45 > F_{0,95}(4;25) = 2{,}8$, *влияние фактора А на поведение наблюдаемых случайных величин следует признать значимым.*

5.1.2. Двухфакторный дисперсионный анализ

Предположим, что анализируется влияние одновременно двух факторов А и В на уровнях $A_1, A_2, ..., A_k$ и $B_1, B_2, ..., B_m$ соответственно. Предполагается, что факторы А и В независимы. Пусть результаты эксперимента представлены таблицей:

B	A						$\sum$
	A_1	A_2	$\dots$	A_i	$\dots$	A_k	
B_1 B_2 . B_j . B_m	x_{11} x_{12} . x_{1j} . x_{1m}	x_{21} x_{22} . x_{2j} . x_{2m}	$\dots$ $\dots$. $\dots$. $\dots$	x_{i1} x_{i2} . x_{ij} . x_{im}	$\dots$ $\dots$. $\dots$. $\dots$	x_{k1} x_{k2} . x_{kj} . x_{km}	Y_1 Y_2 . Y_j . Y_m
$\sum$	X_1	X_2	$\dots$	X_i	$\dots$	X_k	

Дисперсионный анализ для двухфакторных таблиц проводится в следующей последовательности. Вычисляются суммы:

$$Q_1 = \sum_{i=1}^{k}\sum_{j=1}^{m} x_{ij}^2;\quad Q_2 = \frac{1}{m}\sum_{i=1}^{k} X_i^2;\ Q_3 = \frac{1}{k}\sum_{j=1}^{m} Y_j^2;\ Q_4 = \frac{1}{mk}\left(\sum_{i=1}^{k} X_i\right)^2 = \frac{1}{mk}\left(\sum_{j=1}^{m} Y_j\right)^2.$$

Далее находим оценки дисперсий: $S_0^2 = \dfrac{Q_1 + Q_4 - Q_2 - Q_3}{(k-1)(m-1)}$; $\quad S_A^2 = \dfrac{Q_2 - Q_4}{k-1}$; $S_B^2 = \dfrac{Q_3 - Q_4}{m-1}$.

Если $\dfrac{S_A^2}{S_0^2} > F_\alpha[f_1; f_2]$, где $f_1 = k-1$ и $f_2 = (k-1)(m-1)$, то влияние фактора А с достоверностью α признается значимым.

Если $\dfrac{S_B^2}{S_0^2} > F_\alpha[f_1; f_2]$, где $f_1 = m-1$ и $f_2 = (k-1)(m-1)$, то влияние фактора В с достоверностью α признается значимым.

Если факторы А и В зависимы, то взаимодействие факторов С=АВ также является фактором, которому соответствует своя дисперсия. Для того, чтобы выделить такое взаимодействие, необходимы параллельные наблюдения в каждой клетке таблицы, то есть при каждом сочетании факторов А и В на уровнях A_i и B_j соответственно необходимо не одно наблюдение, а серия наблюдений $x_{ij1}, x_{ij2}, ..., x_{ijn}$. Пусть x_{ij} теперь является средним из n наблюдений, то есть $x_{ij} = \dfrac{1}{n}\sum_{v=1}^{n} x_{ijv}$. Для оценки влияния взаимодействия факторов АВ вычисляем дополнительную сумму: $Q_5 = \sum_{i=1}^{k}\sum_{j=1}^{m}\sum_{v=1}^{n} x_{ijv}^2$.

Замечание. Далее анализ проводится аналогично, только в клетках таблицы вместо отдельных значений x_{ijv} используется их среднее значения x_{ij}.

Вычисляется дисперсия: $S_{AB}^2 = \dfrac{Q_5 - nQ_1}{mk(n-1)}$, и проверяется значимость взаимодействия факторов АВ критерием $\dfrac{nS_0^2}{S_{AB}^2} > F_\alpha(f_1; f_2)$, где $f_1 = (k-1)(m-1)$ и $f_2 = mk(n-1)$.

Задача 31. Провести двухфакторный дисперсионный анализ данных, представленных следующей таблицей, при доверительной вероятности α=0,95:

B	A		
	A_1	A_2	A_3
B_1	3,6 3,8 4,1	2,9 3,1 3,0	2,6 2,5 2,9
B_2	4,2 4,0 4,1	3,3 2,9 3,2	3,7 3,5 3,6
B_3	3,8 3,5 3,6	3,6 3,7 3,5	3,2 3,0 3,4
B_4	3,4 3,2 3,2	3,4 3,6 3,5	3,6 3,8 3,7

Решение.

Заменим в клетках таблицы значения их средними, получаем следующую таблицу:

B	A			$\sum$
	A_1	A_2	A_3	
B_1	3,83	3,00	2,67	9,50
B_2	4,10	3,13	3,60	10,83
B_3	3,63	3,60	3,20	10,43
B_4	3,27	3,50	3,70	10,47
$\sum$	14,83	13,23	13,17	41,23

Использую данные таблицы, вычисляем суммы:

$$Q_1 = \sum_{i=1}^{3}\sum_{j=1}^{4} x_{ij}^2 = 143{,}34;\ Q_2 = \frac{1}{4}\sum_{i=1}^{3} X_i^2 = 142{,}102675;\ Q_3 = \frac{1}{3}\sum_{j=1}^{4} X_j^2 = 141{,}98157;$$

$$Q_4 = \frac{1}{12}\left(\sum_{i=1}^{3} X_i\right)^2 = 141{,}6594;\ Q_5 = \sum_{i=1}^{3}\sum_{j=1}^{4}\sum_{v=1}^{3} x_{ijv}^2 = 430{,}79.$$

Тогда

$$S_0^2 = \frac{Q_1 + Q_4 - Q_2 - Q_3}{(k-1)(m-1)} = \frac{143{,}3745 + 141{,}6594 - 142{,}102675 - 141{,}98157}{6} = 0{,}1582;$$

$$S_A^2 = \frac{Q_2 - Q_4}{k-1} = \frac{142{,}102675 - 141{,}6594}{2} = 0{,}223675;$$

$$S_B^2 = \frac{Q_3 - Q_4}{m-1} = \frac{141{,}98157 - 141{,}6594}{3} = 0{,}10739;$$

$$S_{AB}^2 = \frac{Q_5 - nQ_1}{mk(n-1)} = \frac{430{,}79 - 3\cdot 143{,}34}{24} = 0{,}02777;$$

$$\frac{S_A^2}{S_0^2} = \frac{0{,}223675}{0{,}1582} = 1{,}41;\ \frac{S_B^2}{S_0^2} = \frac{0{,}10739}{0{,}1582} = 0{,}679;\ \frac{nS_0^2}{S_{AB}^2} = \frac{3\cdot 0{,}1582}{0{,}02777} = 17{,}09.$$

Из таблицы приложения 9 имеем: $F_{0,95}(k-1;(k-1)(m-1)) = F_{0,95}(2;6) = 5{,}1$;

$F_{0,95}(m-1;(k-1)(m-1)) = F_{0,95}(3;6) = 4{,}8$; $F_{0,95}((k-1)(m-1);km(n-1)) = F_{0,95}(6;24) = 2{,}5$.

Получаем: $\frac{S_A^2}{S_0^2} = 1{,}41 < F_{0,95}(2;6) = 5{,}1$;

$$\frac{S_B^2}{S_0^2} = 0{,}679 < F_{0,95}(3;6) = 4{,}8;$$

$$\frac{nS_0^2}{S_{AB}^2} = 17{,}09 > F_{0,95}(6;24) = 2{,}5.$$

Следовательно, влияние факторов A и B должно быть признано незначительным. Однако существенно значимым является взаимодействие факторов A и B.

5.2. Корреляционный анализ

Корреляционный анализ предполагает изучение зависимости между случайными величинами с одновременной количественной оценкой степени неслучайности их совместного изменения.

5.2.1. Оценка коэффициента корреляции

Рассмотрим нормально распределенные случайные величины x и y - $(x_1;\, y_1)$, $(x_2;\, y_2)$, …, $(x_n;\, y_n)$. Выборочной оценкой коэффициента корреляции $\rho = \dfrac{M\big((x - M(x))(y - M(y))\big)}{\sqrt{D(x)D(y)}}$ является случайная величина

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}},$$ где $\overline{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$, $\overline{y} = \dfrac{1}{n}\sum_{i=1}^{n} y_i$, n – объем выборки.

Гипотеза о значимости корреляционной связи между случайными величинами проверяется сравнением выборочного значения коэффициента корреляции r с его критическим значением r_α, являющимся α-квантилью распределения r при $\rho = 0$. Корреляция между случайными величинами признается значимой, если $|r| \geq r_\alpha$. Критические значения r_α определяются из таблицы 8.

Таблица 8

Критические значения r_α
выборочного коэффициента корреляции для $\rho = 0$

n	Доверительная вероятность α			n	Доверительная вероятность α		
	0,90	0,95	0,99		0,90	0,95	0,99
3	0,988	0,997	1,000	13	0,476	0,553	0,684
4	0,900	0,950	0,990	14	0,457	0,532	0,661
5	0,805	0,878	0,959	15	0,441	0,514	0,641
6	0,729	0,811	0,917	16	0,426	0,497	0,623
7	0,669	0,754	0,874	17	0,412	0,482	0,606
8	0,621	0,707	0,834	18	0,400	0,468	0,590
9	0,582	0,666	0,798	19	0,389	0,456	0,575
10	0,549	0,632	0,765	20	0,378	0,444	0,561
11	0,521	0,602	0,735	21	0,369	0,433	0,549
12	0,497	0,576	0,708	22	0,360	0,423	0,537

Замечание. Критические значения r_α можно вычислить с помощью аппроксимаций (где u_α и t_α - α-квантили соответственно стандартного нормального распределение и распределения Стьюдента с *f*=*n*-2 степенями свободы):

- при n>5 $$r_\alpha = \frac{e^{\frac{2}{\sqrt{n-3}}u_{\frac{1+\alpha}{2}}}-1}{e^{\frac{2}{\sqrt{n-3}}u_{\frac{1+\alpha}{2}}}+1};$$

- при n>10 $$r_\alpha = \sqrt{\frac{t^2_{\frac{1+\alpha}{2}}}{n-2+t^2_{\frac{1+\alpha}{2}}}};$$

-при n>200 $$r_\alpha = \frac{1}{\sqrt{n-1}}u_{\frac{1+\alpha}{2}}.$$

Задача 32. В результате наблюдений над случайными величинами *x* и *y* получена следующая совокупность данных (*n*=10):

x	2	4	1	7	3	11	14	15	21	4
y	7	6	4	11	2	21	31	23	40	15

Необходимо проверить гипотезу о наличии корреляции между случайными величинами *x* и *y* с достоверностью $\alpha = 0{,}95$.

Решение.

Находим $\bar{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = 8{,}2$, $\bar{y} = \frac{1}{10}\sum_{i=1}^{10} y_i = 16{,}0$,

$\sum_{i=1}^{10}(x_i-\bar{x})^2 = 405{,}6$, $\sum_{i=1}^{10}(y_i-\bar{y})^2 = 1422$, $\sum_{i=1}^{10}(x_i-\bar{x})(y_i-\bar{y}) = 723$.

Получаем оценку коэффициента корреляции: $r = \frac{723}{\sqrt{405{,}6\cdot 1422}} = 0{,}952$.

Из таблицы 8 для n=10 и $\alpha = 0{,}95$ *получаем* $r_{0,95} = 0{,}632$.

Так как $r = 0{,}952 > r_{0,95} = 0{,}632$, *наличие зависимости между величинами x и y следует признать значимой с достоверностью* $\alpha = 0{,}95$.

5.2.2. Оценка корреляционного отношения

Пусть дано n значений случайной величины $y: y_1, y_2, ..., y_k$. При $y = y_i$ наблюдаются n_i значений случайной величины x. Тогда выборочная оценка корреляционного отношения x и y вычисляется по формуле:

$$\eta_{xy}^2 = \frac{\sum_{i=1}^{k} n_i (\overline{x_i} - \overline{x})^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i} (x_{ij} - \overline{x})^2} = \frac{\sum_{i=1}^{k} n_i \overline{x_i}^2 - n\overline{x}^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{xy}^2 - n\overline{x}^2},$$

где $n = \sum_{i=1}^{k} n_i$, x_{ij} - j-е значение величины x, наблюдаемое при $y = y_i$,

$\overline{x_i} = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$, $\overline{x} = \frac{1}{k}\sum_{i=1}^{k}\overline{x_i}$.

Проверка гипотезы H_0: $\eta^2 = 0$ против альтернативы H_1: $\eta^2 \neq 0$ производиться с помощью статистики $l = \frac{\eta^2 (n-k)}{(k-1)(1-\eta^2)}$.

Если $l \geq F_\alpha(f_1; f_2)$, то нулевая гипотеза отклоняется с достоверностью α, где $F_\alpha(f_1; f_2)$ - α-квантиль F-распределения с $f_1 = k-1$ и $f_2 = n-k$ степенями свободы.

Задача 33. Проверить линейность корреляционной связи для выборки

y_i	2	4	9	13	15
x_{ij}	1, 3, 4	7, 8, 12	14, 19, 21	11, 9, 6	8, 7, 3

при доверительной вероятности $\alpha = 0{,}95$.

Решение.

Имеем k=5, n_i=3 *и* n=15. *Вычислим:* $\overline{x_1} = \frac{1+3+4}{3} = 2{,}66$; $\overline{x_2} = \frac{7+8+12}{3} = 9$;

$\overline{x_3} = \frac{14+19+21}{3} = 18$; $\overline{x_4} = \frac{11+9+6}{3} = 8{,}67$; $\overline{x_5} = \frac{8+7+3}{3} = 6$; $\sum_{i=1}^{5}\sum_{j=1}^{3} x_{ij}^2 = 1641$;

$\overline{x} = \frac{2{,}66+9+18+8{,}67+6}{5} = 8{,}864$; $\sum_{i=1}^{5} n_i \overline{x_i}^2 = 3\left(2{,}66^2 + 9^2 + 18^2 + 8{,}67^2 + 6^2\right) = 1569{,}2136$.

Получаем: $\eta_{xy}^2 = \frac{1569{,}2136 - 15\cdot 8{,}864^2}{1641 - 15\cdot 8{,}864^2} = 0{,}8448$.

Из таблицы приложения 9 находим

$F_{0,95}(f_1; f_2) = F_{0,95}(5-1;15-5) = F_{0,95}(4;10) = 3{,}5$

Вычислим $l = \frac{0{,}845\cdot 10}{4\cdot(1-0{,}845)} = 13{,}629$.

Полученная величина больше критического значения $F_{0,95}(4;10)=3{,}5$, *следовательно, необходимо признать наличие существенной нелинейной связи между x и y.*

5.2.3. Частная и множественная корреляция

Рассмотрим случай трех переменных – x, y, z.

Зависимость между двумя переменными x и y при фиксированной третьей переменной – z оцениваются с помощью частного коэффициента корреляции $\rho_{xy,z}$. Аналогично определяются $\rho_{xz,y}$ и $\rho_{zy,x}$.

Выборочные частные коэффициенты корреляции определяется с помощью соотношений:

$$r_{xy,z}=\frac{r_{xy}-r_{xz}r_{yz}}{\sqrt{(1-r_{xz}^2)(1-r_{yz}^2)}};\ r_{xz,y}=\frac{r_{xz}-r_{xy}r_{zy}}{\sqrt{(1-r_{xy}^2)(1-r_{zy}^2)}};\ r_{zy,x}=\frac{r_{zy}-r_{zx}r_{yx}}{\sqrt{(1-r_{zx}^2)(1-r_{yx}^2)}};$$

$$r_{yz,x}=r_{zy,x};\ r_{xy,z}=r_{yx,z};\ r_{xz,y}=r_{zx,y}.$$

Гипотеза H_0: $\rho_{xy,z}=0$ для коэффициента корреляции $\rho_{xy,z}$ проверяется с помощью статистики: $t=\dfrac{\sqrt{n-k}\cdot r_{xy,z}}{\sqrt{1-r_{xy,z}^2}}$, где k – число переменных, t имеет распределение Стьюдента при $f=n-k$ степенях свободы. При $|t|>t_{\frac{1+\alpha}{2}}(n-k)$ нулевая гипотеза H_0 отклоняется с вероятностью α.

Множественная корреляция исследуется в том случае, когда необходимо установить существенность взаимосвязи одной переменной с совокупностью остальных. Выборочные множественные коэффициенты корреляции обозначаются $r_{x,yz}$, $r_{y,xz}$, $r_{z,yx}$ и вычисляются с помощью соотношений:

$$r_{x,yz}^2=\frac{r_{xy}^2+r_{xz}^2-2r_{xy}r_{xz}r_{yz}}{1-r_{yz}^2};\ r_{y,xz}^2=\frac{r_{yx}^2+r_{yz}^2-2r_{yx}r_{yz}r_{xz}}{1-r_{xz}^2};\ r_{z,xy}^2=\frac{r_{zx}^2+r_{zy}^2-2r_{zx}r_{yz}r_{xy}}{1-r_{xy}^2}.$$

Замечание. Между частными, множественными и обыкновенными коэффициентами корреляции имеют место следующие соотношения:

$$r_{x,yz}^2=1-\left(1-r_{xz}^2\right)\left(1-r_{xy,z}^2\right)=1-\left(1-r_{xy}^2\right)\left(1-r_{xz,y}^2\right);$$

$$r_{y,xz}^2=1-\left(1-r_{yz}^2\right)\left(1-r_{yx,z}^2\right)=1-\left(1-r_{yx}^2\right)\left(1-r_{zy,x}^2\right);$$

$$r_{z,xy}^2=1-\left(1-r_{zy}^2\right)\left(1-r_{zx,y}^2\right)=1-\left(1-r_{zx}^2\right)\left(1-r_{zy,x}^2\right).$$

Для проверки гипотезы H_0: $\rho_{x,yz}=0$ используется статистика:

$F = \frac{r_{x,yz}^2}{1-r_{x,yz}^2}\frac{n-k}{k-1}$, где F-распределение с $f_1 = k-1$ и $f_2 = n-k$ степенями свободы (k – число переменных).

Если $F > F_\alpha(f_1; f_2)$, то соответствующая корреляция признается значимой. Критическое значение коэффициента множественной корреляции равно $r_{x,yz}(\alpha) = \sqrt{\frac{(k-1)F_\alpha(f_1; f_2)}{n-k+(k-1)F_\alpha(f_1; f_2)}}$. Корреляция признается значимой при $r_{x,yz} \geq r_{x,yz}(\alpha)$, где критические значения $r_{x,yz}(\alpha)$ приведены в таблице 9.

Таблица 9

Критические значения $r_{1,23...k}$ коэффициента множественной корреляции (k – число переменных, n – объем выборки)

n-k	Доверительная вероятность α							
	0,95				0,99			
	K				*k*			
	3	4	5	6	3	4	5	6
1	0,999	0,999	0,999	1,000	1,000	1,000	1,000	1,000
2	0,975	0,983	0,987	0,990	0,995	0,997	0,997	0,998
3	0,930	0,950	0,961	0,968	0,977	0,983	0,987	0,990
4	0,881	0,912	0,930	0,942	0,949	0,962	0,970	0,975
5	0,836	0,874	0,898	0,914	0,917	0,937	0,949	0,957
6	0,795	0,839	0,867	0,886	0,886	0,911	0,927	0,938
7	0,758	0,807	0,838	0,860	0,855	0,885	0,904	0,918
8	0,726	0,777	0,811	0,835	0,827	0,860	0,882	0,898
9	0,697	0,750	0,786	0,812	0,800	0,837	0,861	0,878
10	0,671	0,726	0,763	0,790	0,776	0,814	0,840	0,859
11	0,648	0,703	0,741	0,770	0,753	0,793	0,821	0,841
12	0,627	0,683	0,722	0,751	0,732	0,773	0,802	0,824
13	0,608	0,664	0,703	0,733	0,712	0,755	0,785	0,807
14	0,590	0,646	0,686	0,717	0,694	0,737	0,768	0,791
15	0,574	0,630	0,670	0,701	0,677	0,721	0,752	0,776
16	0,559	0,615	0,655	0,687	0,662	0,706	0,738	0,762
17	0,545	0,601	0,641	0,673	0,647	0,691	0,724	0,749
18	0,532	0,587	0,628	0,660	0,633	0,678	0,710	0,736
19	0,520	0,575	0,615	0,647	0,620	0,665	0,697	0,723
20	0,509	0,563	0,604	0,636	0,607	0,652	0,685	0,712
22	0,488	0,542	0,582	0,614	0,585	0,630	0,663	0,690
24	0,470	0,523	0,562	0,594	0,565	0,609	0,643	0,669
26	0,454	0,506	0,545	0,576	0,546	0,590	0,624	0,651
28	0,439	0,490	0,529	0,560	0,529	0,573	0,607	0,633
30	0,425	0,476	0,514	0,545	0,514	0,557	0,591	0,618
40	0,373	0,419	0,455	0,484	0,454	0,494	0,526	0,552
60	0,308	0,348	0,380	0,406	0,377	0,414	0,442	0,467

Задача 34. Вычислить коэффициенты частной и множественной корреляций и проверить их значимость при доверительной вероятности $\alpha = 0{,}95$ для данных:

x_i	1	3	4	7	12	4	19	21	1	3
y_i	12	42	58	71	68	50	49	85	18	26
z_i	41	12	7	3	14	27	38	13	64	75

Решение.

Найдем парные коэффициенты корреляции и вычислим коэффициент r_{xy}*:*

$$\sum_{i=1}^{10}\left(x_i-\bar{x}\right)^2 = 484{,}5;\ \sum_{i=1}^{10}\left(y_i-\bar{y}\right)^2 = 6882{,}1;\ \sum_{i=1}^{10}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right) = 1091;$$

$$\bar{x} = \frac{1}{10}\sum_{i=1}^{10}x_i = 7{,}5;\ \bar{y} = \frac{1}{10}\sum_{i=1}^{10}y_i = 47{,}9;\ r_{xy} = \frac{1091}{\sqrt{484{,}5\cdot 6882{,}1}} = 0{,}597.$$

Вычислим коэффициент r_{xz}*:*

$$\sum_{i=1}^{10}\left(z_i-\bar{z}\right)^2 = 5498{,}4;\ \sum_{i=1}^{10}\left(x_i-\bar{x}\right)\left(z_i-\bar{z}\right) = -519;\ \bar{z} = \frac{1}{10}\sum_{i=1}^{10}z_i = 29{,}4;$$

$$r_{xz} = -\frac{519}{\sqrt{484{,}5\cdot 5498{,}4}} = -0{,}318.$$

Вычислим коэффициент r_{yz}*:*

$$\sum_{i=1}^{10}\left(y_i-\bar{y}\right)\left(z_i-\bar{z}\right) = -3172{,}66;\ r_{yz} = -\frac{3172{,}66}{\sqrt{6862{,}1\cdot 5498{,}4}} = -0{,}516.$$

Вычислим частные коэффициенты корреляции:

$$r_{xy,z} = \frac{r_{xy}-r_{xz}\cdot r_{yz}}{\sqrt{\left(1-r_{xz}^2\right)\left(1-r_{yz}^2\right)}} = \frac{0{,}597-(-0{,}318)\cdot(-0{,}516)}{\sqrt{(1-0{,}318^2)(1-0{,}516^2)}} = 0{,}533;$$

$$r_{xz,y} = \frac{r_{xz}-r_{xy}\cdot r_{yz}}{\sqrt{\left(1-r_{xy}^2\right)\left(1-r_{yz}^2\right)}} = \frac{-0{,}318-0{,}597\cdot(-0{,}516)}{\sqrt{(1-0{,}597^2)(1-0{,}516^2)}} = -0{,}014;$$

$$r_{yz,x} = \frac{r_{yz}-r_{yx}\cdot r_{zx}}{\sqrt{\left(1-r_{yx}^2\right)\left(1-r_{zx}^2\right)}} = \frac{-0{,}516-(-0{,}318)\cdot 0{,}597}{\sqrt{(1-0{,}318^2)(1-0{,}597^2)}} = -0{,}429.$$

Вычислим множественные коэффициенты корреляции:

$$r_{x,yz}^2 = \frac{0{,}597^2+0{,}318^2-2\cdot 0{,}597\cdot(-0{,}318)\cdot(-0{,}516)}{1-0{,}516^2} = 0{,}356;$$

$$r_{y,xz}^2 = \frac{0{,}597^2+0{,}516^2-2\cdot 0{,}597\cdot(-0{,}318)\cdot(-0{,}516)}{1-0{,}318^2} = 0{,}475;$$

$$r_{z,xy}^2 = \frac{0{,}516^2+0{,}318^2-2\cdot 0{,}597\cdot(-0{,}318)\cdot(-0{,}516)}{1-0{,}597^2} = 0{,}266.$$

Вычислим t-статистики для проверки значимости частных коэффициентов корреляции:

- для проверки $r_{xy,z}$ *:* $t_{xy,z} = \dfrac{\sqrt{10-3}\cdot 0{,}533}{\sqrt{1-0{,}533^2}} = 1{,}667$ *;*

- для проверки $r_{xz,y}$ *:* $t_{xz,y} = \dfrac{\sqrt{10-3}\cdot(-0{,}014)}{\sqrt{1-0{,}014^2}} = -0{,}037$ *;*

- для проверки $r_{zy,x}$ *:* $t_{zy,x} = \dfrac{\sqrt{10-3}\cdot(-0{,}429)}{\sqrt{1-0{,}429^2}} = -1{,}256$.

Для $\alpha = 0{,}95$ *и* $f = n-k = 7$ *из таблицы для t-распределения имеем* $t_{\frac{1+0{,}95}{2}} = t_{0{,}975}(7) = 2{,}37$. *Получаем* $|t_{xy,z}| < 2{,}37$, $|t_{xz,y}| < 2{,}37$, $|t_{zy,x}| < 2{,}37$.

Для коэффициентов множественной корреляции находим критическое значение из таблицы при k=3, n-k=7 и $\alpha = 0{,}95$. *Имеем* $r_{1,23}(0{,}95) = 0{,}758$.

Так как ни один множественный коэффициент корреляции не превышает критическое значение 0,758, то и наличие множественной корреляции отклоняется с достоверностью 0,95.

5.3. Регрессионный анализ

Любая функция распределения определяется своими параметрами. Изменение функции распределения случайной величины y и x можно задать зависимостями: $\mu_{1y} = f(x)$, $\mu_{2y} = \sigma^2 = f_2(x)$, $\mu_{3y} = f_3(x)$, $\mu_{4y} = f_4(x)$, называемые регрессионной, скедастической, клитической, синагической. Зависимость средних значений называется регрессией y по x.

В основе регрессионного анализа лежит принцип наименьших квадратов, в соответствии с которым в качестве уравнения регрессии $y = f(x)$ выбирается функция, достигается минимум суммы квадратов разностей $s = \sum_{i=1}^{n}[y_i - f(x_i)]^2$. При этом вид функции определяется заранее, а методом наименьших квадратов определяются коэффициенты, минимизирующие s. Количественной мерой рассеяния значений y_i вокруг регрессии $f(x)$ является дисперсия $D = \dfrac{1}{n-k}\sum_{i=1}^{n}(y_i - f(x_i))^2$, где k – число коэффициентов, входящих в аналитическое выражение регрессии.

В зависимости от вида уравнения регрессии различают линейную и нелинейную регрессии.

5.3.1. Оценка коэффициентов регрессии линейного регрессионного анализа

Линейный регрессионный анализ исходит из наличия зависимости $y = a + bx$, где a и b - неизвестные коэффициенты регрессии, вычисляемые по формулам: $b = \dfrac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$, $a = \dfrac{\sum_{i=1}^{n} y_i - b\sum_{i=1}^{n} x_i}{n}$.

Замечание. Для проверки правильности вычислений можно использовать соотношения: $\overline{y} = a + b\overline{x}$, где $\overline{y} = \dfrac{1}{n}\sum_{i=1}^{n} y_i$, $\overline{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$.

Задача 35. В результате наблюдений за зависимостью $y = f(x)$ получены следующие данные:

y_i	2	3	7	10	11	13	18	21	25	31
x_i	8	11	14	18	24	26	31	32	34	41

Необходимо найти оценку коэффициентов регрессии y по x методом наименьших квадратов.

Решение.

Вычислим: $\sum_{i=1}^{n} x_i = 239$, $\sum_{i=1}^{n} y_i = 141$, $\sum_{i=1}^{n} x_i^2 = 6779$, $\sum_{i=1}^{n} x_i y_i = 4280$.

Получаем: $b = \dfrac{10 \cdot 4280 - 239 \cdot 141}{10 \cdot 6779 - 239^2} = 0{,}853032$, $a = \dfrac{141 - 0{,}853032 \cdot 239}{10} = -6{,}28747$.

Таким образом, имеем уравнение регрессии y по x: $y = -6{,}28747 + 0{,}853032x$.

Проверка: $\overline{x} = 23{,}9$, *тогда* $\overline{y} = -6{,}28747 + 0{,}853032 \cdot 23{,}9 = 14{,}1$. *Следовательно, коэффициенты вычислены верно.*

5.3.2. Статистический анализ коэффициентов регрессии

Статистические выводы относительно коэффициента β регрессии $y = \alpha + \beta x$ могут быть получены с помощью статистики $t_\beta = \dfrac{b - \beta}{S_\beta}$, где $S_\beta = \dfrac{S}{S_x\sqrt{n-1}}$, $S^2 = \dfrac{1}{n-2}\sum_{i=1}^{n}(y_i - a - bx_i)^2$, $S_x^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$, $\overline{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$, β - истинное значение коэффициента регрессии, b – выборочная оценка коэффициента регрессии.

Статистика t_β при справедливости нулевой гипотезы H_0: $\beta = b$ имеет распределение Стьюдента с $f = n-2$ степенями свободы.

Таким образом, с помощью квантилей распределения Стьюдента можно проверить гипотезу равенства β заданному значению, гипотезу о значимости коэффициента регрессии, построить доверительный интервал для коэффициента β:

1. Гипотеза о равенстве коэффициента β заданному значению β_0 принимается, если $|\beta - b| < t_{\frac{1+\alpha}{2}} S_\beta$.
2. Значение коэффициента β регрессии является значимым с достоверность α, если $|b| > t_{\frac{1+\alpha}{2}} S_\beta$.
3. Доверительный интервал для β имеет вид: $b - S_\beta t_{\frac{1+\alpha}{2}} \le \beta \le b + S_\beta t_{\frac{1+\alpha}{2}}$.

Выводы относительно коэффициента α могут быть получены с помощью статистики: $t_\alpha = \dfrac{a-\alpha}{S_\alpha}$, где $S_\alpha = S\sqrt{\dfrac{1}{n} + \dfrac{\bar{x}^2}{(n-1)S_x^2}}$, a, α - соответственно выборочная оценка и истинное значение коэффициента α. При H_0: $a = \alpha$ статистика t_α имеет распределение Стьюдента с $f = n-2$ степенями свободы.

Замечание. Проверка гипотез о значениях коэффициента α и построении доверительных интервалов выполняется по аналогии с коэффициентом β.

Задача 36. Для совокупности данных

x_i	1,2	2,4	2,8	4,2	5,9	6,8	8,1	9,2	10,1	11,0
y_i	7	12	17	24	29	38	46	45	54	68

Найти оценки коэффициентов α и β регрессии $y = \alpha + \beta x$ и провести их статистический анализ при доверительной вероятности $\alpha = 0{,}95$.

Решение.

Вычислим: $\sum_{i=1}^{10} x_i = 61{,}7$, $\left(\sum_{i=1}^{10} x_i\right)^2 = 3806{,}89$, $\sum_{i=1}^{10} x_i^2 = 486{,}99$, $\sum_{i=1}^{10} y_i = 340$, $\sum_{i=1}^{10} x_i y_i = 2695{,}1$.

Получаем: $b = \dfrac{10 \cdot 2695{,}1 - 61{,}7 \cdot 340}{10 \cdot 486{,}99 - 3806{,}89} = 5{,}6189$ *и* $a = \dfrac{340 - 5{,}6189 \cdot 61{,}7}{10} = -0{,}668$.

Таким образом, имеем $\hat{y}_i = a + bx_i$*:*

$\hat{y}_i \to$ 6,075; 12,818; 15,065; 22,932; 32,484; 37,541; 44,846; 51,027; 56,084; 61,141.

Проверим значимость полученных коэффициентов.

Вычислим: $\overline{x} = 6{,}17$, $S_x^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 = 11{,}8112$, $S_x = 3{,}3467$,

$$S^2 = \frac{1}{8}\sum_{i=1}^{10}\left(y_i - \hat{y}_i\right)^2 = 13{,}4755,\quad S_\beta = \frac{S}{S_x\sqrt{n-1}} = \frac{\sqrt{13{,}4755}}{3{,}3467\cdot 3} = 0{,}3656,$$

$$S_\alpha = S\sqrt{\frac{1}{n} + \frac{\overline{x}^2}{(n-1)S_x^2}} = 3{,}671\sqrt{\frac{1}{10} + \frac{6{,}17^2}{9\cdot 11{,}8112}} = 2{,}486.$$

Для уровня достоверности $\alpha = 0{,}95$ *имеем* $t_{\frac{1+0{,}95}{2}}(n-2) = t_{0{,}975}(8) = 2{,}306$

Проверим значимость коэффициента β:

1. $|b| = 5{,}1689 > t_{0{,}975}(8)\cdot S_\beta = 2{,}306\cdot 0{,}3656 = 0{,}843$, *следовательно, с достоверностью 0,95 получаем вывод о значимости коэффициента регрессии.*
2. *Проверим гипотезу H_0:* $\beta = 5$: $|5{,}619 - 5| = 0{,}619 < t_{0{,}975}(8)\cdot S_\beta = 0{,}843$, *то есть гипотеза о равенстве* $\beta = 5$ *не отклоняется.*
3. *Вычислим доверительный интервал для* β:

   $5{,}9 - 2{,}306\cdot 0{,}3656 = 4{,}776 \le \beta \le 6{,}462 = 5{,}619 + 2{,}306\cdot 0{,}3656$.

Проверим значимость коэффициента α:

1. *Проверим гипотезу H_0:* $\alpha = 0$: $|a| = 0{,}668 < t_{0{,}975}(8)\cdot S_\alpha = 5{,}73$. *Следовательно, коэффициент* α *с вероятностью 0,95 не отличается от нуля, то есть его значение может быть приравнено к нулю.*
2. *Вычислим доверительный интервал для* α:

   $-0{,}668 - 2{,}306\cdot 2{,}485 = -6{,}398 \le \alpha \le 5{,}602 = -0{,}668 + 2{,}306\cdot 2{,}485$.

Таким образом, уравнение регрессии y по x адекватно отображается уравнением $y = 5{,}619\cdot x$.

5.3.3. Оценка адекватности регрессии

Количественной мерой адекватности является отношение дисперсии S^2, определяемой рассеянием значений y_i вокруг линии регрессии, к дисперсии S_y^2 естественного рассеяния значений y_i вокруг своих средних $\hat{y}_i$.

Если $\dfrac{S^2}{S_y^2} > F_\alpha$, где F_α – α-квантиль распределения Фишера с $f_1 = n-2$ и $f_2 = m-1$ степенями свободы, то ошибка в определении регрессии с доверительной вероятностью α признается статистически значимой (m - объем выборки, по которой выполнена оценка дисперсии S_y^2).

Замечание.

1. Если дисперсия S_y^2 известна заранее, то $f_2 \to \infty$. Если дисперсия S_y^2 определяется по дублируемым значениям y_i, то ее оценкой является средневзвешенная дисперсия.

2. Дисперсия S_y^2 вычисляется по формуле: $S_y^2 = \frac{1}{n}\sum_{i=1}^{n} S_i^2$, где $S_i^2 = \frac{1}{m-1}\sum_{j=1}^{m}\left(y_{ij} - \hat{y}_i\right)^2$, $\hat{y}_i = \frac{1}{m}\sum_{j=1}^{m} y_j$.

Задача 37. Проверить адекватность регрессии для данных:

x_i	1,2	2,4	2,8	4,2	5,9	6,8	8,1	9,2	10,1	11,0
y_i	7	12	17	24	29	38	46	45	54	68

при доверительной вероятности $\alpha = 0{,}95$, если для оценки S_y^2 предварительно проводилась серия наблюдений над случайной величиной y при неизменной величине x (m=10): y_{ij} : 12 14 11 10 8 6 7 15 13.

Решение.

Имеем: $\sum_{i=1}^{10} x_i = 61{,}7$, $\left(\sum_{i=1}^{10} x_i\right)^2 = 3806{,}89$, $\sum_{i=1}^{10} x_i^2 = 486{,}99$, $\sum_{i=1}^{10} y_i = 340$, $\sum_{i=1}^{10} x_i y_i = 2695{,}1$,

$b = \frac{10 \cdot 2695{,}1 - 61{,}7 \cdot 340}{10 \cdot 486{,}99 - 3806{,}89} = 5{,}6189$, $a = \frac{340 - 5{,}6189 \cdot 61{,}7}{10} = -0{,}668$.

Тогда $\hat{y}_i = a + bx_i$*:*

$\hat{y}_i \to$ 6,075; 12,818; 15,065; 22,932; 32,484; 37,541; 44,846; 51,027; 56,084; 61,141.

Получаем: $S^2 = \frac{1}{8}\sum_{i=1}^{10}\left(y_i - \hat{y}_i\right)^2 = 13{,}4755$.

По отдельной серии наблюдений находим оценку: $S_y^2 = \frac{1}{m-1}\sum_{j=1}^{m}\left(y_{ij} - \bar{y}_i\right)^2 = 11{,}733$, $S_y = 3{,}425$.

Вычислим $F = \frac{S^2}{S_y^2} = \frac{13{,}4755}{11{,}733} = 1{,}148$. *Из таблицы F-распределения:* $F_{0,95}(f_1 = n-2 = 8;\ f_2 = m-2 = 8) = 3{,}438$.

Так как $F = 1{,}148 < F_{0,95}(8;8) = 3{,}438$, *с вероятностью* $\alpha = 0{,}95$ *следует сделать вывод о статистической неразличимости сравниваемых дисперсий, а следовательно, об адекватности уравнения регрессии.*

5.3.4. Сравнение линейных регрессий

Сравнение двух регрессионных моделей включает в себя проверку нулевой гипотезы H_0: $y_1 = y_2$ о статистической неразличимости линейных регрессий $y_1 = a_1 + b_1 x$ и $y_2 = a_2 + b_2 x$. Для проверки H_0 необходимо последовательно проверить три гипотезы: о неразличимости коэффициентов a_1 и a_2 ($H_0' : a_1 = a_2$); b_1 и b_2 ($H_0'' : b_1 = b_2$), и о равенстве остаточных дисперсий, характеризующих разброс значений y_i вокруг линии регрессии ($H_0''' : S_1^2 = S_2^2$).

В первую очередь проверяется гипотеза $H_0''' : S_1^2 = S_2^2$.

Для проверки гипотезы подсчитываются соответствующие дисперсии:

$S_1^2 = \frac{1}{n_1 - 2}\sum_{i=1}^{n_1}(y_i - a_1 - b_1 x_i)^2$ и $S_2^2 = \frac{1}{n_2 - 2}\sum_{i=1}^{n_2}(y_i - a_2 - b_2 x_i)^2$, где n_1, n_2 - объемы выборок, по которым найдены регрессии y_1 и y_2 соответственно; x_{1i} $(i = \overline{1, n_1})$ и x_{2i} $(i = \overline{1, n_2})$ - значения переменной x в двух выборках. Дисперсии S_1^2 и S_2^2 сравниваются между собой с помощью критерия Фишера. Если $\frac{S_1^2}{S_2^2} > F_\alpha$, где $F_\alpha - \alpha$-квантиль распределения Фишера с $f_1 = n_1 - 2$ и $f_2 = n_2 - 2$ степенями свободы, то гипотеза $H_0''' : S_1^2 = S_2^2$ отклоняется, и регрессии признаются статистически различными.

Если гипотеза $H_0''' : S_1^2 = S_2^2$ не отклоняется, можно приступить к проверке гипотез $H_0' : a_1 = a_2$ и $H_0'' : b_1 = b_2$. Для проверки равенства угловых коэффициентов b_1 и b_2 используется статистика:

$$t_b = \frac{b_1 - b_2}{S^* \sqrt{\frac{1}{(n_1 - 1)S_{\overline{x_1}}^2} + \frac{1}{(n_2 - 1)S_{\overline{x_2}}^2}}},$$

где $S_{\overline{x_1}}^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}\left(x_{1i} - \overline{x_1}\right)^2$, $S_{\overline{x_2}}^2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}\left(x_{2i} - \overline{x_2}\right)^2$, $\overline{x_1} = \frac{1}{n_1}\sum_{i=1}^{n_1} x_{1i}$, $\overline{x_2} = \frac{1}{n_2}\sum_{i=1}^{n_2} x_{2i}$, $S^* = \left(\frac{(n_1 - 2)S_{x_1}^2 + (n_2 - 2)S_{x_2}^2}{n_1 + n_2 - 4}\right)^{\frac{1}{2}}$.

При справедливости нулевой гипотезы $H_0'' : b_1 = b_2$ статистика t_b имеет распределение Стьюдента с $f = n_1 + n_2 - 4$ степенями свободы. Если $|t_b| \geq t_{\frac{1+\alpha}{2}}$, то с достоверной вероятностью α гипотеза $H_0'':b_1=b_2$ отклоняется и сравниваемые угловые коэффициенты b_1 и b_2 признаются различными.

Если гипотеза $H_0'':b_1=b_2$ не отклоняется, то проводится проверка гипотезы $H_0':a_1=a_2$. Для проверки этой гипотезы используется статистика:

$$t_a=\frac{\bar{b}-\tilde{b}}{\tilde{S}},\quad \text{где}\quad \bar{b}=\frac{(n_1-1)S_{x_1}^2 b_1+(n_2-1)S_{x_2}^2 b_2}{(n_1-1)S_{x_1}^1+(n_2-1)S_{x_2}^2},\quad \tilde{b}=\frac{\overline{y_1}-\overline{y_2}}{\overline{x_1}-\overline{x_2}},\quad \overline{y_1}=\frac{1}{n_1}\sum_{i=1}^{n_1}y_{1i},$$

$$\overline{y_2}=\frac{1}{n_2}\sum_{i=1}^{n_2}y_{2i},\quad \tilde{S}=S^*\left(\frac{1}{(n_1-1)S_{x_1}^2+(n_2-1)S_{x_2}^2}+\frac{1}{(\overline{x_1}-\overline{x_2})}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\right),\quad S^*,\ S_{x_1}^2,\ S_{x_2}^2$$

определяются при проверки гипотезы $H_0'':b_1=b_2$.

При справедливости $H_0':a_1=a_2$ статистика t_a имеет распределение Стьюдента с $f=n_1+n_2-4$ степенями свободы. Если $|t_a|>t_{\frac{1+\alpha}{2}}$, то с достоверной вероятностью α гипотеза $H_0':a_1=a_2$.

Если все гипотезы $H_0':a_1=a_2$, $H_0'':b_1=b_2$, $H_0''':S_1^2=S_2^2$ не отклоняются, то сравниваемые регрессии y_1 и y_2 признаются неразличимыми.

Задача 38. В результате двух независимых экспериментов получены следующие результаты ($n_1=10, n_2=6$):

x_{1i}	2	4	6	9	11	16	17	20	25	31
y_{1i}	9	19	22	41	49	61	69	83	98	128

x_{2i}	12	16	21	23	28	31
y_{2i}	54	68	87	93	112	130

Проверить гипотезу о статистической неразличимости регрессионных моделей, полученных по обеим выборкам при вероятности $\alpha=0{,}95$.

Решение.

Вычислим для первой выборки: $\sum_{i=1}^{10}x_{1i}=141$, $\left(\sum_{i=1}^{10}x_{1i}\right)^2=19881$, $\sum_{i=1}^{10}x_{1i}^2=2789$, $\sum_{i=1}^{10}y_{1i}=589$, $\overline{x_1}=14{,}1$, $\overline{y_1}=57{,}9$, $\sum_{i=1}^{10}x_{1i}\cdot y_{1i}=11361$, $b_1=\frac{10\cdot 11361-141\cdot 579}{10\cdot 2789-19881}=3{,}992$,

$a_1=\frac{179-3{,}992\cdot 14{,}1}{10}=1{,}613$, $S_{x_1}^2=\frac{1}{n_1-1}\sum_{i=1}^{10}\left(x_{1i}-\overline{x_1}\right)^2=88{,}989$, $S_{x_1}=9{,}439$.

Вычислим для второй выборки: $\sum_{i=1}^{6}x_{2i}=131$, $\left(\sum_{i=1}^{6}x_{2i}\right)^2=17161$, $\sum_{i=1}^{6}x_{2i}^2=3115$, $\sum_{i=1}^{6}y_{2i}=544$, $\overline{x_2}=21{,}83$, $\overline{y_2}=90{,}667$, $\sum_{i=1}^{6}x_{2i}\cdot y_{2i}=12868$, $b_2=\frac{6\cdot 12868-131\cdot 544}{6\cdot 3115-17161}=3{,}887$,

$a_2=\frac{544-3{,}887\cdot 21{,}83}{6}=5{,}8$, $S_{x_2}^2=\frac{1}{n_2-1}\sum_{i=1}^{6}\left(x_{2i}-\overline{x_2}\right)^2=50{,}967$, $S_{x_2}=7{,}39$.

Вычислим дисперсию рассеяния значений y_i *вокруг линии регрессии:*

$S_1^2 = \frac{1}{8}\sum_{i=1}^{10}\left(y_{1i} - 1{,}613 - 3{,}992 \cdot x_{1i}\right)^2 = 10{,}056$, $\quad S_2^2 = \frac{1}{4}\sum_{i=1}^{6}\left(y_{2i} - 5{,}8 - 3{,}887 \cdot x_{2i}\right)^2 = 7{,}0271$,

тогда $S_1 = 3{,}171$, $S_2 = 2{,}651$.

Проверим гипотезу $H_0''' : S_1^2 = S_2^2$. *Имеем* $\frac{S_1^2}{S_2^2} = \frac{10{,}056}{7{,}0271} = 1{,}431 < F_{0{,}95}(8;4) = 3{,}838$, *следовательно, гипотеза* $H_0''' : S_1^2 = S_2^2$ *принимается и дисперсии* S_1^2 *и* S_2^2 *с вероятностью 0,95 признаются статистически неразличимыми.*

Проверим гипотезу $H_0'' : b_1 = b_2$.

Для этого вычислим: $S^* = \left(\frac{8 \cdot 10{,}056 + 4 \cdot 7{,}0271}{12}\right)^{\frac{1}{2}} = 3{,}008$.

Тогда $t_b = \frac{3{,}992 - 3{,}887}{3{,}008\sqrt{\frac{1}{9 \cdot 88{,}989} + \frac{1}{5 \cdot 50{,}967}}} = 0{,}0078$. *Из таблицы* $t_{0{,}95}(10 + 6 - 4) = t_{0{,}95}(12) = 2{,}17$. *Так как* $|t_b| = 0{,}0078 < t_{0{,}95}(12) = 2{,}17$, *гипотеза* $H_0'' : b_1 = b_2$ *не отклоняется, и угловые коэффициенты* b_1 *и* b_2 *признаются неразличимыми с достоверностью 0,95.*

Проверим гипотезу $H_0' : a_1 = a_2$.

Для этого вычислим: $\bar{b} = \frac{9 \cdot 88{,}989 \cdot 3{,}992 + 5 \cdot 50{,}967 \cdot 3{,}887}{9 \cdot 88{,}989 + 5 \cdot 50{,}967} = 3{,}967$,

$$\tilde{b} = \frac{\overline{y_1} - \overline{y_2}}{\overline{x_1} - \overline{x_2}} = \frac{57{,}9 - 50{,}967}{14{,}1 - 21{,}83} = 4{,}239,$$

$$\tilde{S} = 3{,}008\left(\frac{1}{9 \cdot 88{,}989 + 5 \cdot 50{,}967} + \frac{1}{(14{,}1 - 21{,}83)^2}\left(\frac{1}{10} + \frac{1}{6}\right)\right)^{\frac{1}{2}} = 0{,}2212,$$

$$t_a = \frac{3{,}967 - 4{,}239}{0{,}2212} = -1{,}23.$$

Так как $|t_a| = 1{,}23 < t_{0{,}95}(12) = 2{,}170$, *то гипотеза* $H_0' : a_1 = a_2$ *не отклоняется. Следовательно, сравниваемые эмпирические уравнения регрессии с достоверностью 0,95 статистически эквивалентны.*

5.3.5. Линеаризация нелинейной модели

Смысл метода заключается в том, что нелинейное уравнение преобразуется в линейное с помощью перехода от исследуемых переменных к новым переменным. Обработка результатов наблюдений, вычисление регрессии и ее статистический анализ для линейно преобразованного уравнения выполняются методами линейного регрессионного анализа.

Замечание. Подбор необходимого линеаризующего преобразования выполняется на основании графика.

Задача 39. В результате исследования зависимости между случайными величинами x и y получены следующие результаты (n=10):

x_i	1	2	3	4	5	6	7	8	9	10
y_i	81	19	9	8,5	7	6,6	5	5,8	5,6	5,4

Необходимо найти регрессию $y = f(x)$ и провести ее статистический анализ.

Решение.

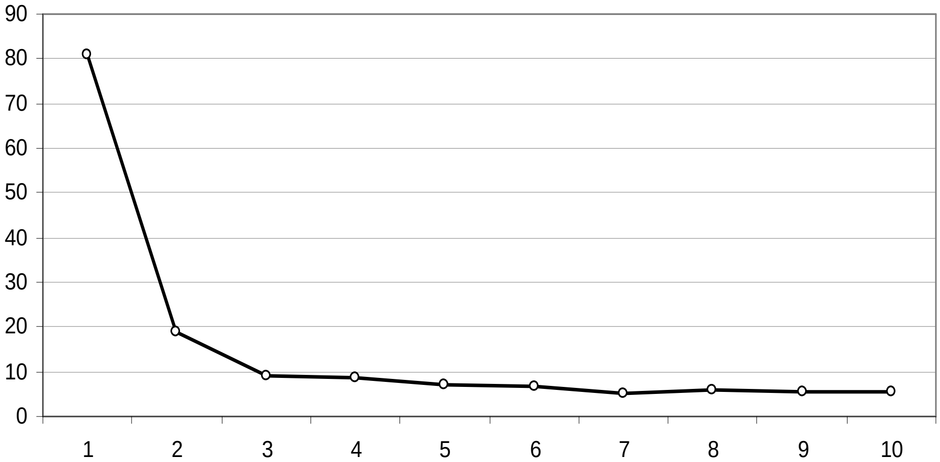


Рисунок 24 – График зависимости x и y

Для наглядности создадим график на основании данных x и y.

На основании графического изображения убеждаемся, что можно использовать линеаризующее преобразование для функции типа $y = ae^{\frac{b}{x}}$.

Обозначим $y^* = \ln(y)$ *и* $x^* = \dfrac{1}{x}$. *Тогда будем искать линейную регрессию* $y^* = a^* + b^* x^*$. *Для* x^* *и* y^* *имеем ряд:*

$$x^* \to \quad 1 \quad \frac{1}{2} \quad \frac{1}{3} \quad \frac{1}{4} \quad \frac{1}{5} \quad \frac{1}{6} \quad \frac{1}{7} \quad \frac{1}{8} \quad \frac{1}{9} \quad \frac{1}{10}$$

$$y^* \to 4{,}39 \quad 2{,}94 \quad 2{,}20 \quad 2{,}14 \quad 1{,}94 \quad 1{,}89 \quad 1{,}79 \quad 1{,}76 \quad 1{,}72 \quad 1{,}69$$

Вычислим: $\sum_{i=1}^{10} x_i^* = 2{,}92896$, $\left(\sum_{i=1}^{10} x_i^*\right)^2 = 8{,}578806$, $\sum_{i=1}^{10} x_i^{*2} = 1{,}549768$, $\sum_{i=1}^{10} y_i^* = 22{,}46$,

$\sum_{i=1}^{10} x_i^* y_i^* = 8{,}6671587$, $b^* = \dfrac{10 \cdot 8{,}6671587 - 2{,}92896 \cdot 22{,}46}{10 \cdot 1{,}549768 - 8{,}578806} = 3{,}019$,

$a^* = \dfrac{22{,}46 - 3{,}019 \cdot 2{,}92896}{10} = 1{,}361$. *Таким образом, искомая регрессия имеет вид:* $y^* = 1{,}361 + 3{,}019x^*$ *или* $y = 3{,}9e^{\frac{3{,}019}{x}}$.

Проверим значимость коэффициентов регрессии.

Рассмотрим коэффициент b^*.

Получаем: $S^2 = \dfrac{1}{n-2}\sum_{i=1}^{10}\left(y_i^* 1{,}361 - 3{,}019x_i^*\right)^2 = 0{,}004$, $S = 0{,}067$,

$$S_{x^*}^2 = \frac{1}{n-1}\sum_{i=1}^{10}\left(x_i^* - \overline{x^*}\right)^2 = 0{,}0768,\ S_{x^*} = 0{,}2772,$$

$$S_{b^*} = \frac{S}{S_{x^*}\sqrt{n-1}} = 0{,}0805.$$

Из таблицы $t_{0,975}(8) = 2{,}31$. *Так как* $\left|b^*\right| = 3{,}019 > t_{0,975}(8) \cdot S_{b^*} = 2{,}31 \cdot 0{,}0805 = 0{,}186$, *коэффициент регрессии признается значимым.*

Рассмотрим коэффициент a^*.

Получаем: $S_{a^*} = S\left(\dfrac{1}{n} + \dfrac{\overline{x^*}^2}{(n-1)S_{x^*}^2}\right)^{\frac{1}{2}} = 0{,}067\left(\dfrac{1}{10} + \dfrac{0{,}292896^2}{9 \cdot 0{,}0768}\right)^{\frac{1}{2}} = 0{,}0317$.

Так как $\left|a^*\right| = 1{,}361 > t_{0,975}(8) \cdot S_{a^*} = 2{,}31 \cdot 0{,}0317 = 0{,}073$, *то коэффициент регрессии признается значимым.*

Глава 6. Контрольные карты

Термин «контрольная карта», перевод с английского термина «control card», означает, что это карта управления или регулирования технологических процессов.

Совокупность контролируемых показателей качества изделия представляется, если отвлечься от их физической природы, потоком случайных величин, и задачей контролера становится своевременное распознание ухудшения качества изделия или признаков нарушения технологического процесса его изготовления.

Сегодня статистические карты контроля качества стали настолько общепринятыми, что любая фирма, не пользующаяся им в том или ином виде, оказывается в невыгодном положении по сравнению с конкурентами.

Контрольные карты применяются следующим образом. Делается предположение о распределении совокупности или выборочных статистик совокупности. Отбираются независимые случайные повторные выборки из текущей продукции. Подсчитываются и наносятся на специально подготовленную карту выборочные значения контролируемых статистик вместе с дву- или односторонними (в зависимости от специфики контроля) доверительными интервалами. Если текущие изменения показателей качества продукции находятся внутри доверительного интервала – процесс производства статистически управляем, качество продукции находится в допустимых пределах; если нет – требуется регулирующее вмешательство в производственный процесс.

Контрольные карты дают возможность фиксировать, когда и где наблюдаемые величины выходят за рамки случайных колебаний, указывая на необходимость регулирования технологического процесса.

Класс контрольных карт Шухарта позволяет определить, находится ли технологический процесс в одном и том же статистически устойчивом состоянии или в отдельные периоды времени было статистически значимое нарушение этого состояния. Контрольная карта Шухарта может применяться в двух видах задач:

- анализа состояния технологического процесса с целью обнаружения дестабилизирующих воздействий (факторов);

- слежения за текущем состоянием технологического процесса и остановками или регулировками технологического процесса в случае необходимости, попутно также можно обнаруживать дестабилизирующие факторы.

Характерным отличаем контрольных карт Шухарта является способ определения контрольных границ. Они определяются на основе оценки σ: в первой задаче – по той же выборке; во второй задаче – по предварительному исследованию. Устанавливаются контрольные границы на расстоянии -3σ и $+3\sigma$ соответственно от «центра», где «центром» в первой задаче является выборочное среднее, а во второй задаче – целевое значение.

Замечание. Если точки на контрольной карте лежат внутри контрольных границ, то считается, что все колебания точек объясняются чисто случайными факторами. Если же одна или несколько точек выходят за контрольные границы, то считается, что такие сильные отклонения невозможны за счет чистой случайности, то есть имеется воздействие неслучайного фактора: в первой задаче – это повод для обнаружения такого фактора, во второй задаче – повод для остановки и регулировки технологического процесса.

6.1. Контрольные карты по количественному признаку

6.1.1. $\bar{x}-$ *и R- карты*

Пусть в последовательно контролируемых выборках объёма n измеряется некоторая характеристика изделия - x. В каждой выборке измеряется среднее значение контролируемого признака $\bar{x}$ и его размах: $R=x_{\max}-x_{\min}$. Обозначим их для i-ой выборки $\bar{x}_i$ и R_i. В $\bar{x}-$ карте контрольные границы для среднего при достоверности $\alpha=0{,}997$ будут:

$$x_н = \bar{\bar{x}} - A\bar{R} \text{ (нижняя граница) ; } \quad x_в = \bar{\bar{x}} + A\bar{R} \text{ (верхняя граница) ,}$$

где $\bar{\bar{x}} = \frac{1}{k}\sum_{i=1}^{k}\bar{x}_i$ - среднее по k текущим выборкам; $\bar{R} = \frac{1}{k}\sum_{i=1}^{k}R_i$ - средний размах по k текущим выборкам; A - коэффициент, значения которого приведены в таблице приложения 9.

Для R- карты средним уровнем (центральным) является $\bar{R}$, нижняя и верхняя границы регулирования равны соответственно $R_н = D_1\bar{R}$ и $R_в = D_2\bar{R}$, где D_1 и D_2 - коэффициенты, приведённые в таблице приложения 9.

Замечание. $\bar{x}-$ и R- карты Шухарта рекомендуется применять при $n<10\div 20$.

Задача 40. В результате контроля 25 выборок изделий, при объёме каждой выборки n=5, получены следующие значения контролируемого параметра x.

Номер выборки i	Выборочные значения x_{ij}				
	x_1	x_2	x_3	x_4	x_5
1	1	-1	-2	2	0
2	-2	0	-2	2	0

3	0	0	1	-1	3
4	1	2	-1	2	-2
5	2	1	-1	1	0
6	2	-1	2	-1	0
7	-2	-1	0	-2	0
8	1	1	0	1	-1
9	0	-2	1	-1	4
10	-1	-1	0	-2	1
11	1	-1	2	-1	0
12	2	2	-1	-1	0
13	0	-2	2	-2	0
14	0	-1	1	-1	2
15	-3	0	0	0	-1
16	0	-1	0	1	2
17	3	1	-3	-2	0
18	-1	-2	-1	-1	0
19	0	0	0	1	1
20	0	-5	2	2	0
21	2	-1	-1	0	0
22	0	0	-3	2	-3
23	3	1	4	-2	-1
24	1	-3	0	1	-1
25	-1	0	1	0	1

Вычислить граница для контрольных $\bar{x}$ – и R- карт Шухарта (k=25).
Решение.
Заполним таблицу:

Номер выборки i	*Выборочные значения x_{ij}*					*Статистики*			
	x_1	x_2	x_3	x_4	X_5	$\bar{x}_i$	R_i	S_i^2	S_i
1	1	-1	-2	2	0	0,0	4	2,5	1,58
2	-2	0	-2	2	0	-0,4	4	2,8	1,67
3	0	0	1	-1	3	0,6	4	2,3	1,52
4	1	2	-1	2	-2	0,4	4	3,3	1,82
5	2	1	-1	1	0	0,6	3	1,3	1,14
6	2	-1	2	-1	0	0,4	3	2,3	1,52
7	-2	-1	0	-2	0	-1,0	2	1,0	1,00
8	1	1	0	1	-1	0,4	2	0,8	0,89
9	0	-2	1	-1	4	0,4	6	5,3	2,30
10	-1	-1	0	-2	1	-0,6	3	1,3	1,14
11	1	-1	2	-1	0	0,2	3	1,7	1,30
12	2	2	-1	-1	0	0,4	3	2,3	1,52

13	0	-2	2	-2	0	-0,4	4	2,8	1,67
14	0	-1	1	-1	2	0,2	3	1,7	1,30
15	-3	0	0	0	-1	-0,8	3	1,7	1,30
16	0	-1	0	1	2	0,4	3	1,3	1,14
17	3	1	-3	-2	0	-0,2	6	5,7	2,39
18	-1	-2	-1	-1	0	-1,0	2	0,5	0,71
19	0	0	0	1	1	0,4	1	0,3	0,55
20	0	-5	2	2	0	-0,2	7	8,2	2,86
21	2	-1	-1	0	0	0,4	3	1,3	1,14
22	0	0	-3	2	-3	-0,8	5	4,7	2,17
23	3	1	4	-2	-1	1,0	6	6,5	2,55
24	1	-3	0	1	-1	-0,4	4	2,8	1,67
25	-1	0	1	0	1	0,2	2	0,4	0,63

Имеем $\bar{x} = \frac{1}{25}\sum_{i=1}^{25}\bar{x}_i = \frac{0,2}{25} = 0,008$, $\bar{R} = \frac{1}{25}\sum_{i=1}^{k} R_i = \frac{90}{25} = 3,6$.

Контрольные границы для среднего (для n=5 из таблицы приложения 9 находим A=0,577) будут равны

$x_н = \bar{x} - A\bar{R} = 0,008 - 0,577 \cdot 3,6 = -2,069$; $x_в = \bar{x} + A\bar{R} = 0,008 + 0,577 \cdot 3,6 = 2,085$.

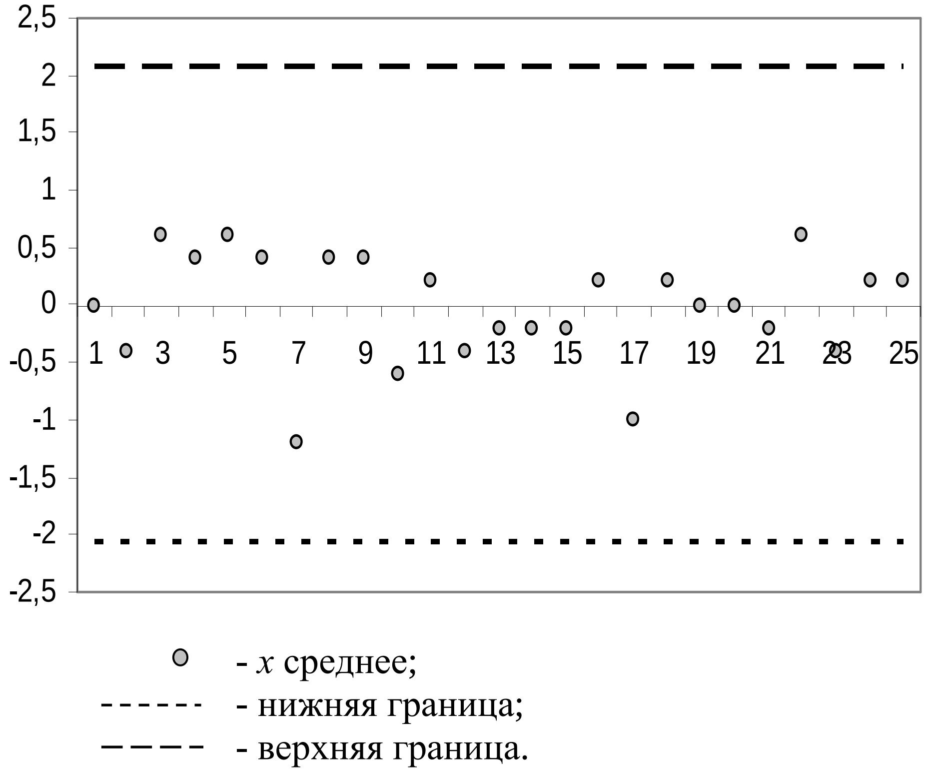


Рисунок 25 – Контрольная карта

Мы видим, что нигде $\bar{x}_i$ *не вышли за контрольные границы.*

Теперь для R- карты из таблицы приложения 9 имеем $D_1=0$ $D_2=2,115$.

Тогда $R_н = D_1\overline{R} = 0$, $R_в = D_2\overline{R} = 2{,}115 \cdot 3{,}6 = 7{,}61$. *Видим, что ни одно значение* R_i *не выходит за эти пределы, то есть процесс статистически управляем.*

6.1.2. s-карта

s-карта для выборочного контроля среднего квадратического отклонения более чувствительна к изменению рассеяния, чем *R*- карта. Алгоритм построения и использования s-карты состоит в следующем. По текущим *k* выборкам вычисляется оценка дисперсии контролируемого признака

$$s^2 = \frac{\sum_{i=1}^{k} v_i s_i^2}{\sum_{i=1}^{k} v_i}, \text{ где } v_i = n_i - 1; \; s_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}(x_{ij} - \overline{x}_i)^2 .$$

Если объёмы выборок n_i одинаковы, то $s^2 = \frac{1}{k}\sum_{i=1}^{k} s_i^2$. Центральная линия *s*-карты равна $s_ц = C \cdot s$, а граница зоны регулирования равны $s_н = B_1 \cdot s$ и $s_в = B_2 \cdot s$, где B_1, B_2 и С – коэффициенты, приведённые в таблице приложения 9.

Задача 41. Для данных задачи 39 построить контрольные границы для s-карты Шухарта.

В результате контроля 25 выборок изделий, при объёме каждой выборки *n*=5, получены следующие значения контролируемого параметра *x*.

Номер Выборки *i*	Выборочные значения x_{ij}				
	x_1	x_2	x_3	x_4	x_5
1	1	-1	-2	2	0
2	-2	0	-2	2	0
3	0	0	1	-1	3
4	1	2	-1	2	-2
5	2	1	-1	1	0
6	2	-1	2	-1	0
7	-2	-1	0	-2	0
8	1	1	0	1	-1
9	0	-2	1	-1	4
10	-1	-1	0	-2	1
11	1	-1	2	-1	0
12	2	2	-1	-1	0
13	0	-2	2	-2	0

14	0	-1	1	-1	2
15	-3	0	0	0	-1
16	0	-1	0	1	2
17	3	1	-3	-2	0
18	-1	-2	-1	-1	0
19	0	0	0	1	1
20	0	-5	2	2	0
21	2	-1	-1	0	0
22	0	0	-3	2	-3
23	3	1	4	-2	-1
24	1	-3	0	1	-1
25	-1	0	1	0	1

Решение.

Имеем $s^2 = \frac{1}{25}\sum_{i=1}^{k} s_i^2 = 2{,}592$ *(s=1,610).*

Следовательно, для контрольных границ имеем $s_ц = 0{,}940 \cdot 1{,}610 = 1{,}513$*;* $s_н = 0$*;* $s_в = 1{,}889 \cdot 1{,}610 = 3{,}041$*(предварительно из таблицы приложения 9 для n=5 получили C=0,940,* B_1*=0,* B_2*=1,889).*

Таким образом, из всех выборок только s_i *в 20-й выборке близко к* $s_в = 3{,}041$ *(впрочем, то же самое было и для R- карты).*

6.1.3. $\bar{x}-$ и s- контрольные карты для выборок неравного объёма

Пусть имеем k выборок объёмами n_i ($i=1,2,\ldots k$). Если объёмы выборок не равны между собой, то, естественно, границы регулирования $\bar{x}-$ и s-карт будут меняться от выборки к выборке. В этом случае следует поступать следующим образом. По всем выборкам находим

$$\bar{\bar{x}} = \frac{\sum_{i=1}^{k} n_i \bar{x}_i}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}}{\sum_{i=1}^{k} n_i} \text{ и } s^2 = \frac{\sum_{i=1}^{k}(n_i - 1)s_i^2}{\sum_{i=1}^{k}(n_i - 1)}.$$

Для $\bar{x}-$ карты имеем: $x_ц = \bar{\bar{x}}$ – центральная линия; $x_н = \bar{\bar{x}} - A_i \cdot s$ – нижняя граница; $x_в = \bar{\bar{x}} + A_i \cdot s$ – верхняя граница.

Для s- карты имеем: $s_ц = C_i \cdot s$ – центральная линия; $s_н = B_{1i} \cdot s$ – нижняя граница; $s_в = B_{2i} \cdot s$ – верхняя граница.

Здесь B_{1i},B_{2i} и C_i –коэффициенты из таблицы (приложение 9), определяемые для объёма каждой выборки; A_i-коэффициенты, приведённые в таблице 11.

Таблица 11

Значения коэффициентов A_i для $\bar{x}$ – карт Шухарта
при неравных объёмах контролируемых выборок

n_i	A_i	n_i	A_i	n_i	A_i
2	2,121	10	0,949	18	0,770
3	1,732	11	0,905	19	0,688
4	1,500	12	0,866	20	0,671
5	1,342	13	0,832	21	0,655
6	1,225	14	0,802	22	0,640
7	1,134	15	0,775	23	0,626
8	1,061	16	0,750	24	0,612
9	1,000	17	0,728	25	0,600

Задача 42. В результате контроля получены результаты измерений параметра x (средние значения и дисперсии, полученные в результате контроля в выборках разного объёма, приведены в таблице). Необходимо найти контрольные границы для $\bar{x}$ – и s- карт.

Номер выборки, i	1	2	3	4	5	6	7	8	9	10
Объём выборки, n_i	8	10	12	7	10	6	8	6	5	12
$\bar{x}_i$	1356	1380	1448	1358	1372	1430	1356	1426	1444	1404
s_i^2	1281	2570	1901	1384	1430	1672	1354	643	3418	1608

Решение.

Для $\bar{x}$ – карты имеем центральную линию

$$\bar{\bar{x}} = x_ц = \frac{\sum_{i=1}^{10} n_i \cdot x_i}{\sum_{i=1}^{10} n_i} = \frac{8 \cdot 1356 + 10 \cdot 1380 + ... + 12 \cdot 1404}{8 + 10 + ...12} = 1396{,}5 .$$

Коэффициенты A_i определяются для каждой выборки по таблице 11. Например, для первой выборки имеем n_1=8 и A_1=1,061, для второй выборки n_2=10 и A_2=0,949, и т.д.

Найдем оценку средней дисперсии: $s^2 = \frac{\sum_{i=1}^{10}(n_i-1)s_i^2}{\sum_{i=1}^{10}(n_i-1)} = 1680{,}4$, *s=40,99.*

Тогда для первой выборки имеем

$x_{н1} = \overline{\overline{x}} - A_1 \cdot s = 1396{,}5 - 1{,}061 \cdot 40{,}99 = 1353$; $x_{в1} = \overline{\overline{x}} + A_1 \cdot s = 1396{,}5 + 1{,}061 \cdot 40{,}99 = 1440$.

Полученные по аналогии результаты для других выборок приведены в таблице.

Для расчёта центральной линии и контрольных границ для s- карты пользуемся коэффициентами B_{1i},B_{2i} *и* C_i *из таблицы приложения 9, взятыми для соответствующих объёмов выборок* n_i.

Например, для первой выборки имеем $n_1=8$, $C_1=0{,}965$, $B_{11}=0{,}215$ *и* $B_{21}=1{,}715$.

Следовательно, центральная линия равна: $s_{ц1} = C_1 \cdot s = 0{,}965 \cdot 40{,}99 = 39{,}55$,

а границы регулирования равны:

$s_{н1} = B_{11} \cdot s = 0{,}215 \cdot 40{,}99 = 8{,}18$; $s_{в1} = B_{21} \cdot s = 1{,}715 \cdot 40{,}99 = 70{,}30$.

Вычисленные по аналогии значения $s_{цi}$, $s_{нi}$, $s_{вi}$ *для остальных выборок приведены в таблице.*

Номер выборки, i	*Объём выборки,* n_i	$\overline{x}$ *– карта*			*s- карта*				
		$\overline{x}_i$	$x_{нi}$	$x_{вi}$	s_i^2	s_i	$s_{цi}$	$s_{нi}$	$s_{вi}$
1	8	1356	1353	1440	1281	35,8	39,55	8,18	70,30
2	10	1380	1357,6	1435,4	2570	50,7	39,88	12,38	67,39
3	12	1448	1361	1432	1901	43,6	40,05	14,96	65,13
4	7	1358	1350	1443	1384	37,2	40,57	16,07	64,27
5	10	1372	1357,6	1435,4	1430	37,8	40,21	16,97	63,45
6	6	1430	1346,3	1446,7	1672	40,9	40,25	17,79	62,71
7	8	1356	1353	1440	1354	36,8	40,33	18,61	62,06
8	6	1426	1346,3	1446,7	643	25,3	40,33	19,22	61,44
9	5	1444	1341,5	1451,5	3418	58,5	40,42	19,92	60,91
10	12	1404	1361	1432	1608	40,1	40,42	20,49	60,34

Таким образом, получаем, что среднее значение в третьей выборке выходит за границы регулирования, а все стандартные отклонения находятся в пределах границ регулирования.

Задача 43. В цехе принято решение перевести на статистическое регулирование технологический процесс изготовления болта на автоматах. За показатель качества выбран диаметр болта, равный 26 мм, и его допускаемые отклонения: e_s = -0,005 мм; e_i = -0,019 мм. Построить $\overline{x}$ – и s- контрольные карты и провести по ним статистический анализ процесса. Для упрощения измерений и вычислений измерительный прибор (рычажная скоба) был настроен на размер 25,980 мм. Результаты измерений (отклонения от размера 25,980 мм в микрометрах) приведены в таблице.

<table>
<tr><td colspan="2">Цех автоматный</td><td colspan="2">Оборудование – токарный автомат 5803</td><td>Контролируемая операция – нарезание резьбы</td><td colspan="2">Контролируемый параметр – $\varnothing 26^{-0,005}_{-0,019}$</td></tr>
<tr><td colspan="2">Объём контроля N=100</td><td colspan="2">Объём выборки n=5</td><td>Средство контроля – рычажная скоба</td><td colspan="2"></td></tr>
<tr><td>Время</td><td>№ выборки</td><td colspan="5">Результаты контроля</td></tr>
<tr><td>7.00</td><td>1</td><td>10</td><td>3</td><td>5</td><td>14</td><td>10</td></tr>
<tr><td>8.00</td><td>2</td><td>2</td><td>14</td><td>8</td><td>13</td><td>11</td></tr>
<tr><td>9.00</td><td>3</td><td>12</td><td>12</td><td>3</td><td>8</td><td>10</td></tr>
<tr><td>10.00</td><td>4</td><td>12</td><td>14</td><td>7</td><td>11</td><td>9</td></tr>
<tr><td>11.00</td><td>5</td><td>10</td><td>11</td><td>9</td><td>15</td><td>7</td></tr>
<tr><td>12.00</td><td>6</td><td>11</td><td>12</td><td>11</td><td>14</td><td>12</td></tr>
<tr><td>13.00</td><td>7</td><td>15</td><td>11</td><td>14</td><td>8</td><td>3</td></tr>
<tr><td>14.00</td><td>8</td><td>12</td><td>14</td><td>12</td><td>11</td><td>11</td></tr>
<tr><td>15.00</td><td>9</td><td>11</td><td>7</td><td>11</td><td>13</td><td>9</td></tr>
<tr><td>16.00</td><td>10</td><td>14</td><td>10</td><td>9</td><td>12</td><td>8</td></tr>
<tr><td>7.00</td><td>11</td><td>9</td><td>11</td><td>14</td><td>10</td><td>13</td></tr>
<tr><td>8.00</td><td>12</td><td>13</td><td>13</td><td>6</td><td>4</td><td>13</td></tr>
<tr><td>9.00</td><td>13</td><td>5</td><td>8</td><td>3</td><td>3</td><td>4</td></tr>
<tr><td>10.00</td><td>14</td><td>8</td><td>5</td><td>6</td><td>9</td><td>13</td></tr>
<tr><td>11.00</td><td>15</td><td>8</td><td>4</td><td>9</td><td>5</td><td>8</td></tr>
<tr><td>12.00</td><td>16</td><td>4</td><td>12</td><td>10</td><td>6</td><td>10</td></tr>
<tr><td>13.00</td><td>17</td><td>10</td><td>6</td><td>13</td><td>10</td><td>5</td></tr>
<tr><td>14.00</td><td>18</td><td>7</td><td>9</td><td>12</td><td>1</td><td>7</td></tr>
<tr><td>15.00</td><td>19</td><td>4</td><td>7</td><td>6</td><td>7</td><td>12</td></tr>
<tr><td>16.00</td><td>20</td><td>10</td><td>10</td><td>6</td><td>9</td><td>3</td></tr>
</table>

Решение.

Вычислим средние значения и средние квадратичные отклонения для полученных данных в результате контроля в выборках объема n=5. Результаты приведем в таблице.

№ выборки	*Результаты контроля*					$\overline{x}_i$	s_i
1	10	3	5	14	10	8,4	4,393177
2	2	14	8	13	11	9,6	4,827007

3	12	12	3	8	10	9	3,741657
4	12	14	7	11	9	10,6	2,701851
5	10	11	9	15	7	10,4	2,966479
6	11	12	11	14	12	12	1,224745
7	15	11	14	8	3	10,2	4,868265
8	12	14	12	11	11	12	1,224745
9	11	7	11	13	9	10,2	2,280351
10	14	10	9	12	8	10,6	2,408319
11	9	11	14	10	13	11,4	2,073644
12	13	13	6	4	13	9,8	4,438468
13	5	8	3	3	4	4,6	2,073644
14	8	5	6	9	13	8,2	3,114482
15	8	4	9	5	8	6,8	2,167948
16	4	12	10	6	10	8,4	3,286335
17	10	6	13	10	5	8,8	3,271085
18	7	9	12	1	7	7,2	4,024922
19	4	7	6	7	12	7,2	2,949576
20	10	10	6	9	3	7,6	3,04959

Для $\bar{x}$ *– карты имеем центральную линию:* $\bar{\bar{x}} = x_ц = \dfrac{\sum_{i=1}^{10} n_i \cdot x_i}{\sum_{i=1}^{10} n_i} = 9{,}15$.

Для s *– карты оценка средней дисперсии:* $s^2 = \dfrac{\sum_{i=1}^{10}(n_i - 1)s_i^2}{\sum_{i=1}^{10}(n_i - 1)} = 10{,}44$, *s=3,231.*

Коэффициент A_i *определяется по таблице 11, где* n*=5, тогда* A_5*=1,732.*

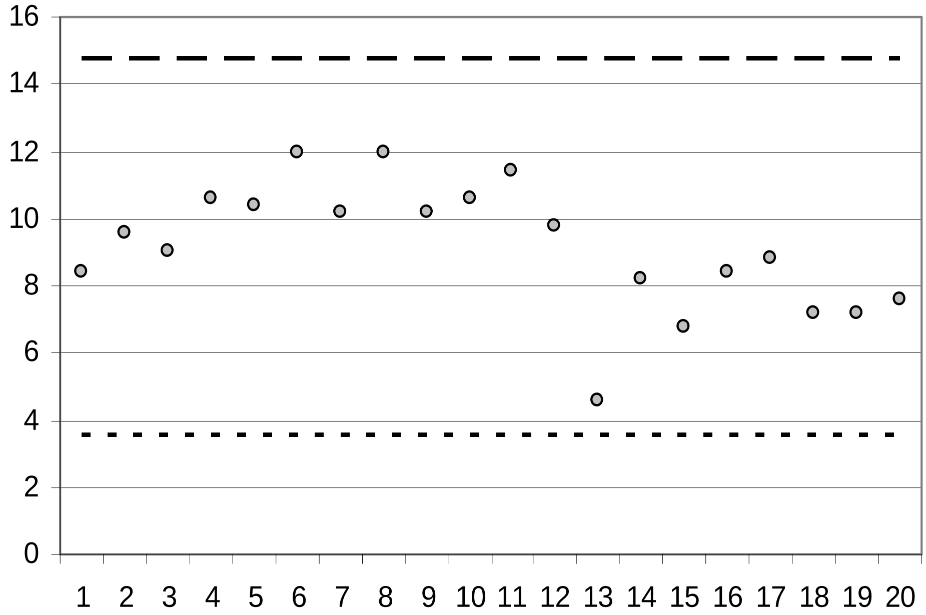
Тогда имеем: $x_{н1} = \bar{\bar{x}} - A_5 \cdot s = 9{,}15 - 1{,}732 \cdot 3{,}231 = 3{,}55$;

$x_{в1} = \bar{\bar{x}} + A_5 \cdot s = 9{,}15 + 1{,}732 \cdot 3{,}231 = 14{,}75$.

Коэффициенты B_{1i},B_{2i} *и* C_i *определяются из таблицы приложения 9, взятыми для соответствующего объёма выборки* n*=5:* C_5*=0,965,* B_{15}*=0 и* B_{25}*=1,889.*

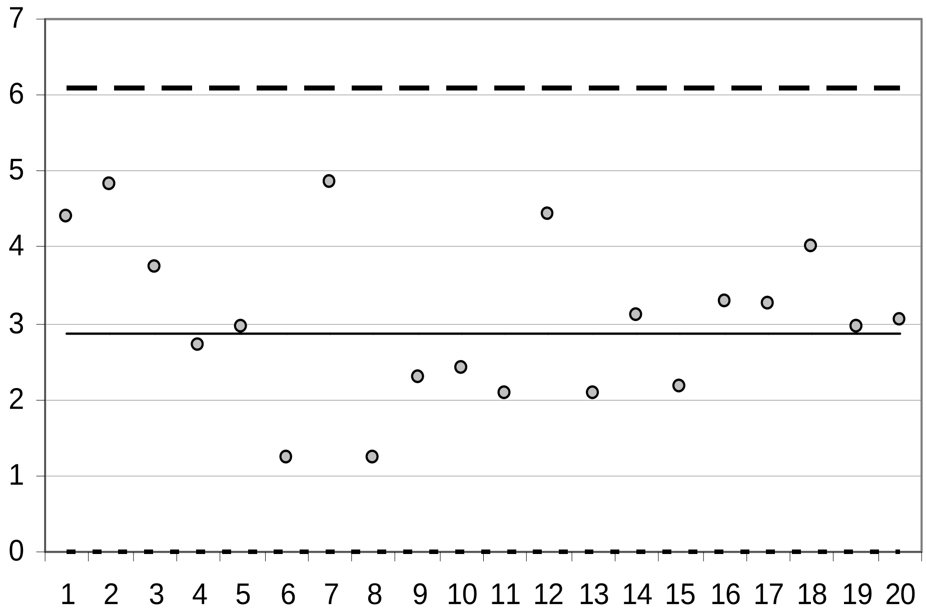
Следовательно, центральная линия равна: $s_{ц1} = C_5 \cdot s = 2{,}863$, *а границы регулирования равны:* $s_{н1} = B_{15} \cdot s = 0$; $s_{в1} = B_{25} \cdot s = 6{,}10$.

Изобразим для наглядности $\bar{x}$ *– контрольную карту (Рисунок26) и* s *- контрольную карту (Рисунок 27).*



- средние значения;
- нижняя граница;
- верхняя граница.

Рисунок 26 – $\bar{x}$ – контрольная карта



- средние квадратические отклонения;
- центральная линия;
- нижняя граница;
- верхняя граница.

Рисунок 27 – s - контрольная карта

Анализ контрольных карт показывает, что рассеяние диаметра болта приемлемо, и по рассеянию процесс стабилен (оборудование настроено достаточно точно), поскольку на s-карте нет показаний разлаженности процесса. Однако на $\bar{x}$-карте имеются серии из девяти точек (с четвёртой по двенадцатую) и из восьми точек (с тринадцатой по двадцатую), расположенных по одну сторону от средней линии. Это указывает на нестабильность процесса. Видимо, в течение процесса, при переходе от двенадцатой к тринадцатой точке изменилось математическое ожидание диаметра. Следует постараться выяснить причину этой нестабильности и провести управляющее воздействие на процесс. После стабилизации контрольную карту следует построить заново.

6.2. Контрольная карта по качественному признаку: р-карта для доли дефектных изделий

Предположим, что наблюдается k выборок объёмами n_i, причём в i-ой выборке фиксируется m_i дефектных изделий ($1 \le i \le k$). Оценкой доли дефектных изделий в i-ой выборке является величина $p_i = \frac{m_i}{n_i}$.

Оценка среднего значения доли дефектных изделий $\bar{p} = \frac{\sum_{i=1}^{k} m_i}{\sum_{i=1}^{k} n_i}$ является центральной линией p-карты.

Границы регулирования определяются формулами:

$$p_{нi} = \bar{p} - 3\left[\frac{\bar{p}(1-\bar{p})}{n_i}\right]^{\frac{1}{2}}; \quad p_{вi} = \bar{p} + 3\left[\frac{\bar{p}(1-\bar{p})}{n_i}\right]^{\frac{1}{2}}.$$

Если все значения n_i отличаются друг от друга незначительно, то используется единая пара контрольных границ:

$$p_{нг} = \bar{p} - 3\left[\frac{\bar{p}(1-\bar{p})}{\bar{n}}\right]^{\frac{1}{2}}; \quad p_{вг} = \bar{p} + 3\left[\frac{\bar{p}(1-\bar{p})}{\bar{n}}\right]^{\frac{1}{2}}, \text{ где } \bar{n} = \frac{1}{k}\sum_{i=1}^{k} n_i.$$

Задача 44. В ходе контроля различных партий получены результаты, приведённые в таблице.

Номер партии, i	1	2	3	4	5	6	7	8	9	10
Объём выборки, n_i	981	1422	1174	1524	1353	847	1535	1248	1296	985

Число дефектов, m_i	27	87	87	76	80	25	37	19	49	42

Найти границы регулирования р-карты.
Решение.
Найдём оценку средней доли дефектных изделий

$$\overline{p} = \frac{\sum_{i=1}^{10} m_i}{\sum_{i=1}^{10} n_i} = \frac{27+87+87+...+42}{981+1422+1174+...985} = 0{,}0428.$$

Далее, для каждой выборки вычисляются границы регулирования. Например, для i=1 имеем n_1=981 и

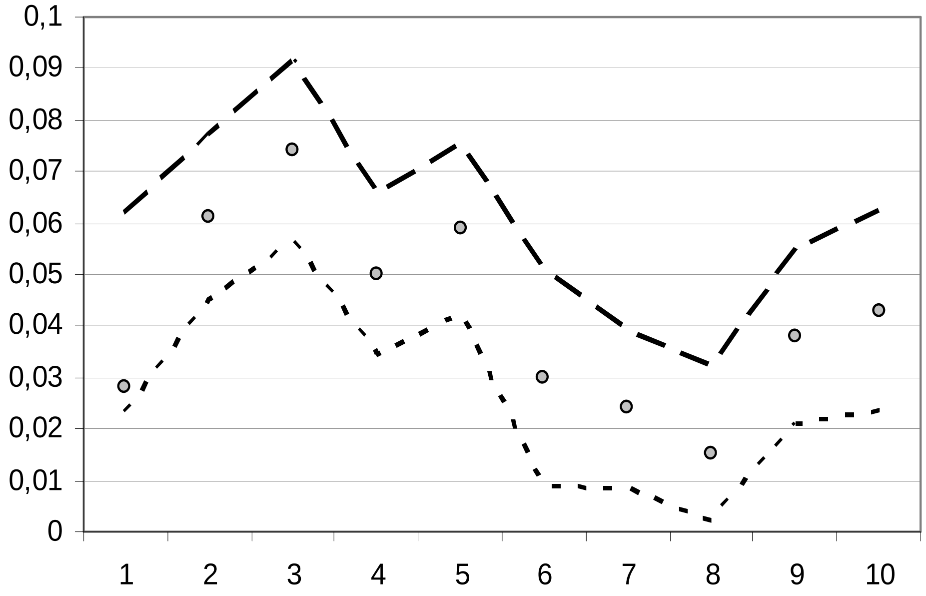
$$p_{н1} = 0{,}0428 - 3\left[\frac{0{,}04287 \cdot (1-0{,}0428)}{981}\right]^{\frac{1}{2}} = 0{,}0234;$$

$$p_{в1} = 0{,}0428 + 3\left[\frac{0{,}0428 \cdot (1-0{,}0428)}{981}\right]^{\frac{1}{2}} = 0{,}0622.$$

Рассчитанные по аналогии величины для остальных выборок приведены в таблице.

Номер партии, i	*Объём выборки, n_i*	*Число дефектов, m_i*	*Доля дефектных изделий, p_i*	*Границы регулирования*	
				$p_{нi}$	$p_{вi}$
1	981	27	0,028	0,0234	0,0622
2	1422	87	0,061	0,0449	0,0771
3	1174	87	0,074	0,0563	0,0917
4	1524	76	0,050	0,0344	0,0655
5	1353	80	0,059	0,0425	0,0755
6	847	25	0,030	0,0091	0,0509
7	1535	37	0,024	0,0085	0,0395
8	1248	19	0,015	0,0022	0,0322
9	1296	49	0,038	0,0211	0,0549
10	985	42	0,043	0,0236	0,0623

Изобразим полученные результаты наглядно (Рисунок 28).



- доля дефектных изделий;
- нижняя граница;
- верхняя граница.

Рисунок 28 – р-карта

Получаем, что доли дефектных изделий во всех выборках находятся в пределах границ регулирования.

Задания для самостоятельной работы
Вариант 1.

Задача 1. В результате статистического исследования получены следующие данные:

x_i	0,8	1,2	1,6	2,0	2,4	2,8	3,2	3,6	4,0	4,4
w_i	0,1	0,16	0,19	0,17	0,13	0,09	0,07	0,05	0,03	0,01

Объем выборки $n = 50$. Определить модель распределения и подтвердить гипотезу критерием χ^2 (проверить три приблизительно подходящие теоретические распределения).

Задача 2. Имеются четыре выборки объема n=5 каждая

x_1	10	11	13	14	16
x_2	9	14	19	24	29
x_3	8	10	12	14	16
x_4	9	12	15	18	21

1) Проверить гипотезу равенства средних при доверительной вероятности α =0,95;

2) Проверить нулевую гипотезу равенства дисперсий критерием Кохрана при доверительной вероятности α =0,95.

Задача 3. Провести двухфакторный дисперсионный анализ данных, представленных следующей таблицей, при доверительной вероятности α =0,95:

B	A		
	A_1	A_2	A_3
B_1	1,1 1,2 1,0	0,9 1,0 1,1	0,8 1,0 1,4
B_2	1,3 1,0 1,5	1,2 1,2 1,0	1,3 1,0 0,9
B_3	1,6 1,4 1,5	0,8 0,9 0,8	1,2 1,1 1,2
B_4	1,3 1,3 1,2	1,1 1,0 0,9	1,0 1,0 1,1

Задача 4. В результате контроля 10 выборок изделий, при объёме каждой выборки n=5, получены следующие значения контролируемого параметра x.

Выборочные	Номер выборки i									
значения x_{ij}	1	2	3	4	5	6	7	8	9	10
x_1	1	-4	0	-1	-2	1	-1	0	0	1
x_2	0	1	3	-2	0	1	4	-1	0	2
x_3	-1	-1	1	0	-2	0	3	1	-1	-2
x_4	2	2	0	-1	-3	2	-2	0	1	-1
x_5	-3	-2	-2	1	0	0	0	-1	3	3

Вычислить граница для контрольных $\bar{x}$ – и R- карт Шухарта.

Вариант 2.

Задача 1. В результате статистического исследования получены следующие данные:

x_i	0,2	0,4	0,6	0,8	1,0	1,2	1,4	1,6	1,8	2,0
w_i	0,09	0,2	0,18	0,14	0,12	0,1	0,08	0,04	0,03	0,02

Объем выборки $n = 50$. Определить модель распределения и подтвердить гипотезу критерием χ^2 (проверить три приблизительно подходящие теоретические распределения).

Задача 2. Имеются четыре выборки объема n=5 каждая

x_1	13	18	23	28	33
x_2	4	6	8	10	12
x_3	8	12	16	20	24
x_4	5	8	11	14	17

1) Проверить гипотезу равенства средних при доверительной вероятности α =0,95;

2) Проверить нулевую гипотезу равенства дисперсий критерием Кохрана при доверительной вероятности α =0,95.

Задача 3. Провести двухфакторный дисперсионный анализ данных, представленных следующей таблицей, при доверительной вероятности α =0,95:

B	A		
	A_1	A_2	A_3
B_1	0,5 0,6 0,4	0,6 0,7 0,8	0,3 0,5 0,7
B_2	0,4 0,3 0,4	0,7 0,7 0,9	0,7 0,6 0,8
B_3	0,6 0,7 0,5	0,8 0,9 0,8	0,5 0,5 0,7
B_4	0,4 0,5 0,5	1,0 0,7 0,9	0,9 0,9 0,7

Задача 4. В результате контроля 10 выборок изделий, при объёме каждой выборки n=5, получены следующие значения контролируемого параметра x.

Выборочные значения x_{ij}	Номер выборки i									
	1	2	3	4	5	6	7	8	9	10
x_1	2	-3	-1	1	-2	-1	0	1	1	0
x_2	0	0	2	2	1	1	1	-1	0	-2
x_3	1	-1	0	3	-1	2	-1	0	1	2
x_4	0	-1	0	0	-1	0	2	0	-1	-1
x_5	-2	1	2	0	1	-2	0	1	1	0

Вычислить граница для контрольных $\bar{x}$ – и R- карт Шухарта.

Вариант 3.

Задача 1. В результате статистического исследования получены следующие данные:

x_i	3,2	3,4	3,6	3,8	4,0	4,2	4,4	4,6	4,8	5,0
w_i	0,09	0,17	0,15	0,14	0,13	0,1	0,08	0,06	0,05	0,03

Объем выборки $n = 50$. Определить модель распределения и подтвердить гипотезу критерием χ^2 (проверить три приблизительно подходящие теоретические распределения).

Задача 2. Имеются четыре выборки объема n=5 каждая

x_1	21	24	27	30	33
x_2	20	22	24	26	28
x_3	18	22	26	30	34
x_4	17	19	21	23	25

1) Проверить гипотезу равенства средних при доверительной вероятности α =0,95;

2) Проверить нулевую гипотезу равенства дисперсий критерием Кохрана при доверительной вероятности α =0,95.

Задача 3. Провести двухфакторный дисперсионный анализ данных, представленных следующей таблицей, при доверительной вероятности α =0,95:

B	A		
	A_1	A_2	A_3
B_1	2,3 2,3 2,4	2,5 2,6 2,7	2,1 2,1 2,2
B_2	2,1 2,2 2,3	2,6 2,6 2,4	2,3 2,2 2,4
B_3	2,4 2,5 2,1	2,5 2,5 2,6	2,5 2,2 2,3
B_4	2,2 2,3 2,3	2,8 2,5 2,6	2,1 2,2 2,2

Задача 4. В результате контроля 10 выборок изделий, при объёме каждой выборки n=5, получены следующие значения контролируемого параметра x.

Выборочные значения x_{ij}	Номер выборки i									
	1	2	3	4	5	6	7	8	9	10
x_1	1	-3	-1	-1	0	0	3	0	2	1
x_2	0	3	-3	3	2	-1	2	-2	-2	-1
x_3	-1	0	4	0	-2	3	-2	3	0	0
x_4	0	-2	0	1	-3	1	1	-1	-2	-2
x_5	2	1	1	0	3	-1	-1	0	3	0

Вычислить граница для контрольных $\bar{x}$ – и R- карт Шухарта.

Вариант 4.

Задача 1. В результате статистического исследования получены следующие данные:

x_i	1,8	2,4	3,0	3,6	4,2	4,8	5,4	6,0	6,6	7,2
w_i	0,08	0,15	0,18	0,16	0,13	0,1	0,08	0,06	0,04	0,02

Объем выборки $n = 50$. Определить модель распределения и подтвердить гипотезу критерием χ^2 (проверить три приблизительно подходящие теоретические распределения).

Задача 2. Имеются четыре выборки объема n=5 каждая

x_1	10	12	14	16	18
x_2	16	19	22	25	28
x_3	12	16	20	24	28
x_4	14	17	20	23	26

1) Проверить гипотезу равенства средних при доверительной вероятности α =0,95;

2) Проверить нулевую гипотезу равенства дисперсий критерием Кохрана при доверительной вероятности α =0,95.

Задача 3. Провести двухфакторный дисперсионный анализ данных, представленных следующей таблицей, при доверительной вероятности α =0,95:

B	A		
	A_1	A_2	A_3
B_1	3,3 3,1 3,3	3,4 3,5 3,6	3,2 3,2 3,4
B_2	3,0 3,4 3,2	3,1 3,5 3,4	3,5 3,3 3,3
B_3	3,5 3,1 3,1	3,3 3,5 3,5	3,0 3,4 3,2
B_4	3,3 3,2 3,2	3,6 3,4 3,4	3,3 3,4 3,3

Задача 4. В результате контроля 10 выборок изделий, при объёме каждой выборки n=5, получены следующие значения контролируемого параметра x.

Выборочные значения x_{ij}	Номер выборки i									
	1	2	3	4	5	6	7	8	9	10
x_1	1	0	0	-1	-2	1	-1	0	0	1
x_2	0	1	-1	-2	0	1	0	-1	0	2
x_3	-1	-1	1	0	-2	0	1	1	-1	-2
x_4	2	2	0	-1	-1	2	-2	0	1	-1
x_5	1	-2	-2	1	0	0	0	-1	-1	1

Вычислить граница для контрольных $\bar{x}-$ и R- карт Шухарта.

Вариант 5.

Задача 1. В результате статистического исследования получены следующие данные:

x_i	1,6	2,3	3,0	3,7	4,4	5,1	5,8	6,5	7,2	7,9
w_i	0,07	0,09	0,19	0,17	0,15	0,12	0,09	0,07	0,03	0,02

Объем выборки $n=50$. Определить модель распределения и подтвердить гипотезу критерием χ^2 (проверить три приблизительно подходящие теоретические распределения).

Задача 2. Имеются четыре выборки объема n=5 каждая

x_1	31	36	41	46	51
x_2	24	28	32	36	40
x_3	14	16	18	20	22
x_4	20	25	30	35	40

1) Проверить гипотезу равенства средних при доверительной вероятности α =0,95;

2) Проверить нулевую гипотезу равенства дисперсий критерием Кохрана при доверительной вероятности α =0,95.

Задача 3. Провести двухфакторный дисперсионный анализ данных, представленных следующей таблицей, при доверительной вероятности α =0,95:

B	A		
	A_1	A_2	A_3
B_1	0,1 0,2 0,4	0,5 0,4 0,5	0,1 0,2 0,1
B_2	0,2 0,2 0,3	0,3 0,6 0,4	0,3 0,3 0,2
B_3	0,4 0,3 0,2	0,5 0,4 0,3	0,2 0,4 0,3
B_4	0,1 0,4 0,4	0,5 0,5 0,4	0,2 0,3 0,3

Задача 4. В результате контроля 10 выборок изделий, при объёме каждой выборки n=5, получены следующие значения контролируемого параметра x.

Выборочные значения x_{ij}	Номер выборки i									
	1	2	3	4	5	6	7	8	9	10
x_1	0,1	-0,3	-0,1	-0,1	0	0	0,3	0	0,2	0,1
x_2	0	0,3	-0,3	0,3	0,2	-0,1	0,2	-0,2	-0,2	-0,1
x_3	-0,1	0	0,4	0	-0,2	0,3	-0,2	0,3	0	0
x_4	0	-0,2	0	0,1	-0,3	0,1	0,1	-0,1	-0,2	-0,2
x_5	0,2	0,1	0,1	0	0,3	-0,1	-0,1	0	0,3	0

Вычислить граница для контрольных $\bar{x}-$ и R- карт Шухарта.

Вариант 6.

Задача 1. В результате статистического исследования получены следующие данные:

x_i	0,8	1,8	2,8	3,8	4,8	5,8	6,8	7,8	8,8	9,8
w_i	0,1	0,16	0,19	0,18	0,13	0,09	0,07	0,04	0,03	0,01

Объем выборки $n = 50$. Определить модель распределения и подтвердить гипотезу критерием χ^2 (проверить три приблизительно подходящие теоретические распределения).

Задача 2. Имеются четыре выборки объема n=5 каждая

x_1	16	19	22	25	28
x_2	24	29	34	39	44
x_3	13	18	24	30	36
x_4	10	12	14	16	18

1) Проверить гипотезу равенства средних при доверительной вероятности α =0,95;

2) Проверить нулевую гипотезу равенства дисперсий критерием Кохрана при доверительной вероятности α =0,95.

Задача 3. Провести двухфакторный дисперсионный анализ данных, представленных следующей таблицей, при доверительной вероятности α =0,95:

B	A		
	A_1	A_2	A_3
B_1	5,1 5,2 5,2	4,9 5,0 4,8	5,1 5,4 5,1
B_2	5,3 5,2 5,3	5,0 5,0 5,1	5,2 5,1 5,1
B_3	5,1 5,1 5,2	4,9 5,1 5,2	5,1 5,2 5,3
B_4	5,4 5,3 5,2	5,0 4,8 5,0	5,4 5,4 5,6

Задача 4. В результате контроля 10 выборок изделий, при объёме каждой выборки n=5, получены следующие значения контролируемого параметра x.

Выборочные	Номер выборки i									
значения x_{ij}	1	2	3	4	5	6	7	8	9	10
x_1	0,1	-0,2	0	-0,1	-0,2	0,1	-0,1	0	0	0,1
x_2	0	0,1	0,3	-0,2	0	0,1	0	-0,1	0	0,2
x_3	-0,1	-0,1	0,1	0	-0,2	0	0,3	0,1	-0,1	-0,2
x_4	0,2	0,2	0	-0,1	-0,1	0,2	-0,2	0	0,1	-0,1
x_5	-0,1	-0,2	-0,2	0,1	0	0	0	-0,1	0,2	0

Вычислить граница для контрольных $\bar{x}-$ и R- карт Шухарта.

Вариант 7.

Задача 1. В результате статистического исследования получены следующие данные:

x_i	0,3	0,8	1,3	1,8	2,3	2,8	3,3	3,8	4,3	4,8
w_i	0,08	0,18	0,16	0,14	0,12	0,1	0,09	0,06	0,04	0,03

Объем выборки $n=50$. Определить модель распределения и подтвердить гипотезу критерием χ^2 (проверить три приблизительно подходящие теоретические распределения).

Задача 2. Имеются четыре выборки объема n=5 каждая

x_1	1,5	2	2,5	3	3,5
x_2	1	2	3	4	5
x_3	0,75	1,25	1,75	2,25	2,75
x_4	1	1,1	1,2	1,3	1,4

1) Проверить гипотезу равенства средних при доверительной вероятности α =0,95;

2) Проверить нулевую гипотезу равенства дисперсий критерием Кохрана при доверительной вероятности α =0,95.

Задача 3. Провести двухфакторный дисперсионный анализ данных, представленных следующей таблицей, при доверительной вероятности α =0,95:

B	A		
	A_1	A_2	A_3
B_1	3,7 3,8 3,7	4,0 3,9 4,1	4,0 4,3 4,0
B_2	3,9 3,8 3,7	3,9 4,1 4,0	4,4 4,3 4,3
B_3	3,8 3,8 3,7	4,2 3,9 4,2	4,4 4,2 4,1
B_4	4,0 3,7 3,8	4,1 4,1 4,0	4,3 4,5 4,2

Задача 4. В результате контроля 10 выборок изделий, при объёме каждой выборки n=5, получены следующие значения контролируемого параметра x.

Выборочные значения x_{ij}	Номер выборки i									
	1	2	3	4	5	6	7	8	9	10
x_1	-1	-2	1	0	-3	0	2	1	-1	0
x_2	0	0	-2	-1	1	0	-1	0	1	1
x_3	1	-1	3	1	2	-1	0	-1	2	-1
x_4	0	-2	1	-3	-2	0	2	-1	-2	-3
x_5	2	1	2	2	0	-1	-1	1	0	-1

Вычислить граница для контрольных $\bar{x}-$ и R- карт Шухарта.

Вариант 8.

Задача 1. В результате статистического исследования получены следующие данные:

x_i	2,2	2,4	2,6	2,8	3,0	3,2	3,4	3,6	3,8	4,0
w_i	0,14	0,15	0,18	0,12	0,1	0,09	0,08	0,06	0,05	0,03

Объем выборки $n = 50$. Определить модель распределения и подтвердить гипотезу критерием χ^2 (проверить три приблизительно подходящие теоретические распределения).

Задача 2. Имеются четыре выборки объема n=5 каждая

x_1	0,3	0,6	0,9	1,2	1,5
x_2	0,25	0,45	0,65	0,85	1,05
x_3	0,2	0,4	0,6	0,8	1
x_4	0,1	0,2	0,3	0,4	0,5

1) Проверить гипотезу равенства средних при доверительной вероятности α =0,95;

2) Проверить нулевую гипотезу равенства дисперсий критерием Кохрана при доверительной вероятности α =0,95.

Задача 3. Провести двухфакторный дисперсионный анализ данных, представленных следующей таблицей, при доверительной вероятности α =0,95:

B	A		
	A_1	A_2	A_3
B_1	2,2 2,1 2,0	2,0 2,1 2,1	2,1 2,0 2,1
B_2	2,1 2,0 2,0	2,2 2,0 2,1	2,2 2,0 2,1
B_3	2,2 2,2 2,1	2,0 2,0 2,1	2,0 2,1 2,1
B_4	2,0 2,1 2,1	2,2 2,2 2,0	2,1 2,0 2,0

Задача 4. В результате контроля 10 выборок изделий, при объёме каждой выборки n=5, получены следующие значения контролируемого параметра x.

Выборочные	Номер выборки i									
значения x_{ij}	1	2	3	4	5	6	7	8	9	10
x_1	0,1	-0,3	-0,1	-0,1	0	0	0,3	0	0,2	0,1
x_2	0	0,3	-0,3	0,3	0,2	-0,1	0,2	-0,2	-0,2	-0,1
x_3	-0,1	0	0	0	-0,2	0,3	-0,2	0,3	0	0
x_4	0	-0,2	0	0,1	-0,3	0,1	0,1	-0,1	-0,2	-0,2
x_5	0,2	0,1	0,1	0	0,3	-0,1	-0,1	0	0,3	0

Вычислить граница для контрольных $\bar{x}-$ и R- карт Шухарта.

Вариант 9.

Задача 1. В результате статистического исследования получены следующие данные:

x_i	1,1	1,7	2,3	2,9	3,5	4,1	4,7	5,3	5,9	6,5
w_i	0,06	0,08	0,21	0,17	0,15	0,12	0,09	0,07	0,03	0,02

Объем выборки $n = 50$. Определить модель распределения и подтвердить гипотезу критерием χ^2 (проверить три приблизительно подходящие теоретические распределения).

Задача 2. Имеются четыре выборки объема n=5 каждая

x_1	0	0,3	0,6	0,9	1,2
x_2	0,25	0,35	0,45	0,55	0,65
x_3	0,2	0,23	0,26	0,29	0,32
x_4	0,1	0,2	0,3	0,4	0,5

1) Проверить гипотезу равенства средних при доверительной вероятности α =0,95;

2) Проверить нулевую гипотезу равенства дисперсий критерием Кохрана при доверительной вероятности α =0,95.

Задача 3. Провести двухфакторный дисперсионный анализ данных, представленных следующей таблицей, при доверительной вероятности α =0,95:

B	A		
	A_1	A_2	A_3
B_1	4,3 4,1 4,0	4,0 4,1 4,1	4,1 4,0 4,3
B_2	4,1 4,0 4,0	4,2 4,0 4,1	4,2 4,0 4,1
B_3	4,2 4,2 4,1	4,1 4,0 4,3	4,0 4,1 4,1
B_4	4,0 4,2 4,1	4,2 4,2 4,0	4,1 4,0 4,0

Задача 4. В результате контроля 10 выборок изделий, при объёме каждой выборки n=5, получены следующие значения контролируемого параметра x.

Выборочные значения x_{ij}	Номер выборки i									
	1	2	3	4	5	6	7	8	9	10
x_1	-1	-2	1	0	-2	0	2	1	-1	0
x_2	0	0	-2	-1	1	0	-1	0	1	1
x_3	1	-1	1	1	2	-1	0	-1	2	-1
x_4	0	-2	1	0	-2	0	2	-1	-2	-1
x_5	2	1	2	2	0	-1	-1	1	0	-1

Вычислить граница для контрольных $\bar{x}$ – и R- карт Шухарта.

Вариант 10.

Задача 1. В результате статистического исследования получены следующие данные:

x_i	1,6	2,1	2,6	3,1	3,6	4,1	4,6	5,1	5,6	6,1
w_i	0,1	0,19	0,17	0,16	0,13	0,09	0,07	0,05	0,03	0,01

Объем выборки $n = 50$. Определить модель распределения и подтвердить гипотезу критерием χ^2 (проверить три приблизительно подходящие теоретические распределения).

Задача 2. Имеются четыре выборки объема n=5 каждая

x_1	0,13	0,17	0,21	0,25	0,29
x_2	0,25	0,55	0,85	1,15	1,45
x_3	0,2	0,23	0,26	0,29	0,32
x_4	0,01	0,02	0,03	0,04	0,05

1) Проверить гипотезу равенства средних при доверительной вероятности α=0,95;

2) Проверить нулевую гипотезу равенства дисперсий критерием Кохрана при доверительной вероятности α=0,95.

Задача 3. Провести двухфакторный дисперсионный анализ данных, представленных следующей таблицей, при доверительной вероятности α=0,95:

B	A		
	A_1	A_2	A_3
B_1	7,3 7,4 7,2	7,1 7,1 7,2	7,5 7,3 7,4
B_2	7,0 7,1 7,3	7,3 7,2 7,1	7,4 7,3 7,3
B_3	7,2 7,1 7,2	7,0 7,3 7,4	7,5 7,4 7,2
B_4	7,1 7,0 7,0	7,2 7,2 7,3	7,2 7,3 7,2

Задача 4. В результате контроля 10 выборок изделий, при объёме каждой выборки n=5, получены следующие значения контролируемого параметра x.

Выборочные значения x_{ij}	Номер выборки i									
	1	2	3	4	5	6	7	8	9	10
x_1	-0,1	-0,2	0,1	0	-0,3	0	0,2	0,1	-0,1	0
x_2	0	0	-0,2	-0,1	0,1	0	-0,1	0	0,1	0,1
x_3	0,1	-0,1	0,3	0,1	0,2	-0,1	0	-0,1	0,2	-0,1
x_4	0	-0,2	0,1	-0,3	-0,2	0	0,2	-0,1	-0,2	-0,3
x_5	0,2	0,1	0,2	0,2	0	-0,1	-0,1	0,1	0	-0,1

Вычислить граница для контрольных $\bar{x}$ – и R- карт Шухарта.

Приложение 1.

Значения функции Лапласа $\Phi(x)=\dfrac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-\frac{t^2}{2}}dt$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000	**0,26**	0,1026	**0,52**	0,1985	**0,78**	0,2823
0,01	0,0040	**0,27**	0,1064	**0,53**	0,2019	**0,79**	0,2852
0,02	0,0080	**0,28**	0,1103	**0,54**	0,2054	**0,80**	0,2881
0,03	0,0120	**0,29**	0,1141	**0,55**	0,2088	**0,81**	0,2910
0,04	0,0160	**0,30**	0,1179	**0,56**	0,2123	**0,82**	0,2939
0,05	0,0199	**0,31**	0,1217	**0,57**	0,2157	**0,83**	0,2967
0,06	0,0239	**0,32**	0,1255	**0,58**	0,2190	**0,84**	0,2995
0,07	0,0279	**0,33**	0,1293	**0,59**	0,2224	**0,85**	0,3023
0,08	0,0319	**0,34**	0,1331	**0,60**	0,2257	**0,86**	0,3051
0,09	0,0359	**0,35**	0,1368	**0,61**	0,2291	**0,87**	0,3078
0,10	0,0398	**0,36**	0,1406	**0,62**	0,2324	**0,88**	0,3106
0,11	0,0438	**0,37**	0,1443	**0,63**	0,2357	**0,89**	0,3133
0,12	0,0478	**0,38**	0,1480	**0,64**	0,2389	**0,90**	0,3159
0,13	0,0517	**0,39**	0,1517	**0,65**	0,2422	**0,91**	0,3186
0,14	0,0557	**0,40**	0,1554	**0,66**	0,2454	**0,92**	0,3212
0,15	0,0596	**0,41**	0,1591	**0,67**	0,2486	**0,93**	0,3238
0,16	0,0636	**0,42**	0,1628	**0,68**	0,2517	**0,94**	0,3264
0,17	0,0675	**0,43**	0 1664	**0,69**	0,2549	**0,95**	0,3289
0,18	0,0714	**0,44**	0,1700	**0,70**	0,2580	**0,96**	0,3315
0,19	0,0753	**0,45**	0,1736	**0,71**	0,2611	**0,97**	0,3340
0,20	0,0793	**0,46**	0,1772	**0,72**	0,2642	**0,98**	0,3365
0,21	0,0832	**0,47**	0,1808	**0,73**	0,2673	**0,99**	0,3389
0,22	0,0871	**0,48**	0,1844	**0,74**	0,2703	**1,00**	0,3413
0,23	0,0910	**0,49**	0,1879	**0,75**	0,2734	**1,01**	0,3438
0,24	0,0948	**0,50**	0,1915	**0,76**	0,2764	**1,02**	0,3461
0,25	0,0987	**0,51**	0,1950	**0,77**	0,2794	**1,03**	0,3485

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,04	0,3508	**1,43**	0,4236	**1,82**	0,4656	**2,42**	0,4922
1,05	0,3531	**1,44**	0,4251	**1,83**	0,4664	**2,44**	0,4927
1,06	0,3554	**1,45**	0,4265	**1,84**	0,4671	**2,46**	0,4931
1,07	0,3577	**1,46**	0,4279	**1,85**	0,4678	**2,48**	0,4934
1,08	0,3599	**1,47**	0,4292	**1,86**	0,4686	**2,50**	0,4938
1,09	0,3621	**1,48**	0,4306	**1,87**	0,4693	**2,52**	0,4941
1,10	0,3643	**1,49**	0,4319	**1,88**	0,4699	**2,54**	0,4945
1,11	0,3665	**1,50**	0,4332	**1,89**	0,4706	**2,56**	0,4948
1,12	0,3686	**1,51**	0,4345	**1,90**	0,4713	**2,58**	0,4951
1,13	0,3708	**1,52**	0,4357	**1,91**	0,4719	**2,60**	0,4953
1,14	0,3729	**1,53**	0,4370	**1,92**	0,4726	**2,62**	0,4956
1,15	0,3749	**1,54**	0,4382	**1,93**	0,4732	**2,64**	0,4959
1,16	0,3770	**1,55**	0,4394	**1,94**	0,4738	**2,66**	0,4961
1,17	0,3790	**1,56**	0,4406	**1,95**	0,4744	**2,68**	0,4963
1,18	0,3810	**1,57**	0,4418	**1,96**	0,4750	**2,70**	0,4965
1,19	0,3830	**1,58**	0,4429	**1,97**	0,4756	**2,72**	0,4967
1,20	0,3849	**1,59**	0,4441	**1,98**	0,4761	**2,74**	0,4969
1,21	0,3869	**1,60**	0,4452	**1,99**	0,4767	**2,76**	0,4971
1,22	0,3883	**1,61**	0,4463	**2,00**	0,4772	**2,78**	0,4973
1,23	0,3907	**1,62**	0,4474	**2,02**	0,4783	**2,80**	0,4974
1,24	0,3925	**1,63**	0,4484	**2,04**	0,4793	**2,82**	0,4976
1,25	0,3944	**1,64**	0,4495	**2,06**	0,4803	**2,84**	0,4977
1,26	0,3962	**1,65**	0,4505	**2,08**	0,4812	**2,86**	0,4979
1,27	0,3980	**1,66**	0,4515	**2,10**	0,4821	**2,88**	0,4980
1,28	0,3997	**1,67**	0,4525	**2,12**	0,4830	**2,90**	0,4981
1,29	0,4015	**1,68**	0,4535	**2,14**	0,4838	**2,92**	0,4982
1,30	0,4032	**1,69**	0,4545	**2,16**	0,4846	**2,94**	6,4984
1,31	0,4049	**1,70**	0,4554	**2,18**	0,4854	**2,96**	0,4985
1,32	0,4066	**1,71**	0,4564	**2,20**	0,4861	**2,98**	0,4986
1,33	0,4082	**1,72**	0,4573	**2,22**	0,4868	**3,00**	0,49865
1,34	0,4099	**1,73**	0,4582	**2,24**	0,4875	**3,20**	0,49931
1,35	0,4115	**1,74**	0,4591	**2,26**	0,4881	**3,40**	0,49966
1,36	0,4131	**1,75**	0,4599	**2,28**	0,4887	**3,60**	0,499841
1,37	0,4147	**1,76**	0,4608	**2,30**	0,4893	**3,80**	0,499928
1,38	0,4162	**1,77**	0,4616	**2,32**	6,4898	**4,00**	0,499968
1,39	0,4177	**1,78**	0,4625	**2,34**	0,4904	**4,50**	0,499997
1,40	0,4192	**1,79**	0,4633	**2,36**	0,4909	**5,00**	0,5
1,41	0,4207	**1,80**	0,4641	**2,38**	0,4913	**> 5**	0,5
1,42	0,4222	**1,81**	0,4649	**2,40**	0,4918	**∞**	0,5

Приложение 2.

Квантили распределения Стьюдента $t_p(k)$

k \ p	0,750	0,900	0,950	0,975	0,990	0,995	0,999
1	1,000	3,078	6,314	12,706	31,821	63,657	318
2	0,816	1,886	2,920	4,303	6,965	9,925	22,3
3	0,765	1,638	2,353	3,182	4,541	5,841	10,2
4	0,741	1,533	2,132	2,776	3,747	4,604	7,173
5	0,727	1,476	2,015	2,571	3,365	4,032	5,893
6	0,718	1,440	1,943	2,447	3,143	3,707	5,208
7	0,711	1,415	1,895	2,365	2,998	3,499	4,785
8	0,706	1,397	1,860	2,306	2,896	3,355	4,501
9	0,703	1,383	1,833	2,262	2,821	3,250	4,297
10	0,700	1,372	1,812	2,228	2,764	3,169	4,144
11	0,697	1,363	1,796	2,201	2,718	3,106	4,025
12	0,695	1,356	1,782	2,179	2,681	3,055	3,930
13	0,694	1,350	1,771	2,160	2,650	3,012	3,852
14	0,692	1,345	1,761	2,145	2,624	2,977	3,787
15	0,691	1,341	1,753	2,131	2,602	2,947	3,733
16	0,690	1,337	1,746	2,120	2,583	2,921	3,686
17	0,689	1,333	1,740	2,110	2,567	2,898	3,646
18	0,688	1,330	1,734	2,101	2,552	2,878	3,610
19	0,688	1,328	1,729	2,093	2,539	2,861	3,579
20	0,687	1,325	1,725	2,086	2,528	2,845	3,552
21	0,686	1,323	1,721	2,080	2,518	2,831	3,527
22	0,686	1,321	1,717	2,074	2,508	2,819	3,505
23	0,685	1,319	1,714	2,069	2,500	2,807	3,485
24	0,685	1,318	1,711	2,064	2,492	2,797	3,467
25	0,684	1,316	1,708	2,060	2,485	2,787	3,450
26	0,684	1,315	1,706	2,056	2,479	2,779	3,435
27	0,684	1,314	1,703	2,052	2,473	2,771	3,421
28	0,683	1,313	1,701	2,048	2,467	2,763	3,408
29	0,683	1,311	1,699	2,045	2,462	2,756	3,398
30	0,683	1,310	1,697	2,042	2,457	2,750	3,385
40	0,681	1,303	1,687	2,021	2,423	2,704	3,307
60	0,679	1,296	1,671	2,000	2,390	2,660	3,232
120	0,677	1,289	1,658	1,980	2,358	2,617	3,160
$\propto$	0,674	1,232	1,645	1,960	2,326	2,576	3,090

Приложение 3.

Квантили стандартного нормального распределения

P	u_p	P	u_p	P	u_p
0,50	0,00000	0,820	0,91536	0,995000	2,57583
0,52	0,05015	0,840	0,99446	0,996000	2,65207
0,54	0,10043	0,860	1,08032	0,997000	2,74778
0,56	0,15097	0,880	1,17499	0,998000	2,87816
0,58	0,20189	0,900	1,28155	0,999000	3,09023
0,60	0,25335	0,910	1,34075	0,999200	3,15591
0,62	0,30548	0,920	1,40507	0,999400	3,23888
0,64	0,35846	0,930	1,47579	0,999500	3,29053
0,66	0,41246	0,940	1,55477	0,999600	3,35279
0,68	0,46770	0,950	1,64485	0,999700	3,43161
0,70	0,52440	0,960	1,75069	0,999800	3,54008
0,72	0,58284	0,970	1,88079	0,999900	3,71902
0,74	0,64334	0,980	2,05375	0,999950	3,89059
0,76	0,70630	0,990	2,32635	0,999990	4,26489
0,78	0,77219	0,992	2,40891	0,999995	4,41717
0,80	0,84162	0,994	2,51214	0,999999	4,75342

Приложение 4.

Значения функции $P(\lambda) = 1 - \sum_{i=-\infty}^{+\infty} (-1)^i \cdot e^{-2i^2\lambda^2}$

λ	$P(\lambda)$	λ	$P(\lambda)$	λ	$P(\lambda)$	λ	$P(\lambda)$
0,00	1,0000	0,45	0,9874	0,90	0,3927	1,70	0,0062
0,05	1,0000	0,50	0,9639	0,95	0,3275	1,80	0,0032
0,10	1,0000	0,55	0,9228	1,00	0,2700	1,90	0,0015
0,15	1,0000	0,60	0,8643	1,10	0,1777	2,00	0,0007
0,20	1,0000	0,65	0,7920	1,20	0,1122	2,10	0,0003
0,25	1,0000	0,70	0,7112	1,30	0,0681	2,20	0,0001
0,30	1,0000	0,75	0,6272	1,40	0,0397	2,30	0,0001
0,35	0,9997	0,80	0,5441	1,50	0,0222	2,40	0,0000
0,40	0,9972	0,85	0,4653	1,60	0,0120	2,50	0,0000

Приложение 5.

Значения вероятностей для критерия χ^2

χ^2 \ r	1	2	3	4	5	6	7	8
1	0,3172	0,6065	0,8013	0,9098	0,9626	0,9856	0,9948	0,9982
2	1574	3679	5724	7358	8491	9197	9598	9810
3	0833	2231	3916	5578	7000	8088	8850	9344
4	0455	1353	2615	4060	5494	6767	7798	8571
5	0254	0821	1718	2873	4159	5438	6600	7576
6	0143	0498	1116	1991	3062	4232	5398	6472
7	0081	0302	0719	1359	2206	3208	4289	5366
8	0047	0183	0460	0916	1562	2381	3326	4335
9	0027	0111	0293	0611	1091	1736	2527	3423
10	0016	0067	0186	0404	0752	1247	1886	2650
11	0009	0041	0117	0266	0514	0884	1386	2017
12	0005	0025	0074	0174	0348	0620	1006	1512
13	0003	0015	0046	0113	0234	0430	0721	1119
14	0002	0009	0029	0073	0156	0296	0512	0818
15	0001	0006	0018	0047	0104	0203	0360	0591
16	0001	0003	0011	0030	0068	0138	0251	0424
17	0000	0002	0007	0019	0045	0093	0174	0301
18		0001	0004	0012	0029	0062	0120	0212
19		0001	0003	0008	0019	0042	0082	0149
20		0000	0002	0005	0013	0028	0056	0103
21			0001	0003	0008	0018	0038	0071
22			0001	0002	0005	0012	0025	0049
23			0000	0001	0003	0008	0017	0034
24				0001	0002	0005	0011	0023
25				0001	0001	0003	0008	0016
26				0000	0001	0002	0005	0010
27					0001	0001	0003	0007
28					0000	0001	0002	0005
29						0001	0001	0003
30						0000	0001	0002

Продолжение приложении 5.

χ^2 \ r	1	2	3	4	5	6	7	8
1	0,9994	0,9998	0,9899	1,0000	1,0000	1,0000	1,0000	1,0000
2	9915	9963	9985	0,9994	0,9998	0,9999	1,0000	1,0000
3	9643	9814	9907	9955	9979	9991	0,9896	0,9998
4	9114	9473	9699	9834	9912	9955	9977	9989
5	8343	8913	9312	9580	9752	9858	9921	9958
6	7399	8153	8734	9161	9462	9665	9797	9881
7	6371	7254	7991	8576	9022	9347	9576	9733
8	5341	6288	7133	7851	8436	8893	9238	9489
9	4373	5321	6219	7029	7729	8311	8775	9134
10	3505	4405	5304	6160	6939	7622	8197	8666
11	2757	3575	4433	5289	6108	6860	7526	8095
12	2133	2851	3626	4457	5276	6063	6790	7440
13	1626	2237	2933	3690	4478	5265	6023	6728
14	1223	1730	2330	3007	3738	4497	5255	5987
15	0909	1321	1825	2414	3074	3782	4514	5246
16	0669	0996	1411	1912	2491	3134	3821	4530
17	0487	0744	1079	1496	1993	2562	3189	3856
18	0352	0550	0816	1157	1575	2068	2627	3239
19	0252	0403	0611	0885	1231	1649	2137	2687
20	0179	0293	0453	0671	0952	1301	1719	2202
21	0126	0211	0334	0504	0729	1016	1368	1785
22	0089	0151	0244	0375	0554	0786	1078	1432
23	0062	0107	0177	0277	0417	0603	0841	1137
24	0043	0076	0127	0203	0311	0458	0651	0895
25	0030	0053	0091	0148	0231	0346	0499	0698
26	0020	0037	0065	0107	0170	0259	0380	0540
27	0014	0026	0046	0077	0154	0193	0287	0415
28	0010	0018	0032	0055	0090	0142	0216	0316
29	0006	0012	0023	0039	0065	0104	0161	0239
30	0004	0009	0016	0028	0047	0076	0119	0180

Приложение 6.

Критические значения $q_{\alpha}(n, f)$ «стьюдентизированного» размаха

n	F								
	2	5	10	15	20	30	40	60	∞
Доверительная вероятность $\alpha = 0{,}05$									
2	6,08	3,64	3,15	3,01	2,95	2,89	2,86	2,63	2,77
3	8,33	4,60	3,88	3,67	3,58	3,49	3,44	3,40	3,31
4	9,80	5,22	4,33	4,08	3,96	3,85	3,79	3,74	3,63
5	10,88	5,67	4,65	4,37	4,23	4,10	4,04	3,98	3,86
6	11,74	6,03	4,91	4,59	4,45	4,30	4,23	4,16	4,03
7	12,44	6,33	5,12	4,78	4,62	4,46	4,39	4,31	4,17
8	13,03	6,58	5,30	4,94	4,77	4,60	4,52	4,44	4,29
9	13,54	6,80	5,46	5,08	4,90	4,72	4,63	4,55	4,39
10	13,99	6,99	5,60	5,20	5,01	4,82	4,73	4,65	4,47
11	14,39	7,17	5,72	5,31	5,11	4,92	4,82	4,73	4,55
12	14,75	7,32	5,83	5,40	5,20	5,00	4,90	4,81	4,62
13	15,08	7,47	5,93	5,49	5,28	5,08	4,98	4,88	4,68
14	15,38	7,60	6,03	5,57	5,36	5,15	5,04	4,94	4,73
15	15,65	7,72	6,11	5,65	5,43	5,21	5,11	5,00	4,80
16	15,91	7,83	6,19	5,72	5,49	5,27	5,16	5,06	4,85
17	16,14	7,93	6,27	5,78	5,55	5,33	5,22	5,11	4,89
18	16,37	8,03	6,34	5,85	5,61	5,38	5,27	5,15	4,93
19	16,57	8,12	6,40	5,90	5,66	5,43	5,31	5,20	4,97
20	16,77	8,21	6,47	5,96	5,71	5,47	5,36	5,24	5,01
Доверительная вероятность $\alpha = 0{,}01$									
2	14,04	5,70	4,48	4,17	4,02	3,82	3,82	3,76	3,64
3	19,02	6,98	5,27	4,84	4,64	4,45	4,37	4,28	4,12
4	22,29	7,80	5,77	5,25,	5,02	4,80	4,70	4,59	4,40
5	24,72	8,42	6,14	5,56	5,29	5,05	4,93	4,82	4,60
6	26,63	8,91	6,43	5,80	5,51	5,24	5,11	4,99	4,76
7	28,20	9,32	6,47	5,99	5,69	5,40	5,26	5,13	4,88
8	29,53	9,67	6,84	6,16	5,84	5,54	5,39	5,25	4,99
9	30,68	9,97	7,05	6,31	5,97	5,65	5,50	5,36	5,08
10	31,69	10,24	7,21	6,44	6,09	5,76	5,60	5,45	5,16
11	32,59	10,48	7,36	6,55	6,19	5,85	5,69	5,53	5,23
12	33,40	10,70	7,49	6,66	6,28	5,93	5,76	5,60	5,29
13	34,13	10,89	7,60	6,76	6,37	6,01	5,83	5,67	5,35
14	34,61	11,08	7,71	6,84	6,45	6,08	5,90	5,73	5,40
15	35,43	11,24	7,81	6,93	6,52	6,14	5,96	5,78	5,45
16	36,00	11,40	7,91	7,00	6,59	6,20	6,02	5,84	5,49
17	36,53	11,55	7,99	7,07	6,65	6,26	6,07	5,89	5,54
18	37,03	11,68	8,08	7,14	6,71	6,31	6,12	5,93	5,57
19	37,50	11,81	8,15	7,20	6,77	6,36	6,16	5,97	5,61
20	37,93	11,93	8,23	7,26	6,82	6,41	6,21	6,01	5,65

Приложение 7.

Коэффициенты a_i K - критерия согласия Манн-Фертига-Шуера для распределения Вейбулла

n	i	a_i	n	i	a_i	n	i	a_i
3	1	1,216395	11	6	0,251386	15	9	0,180266
	2	0,863046		7	0,243928		10	0,180072
4	1	1,150727		8	0,251548		11	0,180072
	2	0,706698		9	0,283879		12	0,186347
	3	0,679596		10	0,389071		13	0,239842
5	1	1,115718	12	1	1,044137		14	0,344309
	2	0,645384		2	0,547721	16	1	1,032617
	3	0,532445		3	0,385338		2	0,534521
	4	0,583273		4	0,307221		3	0,370021
6	1	1,093929		5	0,263737		4	0,289169
	2	0,612330		6	0,238797		5	0,242049
	3	0,474330		7	0,226264		6	0,212103
	4	0,442920		8	0,224477		7	0,192338
	5	0,522759		9	0,235630		8	0,179407
7	1	1,079055		10	0,269966		9	0,171667
	2	0,591158		11	0,375356		10	0,168476
	3	0,442789	13	1	1,040515		11	0,170026
	4	0,387289		2	0,543556		12	0,177619
	5	0,387714		3	0,380417		13	0,194859
	6	0,480648		4	0,301300		14	0,232350
8	1	1,068252		5	0,256437		15	0,336283
	2	0,577339		6	0,229515	17	1	1,030618
	3	0,422889		7	0,213966		2	0,532290
	4	0,356967		8	0,207205		3	0,367507
	5	0,334089		9	0,209131		4	0,286765
	6	0,349907		10	0,222667		5	0,238765
	7	0,449338		11	0,258323		6	0,208278
9	1	1,060046		12	0,363582		7	0,187813
	2	0,566942	14	1	1,037513		8	0,173951
	3	0,409157		2	0,540059		9	0,164928
	4	0,337763		3	0,376352		10	0,159891
	5	0,304777		4	0,296496		11	0,158624
	6	0,297949		5	0,250650		12	0,161559
	7	0,322189		6	0,222377		13	0,170132
	8	0,424958		7	0,204885		14	0,188005
10	1	1,053606		8	0,195165		15	0,225729
	2	0,559013		9	0,192209		16	0,329085
	3	0,399100		10	0,196679	18	1	1,028850
	4	0,324470		11	0,211875		2	0,530332
	5	0,286163		12	0,248409		3	0,365314
	6	0,269493		13	0,353334		4	0,283846
	7	0,271645	15	1	1,034894		5	0,235958
	8	0,300869		2	0,537085		6	0,205051
	9	0,405316		3	0,372934		7	0,184055
11	1	1,048411		4	0,292518		8	0,169504
	2	0,552769		5	0,255180		9	0,159564
	3	0,391410		6	0,216712		10	0,153263
	4	0,314705		7	0,197893		11	0,150176
	5	0,273245		8	0,186266		12	0,150333

Продолжение приложения 7.

n	i	a_i	n	i	a_i	n	i	a_i
18	13	0,154313	21	10	0,140087	23	20	0,159966
	14	0,163660		11	0,134200		21	0,197679
	15	0,181971		12	0,130451		22	0,297435
	16	0,219825		13	0,128702	24	1	1,021431
	17	0,322580		14	0,129025		2	0,522233
19	1	1,027277		15	0,131756		3	0,356436
	2	0,528594		16	0,137659		4	0,274051
	3	0,363389		17	0,148341		5	0,225086
	4	0,281692		18	0,167481		6	0,192892
	5	0,233535		19	0,205352		7	0,170330
	6	0,202291		20	0,306285		8	0,153877
	7	0,180882	22	1	1,023439		9	0,141549
	8	0,165807		2	0,524405		10	0,132195
	9	0,155189		3	0,358790		11	0,125099
	10	0,147984		4	0,276618		12	0,119811
	11	0,143650		5	0,278950		13	0,116054
	12	0,142012		6	0,195983		14	0,113677
	13	0,143250		7	0,173760		15	0,112638
	14	0,148031		8	0,157692		16	0,113007
	15	0,157921		9	0,145834		17	0,114990
	16	0,176611		10	0,137052		18	0,119014
	17	0,214520		11	0,130662		19	0,125889
	18	0,316666		12	0,126260		20	0,137235
20	1	1,025866		13	0,123640		21	0,156679
	2	0,527046		14	0,122763		22	0,194285
	3	0,361682		15	0,123763		23	0,293473
	4	0,279798		16	0,127019	25	1	1,020551
	5	0,231417		17	0,133316		2	0,521285
	6	0,199905		18	0,144273		3	0,355415
	7	0,178167		19	0,163552		4	0,272945
	8	0,162684		20	0,201355		5	0,223885
	9	0,151549		21	0,301693		6	0,191578
	10	0,143674	23	1	1,022380		7	0,168899
	11	0,138448		2	0,523269		8	0,152286
	12	0,135580		3	0,357557		9	0,139789
	13	0,135306		4	0,275268		10	0,130219
	14	0,137120		5	0,226417		11	0,122871
	15	0,142527		6	0,194351		12	0,117274
	16	0,152861		7	0,171948		13	0,113132
	17	0,171810		8	0,155666		14	0,110268
	18	0,209721		9	0,143549		15	0,108598
	19	0,311257		10	0,134451		16	0,108124
21	1	1,024594		11	0,127667		17	0,108944
	2	0,525657		12	0,122768		18	0,111289
	3	0,360159		13	0,119503		19	0,115596
	4	0,278117		14	0,117764		20	0,122683
	5	0,229551		15	0,117577		21	0,134165
	6	0,197821		16	0,119120		22	0,153650
	7	0,175815		17	0,122799		23	0,191137
	8	0,160009		18	0,129416		24	0,289773
	9	0,148471		19	0,140590			

Приложение 8.

Квантили распределения χ^2 (r - число степеней свободы)

r	α							
	0,01	0,025	0,05	0,10	0,90	0,95	0,975	0,99
1	0,0001	0,0010	0,0039	0,0158	2,706	3,841	5,024	6,6350
2	0,0201	0,0506	0,1030	0,2110	4,605	5,991	7,378	9,210
3	0,1150	0,2160	0,3520	0,5840	6,251	7,815	9,348	11,345
4	0,2970	0,4840	0,7110	1,0640	7,779	9,488	11,143	13,277
5	0,5540	0,8310	1,1450	1,6100	9,236	11,070	12,832	15,086
6	0,8720	1,2370	1,6350	2,2040	10,645	12,592	14,449	16,812
7	1,2390	1,6900	2,1670	2,8330	12,017	14,067	16,013	18,475
8	1,6460	2,1800	2,7330	3,4900	13,362	15,507	17,535	20,090
9	2,0880	2,7000	3,3250	4,1680	14,684	16,919	19,023	21,666
10	2,5580	3,2470	3,9400	4,8650	15,987	18,307	20,483	23,209
11	3,0530	3,8160	4,5750	5,5780	17,275	19,675	21,920	24,725
12	3,5710	4,4040	5,2260	6,3040	18,549	21,026	23,336	26,217
13	4,1070	5,0090	5,8920	7,0420	19,812	22,362	24,736	27,688
14	4,6600	5,6290	6,5710	7,7900	21,064	23,685	26,119	29,141
15	5,2290	6,2620	7,2610	8,5470	22,307	24,996	27,488	30,578
16	5,8120	6,9080	7,9620	9,3120	23,542	26,296	28,845	32,000
17	6,4080	7,5640	8,6720	10,0850	24,769	27,587	30,191	33,409
18	7,0150	8,2310	9,3900	10,8650	25,989	28,869	31,526	34,805
19	7,6330	8,9070	10,1170	11,6510	27,204	30,144	32,852	36,191
20	8,2600	9,5910	10,8510	12,4430	28,412	31,410	34,170	37,566
21	8,8970	10,2830	11,5910	13,2400	29,615	32,671	35,479	38,232
22	9,5420	10,9820	12,3380	14,0410	30,813	33,924	36,781	40,289
23	10,1960	11,6880	13,0910	14,8480	32,007	35,172	38,076	41,638
24	10,8560	12,4010	13,8480	15,6590	33,196	36,415	39,364	42,980
25	11,5240	13,1200	14,6110	16,4730	34,382	37,652	40,646	44,314
26	12,1980	13,8440	15,3790	17,2920	35,563	38,885	41,923	45,642
27	12,8790	14,5730	16,1510	18,1140	36,741	40,113	43,194	46,963
28	13,5650	15,3080	16,9280	18,9390	37,916	41,337	44,461	48,278
29	14,2560	16,0470	17,7080	19,7680	39,087	42,557	45,722	49,588
30	14,9530	16,7910	18,4930	20,5990	40,256	43,773	46,979	50,892

Приложение 9.

Квантили распределения Фишера $F_p(k_1,k_2)$

k_2 \ k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
								p=0,9										
1	39,86	49,50	53,59	55,83	57,24	58,20	58,91	59,44	59,86	60,19	60,71	61,22	61,74	62,00	62,26	62,53	62,79	63,06
2	8,53	9,00	9,16	9,24	9,29	9,33	9,35	9,37	9,38	9,39	9,41	9,42	9,44	9,45	9,46	9,47	9,47	9,48
3	5,54	5,46	5,39	5,34	5,31	5,28	5,27	5,25	5,24	5,23	5,22	5,20	5,18	5,18	5,17	5,16	5,15	5,14
4	4,54	4,32	4,19	4,11	4,05	4,01	3,98	3,95	3,94	3,92	3,90	3,87	3,84	3,83	3,82	3,80	3,79	3,78
5	4,06	3,78	3,62	3,52	3,45	3,40	3,37	3,34	3,32	3,30	3,27	3,24	3,21	3,19	3,17	3,16	3,14	3,12
6	3,78	3,46	3,29	3,18	3,11	3,05	3,01	2,98	2,96	2,94	2,90	2,87	2,84	2,82	2,80	2,78	2,76	2,74
7	3,59	3,26	3,07	2,96	2,88	2,83	2,78	2,75	2,72	2,70	2,67	2,63	2,59	2,58	2,56	2,54	2,51	2,49
8	3,46	3,11	2,92	2,81	2,73	2,67	2,62	2,59	2,56	2,54	2,50	2,46	2,42	2,40	2,38	2,36	2,34	2,32
9	3,36	3,01	2,81	2,69	2,61	2,55	2,51	2,47	2,44	2,42	2,38	2,34	2,30	2,28	2,25	2,23	2,21	2,18
10	3,29	2,92	2,73	2,61	2,52	2,46	2,41	2,38	2,35	2,32	2,28	2,24	2,20	2,18	2,16	2,13	2,11	2,08
11	3,23	2,86	2,66	2,54	2,45	2,39	2,34	2,30	2,27	2,25	2,21	2,17	2,12	2,10	2,08	2,05	2,03	2,00
12	3,18	2,81	2,61	2,48	2,39	2,33	2,28	2,24	2,21	2,19	2,15	2,10	2,06	2,04	2,01	1,99	1,96	1,93
13	3,14	2,76	2,56	2,43	2,35	2,28	2,23	2,20	2,16	2,14	2,10	2,05	2,01	1,98	1,96	1,93	1,90	1,88
14	3,10	2,73	2,52	2,39	2,31	2,24	2,19	2,15	2,12	2,10	2,05	2,01	1,96	1,94	1,91	189	1,86	1,83
15	3,07	2,70	2,49	2,36	2,27	2,21	2,16	2,12	2,09	2,06	2,02	1,97	1,92	1,90	1,87	1,85	1,82	1,79
16	3,05	2,67	2,46	2,33	2,24	2,18	2,13	2,09	2,06	2,03	1,99	1,94	1,89	1,87	1,84	1,81	1,78	1,75
17	3,03	2,64	2,44	2,31	2,22	2,15	2,10	2,06	2,03	2,00	1,96	1,91	1,86	1,84	1,81	1,78	1,75	1,72
18	3,01	2,62	2,42	2,29	2,20	2,13	2,08	2,04	2,00	1,98	1,93	1,89	1,84	1,81	1,78	1,75	1,72	1,69
19	2,99	2,61	2,40	2,27	2,18	2,11	2,06	2,02	1,98	1,96	1,91	1,86	1,81	1,79	1,76	1,73	1,70	1,67
20	2,97	2,59	2,38	2,25	2,16	2,09	2,04	2,00	1,96	1,94	1,89	1,84	1,79	1,77	1,74	1,71	1,68	1,64
21	2,96	2,57	2,36	2,23	2,14	2,08	2,02	1,98	1,95	1,92	1,87	1,83	1,78	1,75	1,72	1,69	1,66	1,62
22	2,95	2,56	2,35	2,22	2,13	2,06	2,01	1,97	1,93	1,90	1,86	1,81	1,76	1,73	1,70	1,67	1,64	1,60
23	2,94	2,55	2,34	2,21	2,11	2,05	1,99	1,95	1,92	1,89	1,84	1,80	1,74	1,72	1,69	1,66	1,62	1,59
24	2,93	2,54	2,33	2,19	2,10	2,04	1,98	1,94	1,91	1,88	1,83	1,78	1,73	1,70	1,67	1,64	1,61	1,57
25	2,92	2,53	2,32	2,18	2,09	2,02	1,97	1,93	1,89	1,87	1,82	1,77	1,72	1,69	1,66	1,63	1,59	1,56
26	2,91	2,52	2,31	2,17	2,08	2,01	1,96	1,92	1,88	1,86	1,81	1,76	1,71	1,68	1,65	1,61	1,58	1,54
27	2,90	2,51	2,30	2,17	2,07	2,00	1,95	1,91	1,87	1,85	1,80	1,75	1,70	1,67	1,64	1,60	1,57	1,53
28	2,89	2,50	2,29	2,16	2,06	2,00	1,94	1,90	1,87	1,84	1,79	1,74	1,69	1,66	1,63	1,59	1,56	1,52
29	2,89	2,50	2,28	2,15	2,06	1,99	1,93	1,89	1,86	1,83	1,78	1,73	1,68	1,65	1,62	1,58	1,55	1,51
30	2,88	2,49	2,28	2,14	2,05	1,98	1,93	1,88	1,85	1,82	1,77	1,72	1,67	1,64	1,61	1,57	1,54	1,50
40	2,84	2,44	2,23	2,09	2,00	1,93	1,87	1,83	1,79	1,76	1,71	1,66	1,61	1,57	1,54	1,51	1,47	1,42
60	2,79	2,39	2,18	2,04	1,95	1,87	1,82	1,77	1,74	1,71	1,66	1,60	1,54	1,51	1,48	1,44	1,40	1,35
120	2,75	2,35	2,13	1,99	1,90	1,82	1,77	1,72	1,68	1,65	1,60	1,55	1,48	1,45	1,41	1,37	1,32	1,26
∝	2,71	2,30	2,08	1,94	1,85	1,77	1,72	1,67	1,63	1,60	1,55	1,49	1,42	1,38	1,34	1,30	1,24	1,17

Продолжение приложения 9.

k_1 / k_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
								p=0,95										
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,90	240,5	241,9	243,9	245,9	248,0	246,1	250,1	251,1	252,2	252,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,49
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	43,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,62	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,7	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06
17	4,45	3,59	3,230	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93
20	4,35	3,49	3,19	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	,15	2,07	2,03	1,98	1,94	1,89	1,84
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,07	1,99	1,95	1,90	1,85	1,80	1,75
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,71
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,03	1,94	1,90	1,5	1,81	1,75	1,70
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35
∝	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22

Продолжение приложения 9.

k_1 / k_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
								p=0,975										
1	647,8	799,5	864,2	899,6	921,8	937,1	948,2	956,7	963,3	968,6	976,7	984,9	993,1	997,2	1001	1006	1010	1014
2	38,51	39,00	39,17	39,25	39,30	39,33	39,36	39,37	39,39	39,40	39,41	39,43	39,45	39,46	39,46	39,47	39,48	39,49
3	17,44	16,04	15,44	15,10	14,88	14,73	14,62	14,54	14,47	14,42	14,34	14,25	14,17	14,12	14,08	14,04	13,99	13,95
4	12,22	10,65	9,98	9,60	9,36	9,20	9,07	8,98	8,90	8,84	8,75	8,66	8,56	8,51	8,46	8,41	8,36	8,31
5	10,01	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,68	6,62	6,52	6,43	6,33	6,28	6,23	6,18	6,12	6,07
6	8,81	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,52	5,46	5,37	5,27	5,17	5,12	5,07	5,01	4,96	4,90
7	8,07	6,54	5,89	5,52	5,29	5,12	4,99	4,90	4,82	4,76	4,67	4,57	4,47	4,42	4,36	4,31	4,25	4,20
8	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,36	4,30	4,20	4,10	4,00	3,95	3,89	3,84	3,78	3,73
9	7,21	5,71	5,08	4,72	4,48	4,32	4,20	4,10	4,03	3,96	3,87	3,77	3,67	3,61	3,56	3,51	3,45	3,39
10	6,94	5,46	4,83	4,47	4,24	4,07	3,95	3,85	3,78	3,72	3,62	3,52	3,42	3,37	3,31	3,26	3,20	3,14
11	6,72	5,26	4,63	4,28	4,04	3,88	3,76	3,66	3,59	3,53	3,43	3,33	3,23	3,17	3,12	3,06	3,00	2,94
12	6,55	5,10	4,47	4,12	3,89	3,73	3,61	3,51	3,44	3,37	3,28	3,18	3,07	3,02	2,96	2,91	2,85	2,79
13	6,41	4,97	4,35	4,00	3,77	3,60	3,48	3,39	3,31	3,25	3,15	3,05	2,95	2,89	2,84	2,78	2,72	2,66
14	6,30	4,86	4,24	3,89	3,66	3,50	3,38	3,29	3,21	3,15	3,05	2,95	2,84	2,79	2,73	2,67	2,61	2,55
15	6,30	4,77	4,15	3,80	3,58	3,41	3,29	3,20	3,12	3,06	2,96	2,86	2,76	2,70	2,64	2,59	2,52	2,46
16	6,12	4,69	4,08	3,73	3,50	3,34	3,22	3,12	3,05	2,99	2,89	2,79	2,68	2,63	2,57	2,51	2,45	2,38
17	6,04	4,62	4,01	3,66	3,44	3,28	3,16	3,06	2,96	2,92	2,82	2,72	2,62	2,56	2,50	2,44	2,38	2,32
18	5,98	4,56	3,95	3,61	3,38	3,22	3,10	3,01	2,93	2,84	2,77	2,67	2,56	2,50	2,44	2,38	2,32	2,26
19	5,92	4,51	3,90	3,56	3,33	3,17	3,05	2,69	2,88	2,82	2,72	2,62	2,51	2,45	2,39	2,33	2,27	2,20
20	5,87	4,46	3,86	3,51	3,29	3,13	3,01	2,91	2,84	2,77	2,68	2,57	2,46	2,41	2,35	2,29	2,22	2,16
21	5,83	4,42	3,82	3,48	3,25	3,09	2,97	2,87	2,80	2,73	2,64	2,53	2,42	2,37	2,31	2,25	2,18	2,11
22	5,79	4,38	3,78	3,44	3,22	3,05	2,93	2,84	2,76	2,70	2,60	2,50	2,39	2,33	2,27	2,21	2,14	2,08
23	5,75	4,35	3,75	3,41	3,18	3,02	2,90	2,81	2,73	2,67	2,57	2,47	2,36	2,30	2,24	2,18	2,11	2,04
24	5,72	4,32	3,72	3,38	3,15	2,99	2,87	2,78	2,70	2,64	2,54	2,44	2,33	2,27	2,21	2,15	2,08	2,01
25	5,69	4,29	3,69	3,35	3,13	2,97	2,85	2,75	2,68	2,61	2,51	2,41	2,30	2,24	2,18	2,12	2,05	1,98
26	5,66	4,27	3,67	3,33	3,10	2,94	2,82	2,73	2,65	2,59	2,49	2,39	2,28	2,22	2,16	2,09	2,03	1,95
27	5,63	4,24	3,65	3,31	3,08	2,92	2,80	2,71	2,63	2,57	2,47	2,36	2,25	2,19	2,13	2,07	2,00	1,93
28	5,61	4,22	3,63	3,29	3,06	2,90	2,78	2,69	2,61	2,55	2,45	2,34	2,23	2,17	2,11	2,05	1,98	1,91
29	5,59	4,20	3,61	3,27	3,04	2,88	2,76	2,67	2,59	2,53	2,43	2,32	2,21	2,15	2,09	2,03	1,96	1,89
30	5,57	4,18	3,53	3,25	3,03	2,87	2,75	2,65	2,57	2,51	2,41	2,31	2,20	2,14	2,07	2,01	1,94	1,87
40	5,42	4,05	3,46	3,13	2,90	2,74	2,62	2,53	2,45	2,39	2,29	2,18	2,07	2,01	1,94	1,88	1,80	1,72
60	5,29	3,93	3,34	3,01	2,79	2,63	2,51	2,41	2,33	2,27	2,17	2,06	1,94	1,88	1,82	1,74	1,67	1,58
120	5,15	3,80	3,23	2,89	2,67	2,52	2,39	2,30	2,22	2,16	2,05	1,94	1,82	1,76	1,69	1,61	1,53	1,43
$\propto$	5,02	3,69	3,12	2,79	2,57	2,41	2,29	2,19	2,11	2,05	1,94	1,83	1,71	1,64	1,57	1,48	1,39	1,27

Продолжение приложения 9.

k_1 / k_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
								p=0,99										
1	4052	4999,5	5403	5626	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6267	6313	6339
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,42	99,43	99,45	99,46	99,47	99,47	99,48	99,49
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	27,05	26,87	26,69	26,60	26,50	26,41	26,32	26,22
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,37	14,20	14,02	13,93	13,84	13,75	13,65	13,56
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,89	9,72	9,55	9,47	9,38	9,29	9,20	9,11
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,72	7,56	7,40	7,31	7,23	7,14	7,06	6,97
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,47	6,31	6,16	6,07	5,99	5,91	5,82	5,74
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,67	5,52	5,36	5,28	5,20	5,12	5,03	4,95
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,11	4,96	4,81	4,73	4,65	4,57	4,48	4,40
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,71	4,56	4,41	4,33	4,25	4,17	4,08	4,00
11	9,56	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,40	4,23	4,10	4,02	3,94	3,86	3,78	3,69
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,16	4,01	3,86	3,78	3,70	3,62	3,54	3,45
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,96	3,82	3,66	3,59	3,51	3,43	3,34	3,25
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,80	3,66	3,51	3,43	3,35	3,27	3,18	3,09
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,67	3,52	3,37	3,29	3,21	3,13	3,05	2,96
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,55	3,41	3,26	3,18	3,10	3,02	2,93	2,84
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,46	3,31	3,16	3,08	3,00	2,92	2,83	2,75
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,37	3,23	3,08	3,00	2,92	2,84	2,75	2,66
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,30	3,15	3,00	2,92	2,84	2,76	2,67	2,58
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,23	3,09	2,94	2,86	2,78	2,69	2,61	2,52
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,17	3,03	2,88	2,80	2,72	2,64	2,55	2,46
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,12	2,98	2,83	2,75	2,67	2,58	2,50	2,40
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,07	2,93	2,78	2,70	2,62	2,54	2,45	2,35
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,03	2,89	2,74	2,66	2,58	2,49	2,40	2,31
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	2,99	2,85	2,70	2,62	2,54	2,45	2,36	2,27
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	2,96	2,81	2,66	2,58	2,50	2,42	2,33	2,23
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06	2,93	2,78	2,63	2,55	2,47	2,38	2,29	2,20
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,90	2,75	2,60	2,52	2,44	2,35	2,26	2,17
29	7,60	5,42	7,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,87	2,73	2,57	2,49	2,41	2,33	2,23	2,14
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,84	2,70	2,55	2,47	2,39	2,30	2,21	2,11
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,66	2,52	2,37	2,29	2,20	2,11	2,02	1,92
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,50	2,35	2,20	2,12	2,03	1,94	1,84	1,73
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,34	2,19	2,03	1,95	1,86	1,76	1,86	1,53
∝	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,18	2,04	1,88	1,79	1,70	1,59	1,47	1,32

k_2 \ k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
								p=0,995										
1	16211	20000	21615	22500	23056	23437	23715	23925	24091	24224	24246	24630	24836	24940	25044	25148	25253	
2	198,5	199,0	199,2	199,2	199,3	199,3	199,4	199,4	199,4	199,4	199,4	199,4	199,4	199,5	199,5	199,5	199,5	
3	55,55	49,80	47,47	46,19	45,39	44,84	44,43	44,13	43,88	43,69	43,39	43,08	42,78	42,62	42,47	42,31	42,15	
4	31,33	26,28	24,26	23,15	22,46	21,97	21,62	21,35	21,14	20,97	20,70	20,44	20,17	20,03	19,89	19,75	19,61	
5	22,78	18,31	16,53	15,56	14,94	14,51	14,20	13,96	13,77	13,62	13,38	13,15	12,90	12,78	12,66	12,53	12,40	
6	18,63	14,54	12,92	12,03	11,46	11,07	10,79	10,57	10,39	10,25	10,03	9,81	9,59	9,47	9,36	9,24	9,12	
7	16,24	12,40	10,88	10,05	9,52	9,16	8,89	8,68	8,51	8,38	8,18	7,97	7,75	7,65	7,53	7,42	7,31	
8	14,69	11,04	9,60	8,81	8,30	7,95	7,69	7,50	7,30	7,21	7,01	6,81	6,61	6,50	6,40	6,29	6,18	
9	13,61	10,11	8,72	7,96	7,47	7,13	6,88	6,69	6,54	6,42	6,23	6,03	5,83	5,73	5,62	5,52	5,41	
10	12,83	9,43	8,08	7,34	6,87	6,54	6,30	6,12	5,97	5,85	5,66	5,47	5,27	5,17	5,02	4,97	4,86	
11	12,23	8,91	7,60	6,88	6,42	6,10	5,86	5,68	5,54	5,42	5,24	5,05	4,86	4,76	4,65	4,55	4,44	
12	11,75	8,51	7,23	6,52	6,07	5,76	5,52	5,35	5,20	5,09	4,91	4,72	4,53	4,43	4,33	4,23	4,12	
13	11,37	8,19	6,93	6,23	5,79	5,48	5,25	5,08	4,94	4,82	4,64	4,46	4,27	4,17	4,07	3,97	3,87	
14	11,06	7,92	6,68	6,00	5,56	5,26	5,03	4,86	4,72	4,60	4,43	4,25	4,06	3,96	3,86	3,76	3,66	
15	10,80	7,70	6,48	5,80	5,37	5,07	4,85	4,67	4,54	4,42	4,25	4,07	3,88	3,79	3,69	3,58	3,48	
16	10,58	7,51	6,30	5,64	5,21	4,91	4,69	4,52	4,38	4,27	4,10	3,92	3,73	3,64	3,54	3,44	3,33	
17	10,38	7,35	6,16	5,50	5,07	4,78	4,56	4,39	4,25	4,14	3,97	3,79	3,61	3,51	3,41	3,31	3,21	
18	10,22	7,21	6,03	5,37	4,96	4,66	4,44	4,28	4,14	4,03	3,86	3,68	3,50	3,40	3,30	3,20	3,10	
19	10,07	7,09	5,92	5,27	4,85	4,56	4,34	4,18	4,04	3,93	3,76	3,59	3,40	3,31	3,21	3,11	3,00	
20	9,94	6,99	5,82	5,17	4,76	4,47	4,26	4,09	3,96	3,85	3,68	3,50	3,32	3,22	3,12	3,02	2,92	
21	9,83	6,89	5,73	5,09	4,68	4,39	4,18	4,01	3,88	3,77	3,60	3,43	3,24	3,15	3,05	2,95	2,84	
22	9,73	6,81	5,65	5,02	4,61	4,32	4,11	3,94	3,81	3,70	3,54	3,36	3,18	3,08	2,98	2,88	2,77	
23	9,63	6,73	5,58	4,95	4,54	4,26	4,05	3,88	3,75	3,64	3,47	3,30	3,12	3,02	2,92	2,82	2,71	
24	9,55	6,66	5,52	4,89	4,49	4,20	3,99	3,83	3,69	3,59	3,42	3,25	3,06	2,97	2,87	2,77	2,66	
25	9,48	6,60	5,46	4,84	4,43	4,15	3,94	3,78	3,64	3,54	3,37	3,20	3,01	2,92	2,82	2,72	2,61	
26	9,41	6,54	5,41	4,79	4,38	4,10	3,89	3,73	3,60	3,49	3,33	3,15	2,97	2,87	2,77	2,67	2,56	
27	9,34	6,49	5,36	4,74	4,34	4,06	3,85	3,69	3,56	3,45	3,28	3,11	2,93	2,83	2,73	2,63	2,52	
28	9,28	6,44	5,32	4,70	4,30	4,02	3,81	3,65	3,52	3,41	3,25	3,07	2,89	2,79	2,69	2,59	2,48	
29	9,23	6,40	5,28	4,66	4,26	3,98	3,77	3,61	3,48	3,38	3,21	3,04	2,86	2,76	2,66	2,56	2,45	
30	9,18	6,35	5,24	4,62	4,23	3,95	3,74	3,58	3,45	3,34	3,18	3,01	2,82	2,73	2,63	2,52	2,42	
40	8,83	6,07	4,98	4,37	3,99	3,71	3,51	3,35	3,22	3,12	2,95	2,78	2,60	2,50	2,40	2,30	2,18	
60	8,49	5,79	4,73	4,14	3,76	3,49	3,29	3,13	3,01	2,90	2,74	2,57	2,39	2,29	2,19	2,08	1,96	
120	8,18	5,54	4,50	3,92	3,55	3,28	3,09	2,93	2,81	2,71	2,54	2,37	2,19	2,09	1,98	1,87	1,75	
∝	7,88	5,30	4,28	3,72	3,35	3,09	2,90	2,74	2,62	2,52	2,36	2,19	2,00	1,90	1,79	1,67	1,53	

Приложение 10.

Критические значения $K_{\alpha}(n,r)$ критерия Манн-Фертинга-Шуера (α-доверительная вероятность)

n	r	α			N	r	α			n	r	α			N	r	α		
		0,90	0,95	0,99			0,90	0,95	0,99			0,90	0,95	0,99			0,90	0,95	0,99
3	3	0,90	0,95	0,99	10	9	0,71	0,76	0,85	14	5	0,80	0,86	0,94	16	14	0,63	0,67	0,75
4	3	0,90	0,95	0,99		10	0,64	0,69	0,79		6	0,68	0,74	0,86		15	0,66	0,70	0,77
	4	0,67	0,76	0,89	11	3	0,90	0,95	0,99		7	0,75	0,81	0,89		16	0,62	0,66	0,73
5	3	0,90	0,95	0,99		4	0,68	0,77	0,90		8	0,66	0,73	0,82	17	3	0,90	0,95	0,99
	4	0,68	0,77	0,89		5	0,80	0,86	0,94		9	0,72	0,77	0,85		4	0,69	0,78	0,90
	5	0,79	0,96	0,94		6	0,68	0,75	0,86		10	0,65	0,70	0,79		5	0,80	0,87	0,94
6	3	0,90	0,95	0,99		7	0,75	0,81	0,89		11	0,69	0,74	0,82		6	0,68	0,74	0,85
	4	0,68	0,76	0,89		8	0,66	0,72	0,82		12	0,64	0,68	0,77		7	0,75	0,80	0,89
	5	0,80	0,86	0,94		9	0,71	0,77	0,85		13	0,67	0,72	0,79		8	0,66	0,72	0,81
	6	0,66	0,73	0,84		10	0,64	0,70	0,79		14	0,62	0,67	0,75		9	0,72	0,77	0,85
7	3	0,90	0,95	0,99		11	0,69	0,74	0,82	15	3	0,90	0,95	0,99		10	0,65	0,70	0,79
	4	0,68	0,77	0,89	12	3	0,90	0,95	0,99		4	0,69	0,78	0,90		11	0,69	0,74	0,82
	5	0,80	0,86	0,94		4	0,68	0,78	0,89		5	0,80	0,86	0,94		12	0,64	0,69	0,77
	6	0,67	0,74	0,85		5	0,80	0,86	0,94		6	0,67	0,75	0,86		13	0,67	0,72	0,80
	7	0,74	0,80	0,80		6	0,67	0,74	0,85		7	0,75	0,81	0,89		14	0,63	0,68	0,75
8	3	0,90	0,95	0,99		7	0,75	0,81	0,89		8	0,66	0,72	0,82		15	0,66	0,70	0,77
	4	0,68	0,77	0,90		8	0,66	0,72	0,82		9	0,72	0,77	0,85		16	0,62	0,66	0,74
	5	0,80	0,88	0,94		9	0,71	0,77	0,85		10	0,65	0,70	0,79		17	0,65	0,69	0,75
	6	0,67	0,74	0,85		10	0,65	0,70	0,79		11	0,69	0,74	0,82	18	3	0,90	0,95	0,99
	7	0,74	0,80	0,89		11	0,69	0,74	0,82		12	0,64	0,68	0,77		4	0,68	0,77	0,90
	8	0,65	0,71	0,81		12	0,63	0,68	0,76		13	0,67	0,72	0,79		5	0,80	0,86	0,94
9	3	0,90	0,95	0,99	13	3	0,90	0,95	0,99		14	0,63	0,67	0,75		6	0,67	0,75	0,86
	4	0,68	0,77	0,89		4	0,68	0,77	0,89		15	0,66	0,70	0,77		7	0,75	0,81	0,89
	5	0,80	0,86	0,94		5	0,80	0,86	0,94	16	3	0,90	0,95	0,99		8	0,66	0,73	0,82
	6	0,67	0,75	0,86		6	0,68	0,75	0,86		4	0,69	0,78	0,89		9	0,72	0,77	0,85
	7	0,74	0,80	0,89		7	0,75	0,81	0,90		5	0,80	0,86	0,94		10	0,65	0,71	0,80
	8	0,66	0,72	0,82		8	0,66	0,72	0,82		6	0,68	0,75	0,86		11	0,69	0,74	0,82
	9	0,71	0,76	0,85		9	0,72	0,77	0,85		7	0,75	0,81	0,89		12	0,64	0,69	0,77
10	3	0,90	0,95	0,99		10	0,65	0,70	0,79		8	0,66	0,72	0,82		13	0,68	0,72	0,80
	4	0,68	0,77	0,90		11	0,69	0,74	0,82		9	0,72	0,77	0,85		14	0,63	0,67	0,76
	5	0,80	0,86	0,94		12	0,64	0,68	0,76		10	0,65	0,71	0,79		15	0,66	0,70	0,78
	6	0,68	0,75	0,85		13	0,67	0,72	0,79		11	0,69	0,74	0,82		16	0,62	0,66	0,74
	7	0,75	0,81	0,89	14	3	0,90	0,95	0,99		12	0,64	0,69	0,77		17	0,65	0,69	0,76
	8	0,66	0,72	0,81		4	0,68	0,77	0,89		13	0,68	0,72	0,80		18	0,62	0,65	0,71

Продолжение приложения 10.

n	r	α			N	r	α			n	r	α			N	r	α		
		0,90	0,95	0,99			0,90	0,95	0,99			0,90	0,95	0,99			0,90	0,95	0,99
19	3	0,90	0,95	0,99	21	3	0,90	0,95	0,99	22	19	0,64	0,68	0,75	24	13	0,68	0,73	0,80
	4	0,69	0,78	0,90		4	0,69	0,78	0,90		20	0,61	0,65	0,72		14	0,64	0,68	0,76
	5	0,81	0,86	0,94		5	0,80	0,86	0,94		21	0,64	0,67	0,73		15	0,67	0,71	0,78
	6	0,68	0,75	0,86		6	0,67	0,74	0,85		22	0,61	0,64	0,70		16	0,63	0,67	0,74
	7	0,75	0,81	0,89		7	0,75	0,80	0,89	23	3	0,90	0,95	0,99		17	0,65	0,69	0,76
	8	0,67	0,72	0,82		8	0,66	0,73	0,82		4	0,68	0,77	0,89		18	0,62	0,66	0,73
	9	0,72	0,77	0,85		9	0,72	0,77	0,85		5	0,80	0,86	0,94		19	0,64	0,68	0,75
	10	0,65	0,71	0,80		10	0,65	0,70	0,80		6	0,68	0,76	0,86		20	0,61	0,65	0,72
	11	0,69	0,74	0,82		11	0,69	0,74	0,82		7	0,76	0,82	0,89		21	0,64	0,67	0,73
	12	0,64	0,69	0,77		12	0,64	0,69	0,77		8	0,67	0,73	0,83		22	0,61	0,64	0,71
	13	0,68	0,72	0,80		13	0,68	0,72	0,79		9	0,72	0,78	0,86		23	0,63	0,66	0,72
	14	0,63	0,68	0,76		14	0,63	0,67	0,75		10	0,66	0,71	0,80		24	0,60	0,64	0,69
	15	0,66	0,70	0,78		15	0,66	0,70	0,78		11	0,70	0,75	0,82	25	3	0,90	0,95	0,99
	16	0,62	0,66	0,74		16	0,63	0,67	0,74		12	0,64	0,69	0,78		4	0,69	0,78	0,91
	17	0,65	0,69	0,76		17	0,65	0,69	0,76		13	0,68	0,73	0,80		5	0,81	0,86	0,94
	18	0,61	0,65	0,71		18	0,62	0,66	0,73		14	0,63	0,68	0,76		6	0,68	0,75	0,86
	19	0,64	0,67	0,74		19	0,64	0,68	0,75		15	0,67	0,71	0,78		7	0,75	0,81	0,89
20	3	0,90	0,95	0,99		20	0,61	0,65	0,72		16	0,63	0,67	0,75		8	0,66	0,72	0,82
	4	0,68	0,78	0,90		21	0,63	0,67	0,73		17	0,65	0,69	0,77		9	0,72	0,77	0,85
	5	0,80	0,86	0,91	22	3	0,90	0,95	0,99		18	0,62	0,66	0,73		10	0,65	0,70	0,80
	6	0,67	0,75	0,86		4	0,68	0,77	0,90		19	0,64	0,68	0,75		11	0,70	0,75	0,82
	7	0,75	0,81	0,89		5	0,80	0,86	0,94		20	0,61	0,65	0,72		12	0,64	0,69	0,78
	8	0,66	0,73	0,82		6	0,68	0,75	0,85		21	0,63	0,67	0,73		13	0,68	0,73	0,81
	9	0,72	0,77	0,85		7	0,75	0,81	0,89		22	0,60	0,64	0,70		14	0,63	0,68	0,76
	10	0,65	0,71	0,80		8	0,66	0,72	0,82		23	0,63	0,66	0,72		15	0,66	0,71	0,78
	11	0,69	0,74	0,83		9	0,72	0,77	0,85	24	3	0,90	0,95	0,99		16	0,63	0,67	0,74
	12	0,64	0,69	0,77		10	0,65	0,70	0,80		4	0,69	0,78	0,90		17	0,65	0,69	0,76
	13	0,68	0,72	0,80		11	0,69	0,74	0,82		5	0,81	0,86	0,94		18	0,62	0,66	0,73
	14	0,63	0,68	0,76		12	0,64	0,69	0,78		6	0,68	0,75	0,85		19	0,64	0,68	0,75
	15	0,66	0,71	0,78		13	0,68	0,72	0,80		7	0,75	0,81	0,89		20	0,61	0,65	0,72
	16	0,62	0,67	0,74		14	0,63	0,68	0,75		8	0,67	0,73	0,83		21	0,63	0,67	0,74
	17	0,65	0,69	0,76		15	0,67	0,71	0,78		9	0,72	0,77	0,86		22	0,61	0,64	0,71
	18	0,62	0,66	0,72		16	0,62	0,67	0,74		10	0,66	0,71	0,80		23	0,63	0,66	0,72
	19	0,64	0,68	0,74		17	0,65	0,69	0,76		11	0,70	0,75	0,83		24	0,60	0,63	0,70
	20	0,61	0,65	0,71		18	0,62	0,66	0,73		12	0,64	0,70	0,78		25	0,62	0,65	0,71

Приложение 11.

Коэффициенты для вычисления границ регулирования контрольных карт Шухарта

Объём выборки, n	$\bar{x}$ – карта	R- карта		S- карта		
	A	D_1	D_2	B_1	B_2	C
2	1,880	0	3,267	0	2,298	0,798
3	1,023	0	2,575	0	2,111	0,886
4	0,729	0	2,282	0	1,982	0,921
5	0,577	0	2,115	0	1,889	0,940
6	0,483	0	2,004	0,085	1,817	0,951
7	0,419	0,076	1,924	0,158	1,762	0,960
8	0,373	0,136	1,864	0,215	1,715	0,965
9	0,337	0,184	1,816	0,262	1,676	0,969
10	0,308	0,223	1,777	0,302	1,644	0,973
11	0,285	0,256	1,744	0,336	1,616	0,976
12	0,266	0,284	1,716	0,365	1,589	0,977
13	0,249	0,308	1,692	0,392	1,568	0,980
14	0,235	0,329	1,671	0,414	1,548	0,981
15	0,223	0,348	1,652	0,434	1,530	0,982
16	0,212	0,364	1,636	0,454	1,514	0,984
17	0,203	0,379	1,621	0,469	1,499	0,984
18	0,194	0,392	1,608	0,486	1,486	0,986
19	0,187	0,404	1,596	0,500	1,472	0,986
20	0,180	0,414	1,586	0,513	1,461	0,987
21	0,173	0,425	1,575	0,525	1,451	0,988
22	0,167	0,434	1,566	0,536	1,440	0,988
23	0,162	0,443	1,557	0,546	1,432	0,989
24	0,157	0,452	1,548	0,556	1,422	0,989
25	0,153	0,459	1,541	0,566	1,414	0,990

Приложение 12.

1. Гамма-функция. *Гамма-функцией(* или *интегралом Эйлера второго рода)* называется интеграл вида $\Gamma(p)=\int\limits_0^{\infty} e^{-x}x^{p-1}dx$. (1)

Интеграл (1)-функция параметра *p*-является несобственным, так как верхний предел равен бесконечности и, кроме того, при $x \to 0$ и *p*<1 подынтегральная функция неограниченно возрастает. Интеграл (1) сходится при *p*>0 и расходится при $p \le 1$. Гамма-функция является одной из важнейших (после элементарных) функций для математики и её приложений.

Основные свойства гамма-функции

1^0.Функция $\Gamma(p)$ непрерывна и имеет непрерывную производную $\Gamma'(p)$ для *p*>0 .

2^0. Имеет место равенство $\Gamma(p+1)=p\Gamma(p)$ (2)

3^0. После n-кратного применения формулы (2) получается соотношение

$$\Gamma(p+n)=(p+n-1)(p+n-2)...(p+1)p\cdot\Gamma(p) \qquad (3)$$

4^0. Если в формуле(3) положить *p*=1 и учесть, что $\Gamma(1)=\int\limits_0^{\infty} e^{-x}dx=1$, то получится равенство $\Gamma(\text{т}+1)=n!$. (4)

Если $n=0$, то $0!=\Gamma(1)=1$.

5^0. Функция $\Gamma(\text{p})$ дает возможность распространить понятие факториала $n!$, определенного лишь для натуральных значений n, на область любых положительных значений аргумента. Из формулы (2) следует, что если $p \to 0$, то $\Gamma(\text{p})=\dfrac{\Gamma(\text{p}+1)}{\text{p}} \to +\infty$, т.е. $\Gamma(0)=+\infty$.

6^0. При $p=-n$ из формулы (2) следует, что

$$\Gamma(-\text{n})=\frac{\Gamma(-\text{n}+1)}{-n}=\frac{\Gamma(-\text{n}+2)}{n(n-1)}=-\frac{\Gamma(-\text{n}+3)}{n(n-1)(n-2)}=...=(-1)^n\frac{\Gamma(0)}{n!}=(-1)^n\cdot\infty,$$

т.е. $\Gamma(-\text{n})=(-1)^n\cdot\infty\,(n=1,2,3,...)$.

7^0. Вообще, функцию $\Gamma(\text{p})$ можно распространить на случай отрицательных значений аргумента p. Так как $\Gamma(\text{p})=\dfrac{\Gamma(\text{p}+1)}{\text{p}}$, то $\Gamma(\text{p}+1)$ имеет смысл при $-1<p<0$.

Если $-n<p<-(n-1)$, то из формулы (3) следует, что

$$\Gamma(\text{p})=\frac{\Gamma(\text{p}+\text{n})}{\text{p}(\text{p}+1)(\text{p}+2)...(\text{p}+n-1)}.$$

С помощью подстановки $p+n=a$, откуда $p=-n+a$, последняя формула преобразуется к виду $\Gamma(\alpha-\text{n})=\dfrac{(-1)^{\text{n}}\Gamma(\alpha)}{(1-\alpha)(2-\alpha)...(n-\alpha)}$ (5)

и для $-n<p<-(n-1)$ знак $\Gamma(\text{p})$ определяется множителем $(-1)^n$.

8^0. Используя формулу (2), можно получить значения $\Gamma(\text{p})$ для полуцелого аргумента:

$$\Gamma\left(m+\frac{1}{2}\right)=\Gamma\left[1+\left(m-\frac{1}{2}\right)\right]=\left(m-\frac{1}{2}\right)\cdot\Gamma\left(m-\frac{1}{2}\right)=\left(m-\frac{1}{2}\right)\left(m-\frac{3}{2}\right)\cdot\Gamma\left(m-\frac{3}{2}\right)=\ldots=$$

$$=\left(m-\frac{1}{2}\right)\left(m-\frac{3}{2}\right)\ldots\frac{5}{2}\cdot\frac{3}{2}\cdot\frac{1}{2}\cdot\Gamma\left(\frac{1}{2}\right)=\frac{(2m-1)(2m-3)\ldots5\cdot3\cdot1}{2^m}\Gamma\left(\frac{1}{2}\right),$$

или $\Gamma\left(m+\frac{1}{2}\right)=\frac{(2m-1)!!}{2^m}\cdot\Gamma\left(\frac{1}{2}\right)=\frac{(2m)!}{m!2^{2m}}\Gamma\left(\frac{1}{2}\right).$ (6)

9^0. Имеет место *формула дополнения*

$$\Gamma(p)\cdot\Gamma(1-p)=\frac{\pi}{\sin p\pi}\quad(0<p<1). \qquad (7)$$

Если в этой формуле положить $p=1/2$, то $\left[\Gamma\left(1/2\right)\right]^2=\pi/\sin(\pi/2)=\pi$, т.е. $\Gamma\left(1/2\right)=\sqrt{\pi}$.

Пользуясь основными свойствами, можно вычислить $\Gamma(p)$ для любого p .

Значения гамма-функции $\Gamma(p)$ (при $1\le p\le 2$)

p	$\Gamma(p)$	p	$\Gamma(p)$	p	$\Gamma(p)$	P	$\Gamma(p)$
1,00	1,0000	1,25	0,9064	1,50	0,8862	1,75	0,9191
1,01	0,9943	1,26	9044	1,51	8866	1,76	9214
1,02	9888	1,27	9025	1,52	8870	1,77	9238
1,03	9835	1,28	9007	1,53	8876	1,78	9262
1,04	9784	1,29	8990	1,54	8882	1,79	9288
1,05	9735	1,30	8975	1,55	8889	1,80	9314
1,06	9687	1,31	8960	1,56	8896	1,81	9341
1,07	9642	1,32	8946	1,57	8905	1,82	9368
1,08	9597	1,33	8934	1,58	8914	1,83	9397
1,09	9555	1,34	8922	1,59	8924	1,84	9426
1,10	9514	1,35	8912	1,60	8935	1,85	9456
1,11	9474	1,36	8902	1,61	8947	1,86	9487
1,12	9436	1,37	8893	1,62	8959	1,87	9518
1,13	9399	1,38	8885	1,63	8972	1,88	9551
1,14	9364	1,39	8879	1,64	8986	1,89	9584
1,15	9330	1,40	8873	1,65	9001	1,90	9618
1,16	9298	1,41	8868	1,66	9017	1,91	9652
1,17	9267	1,42	8864	1,67	9033	1,92	9688
1,18	9237	1,43	8860	1,68	9050	1,93	9724
1,19	9209	1,44	8858	1,69	9068	1,94	9761
1,20	9182	1,45	8857	1,70	9086	1,95	9799
1,21	9156	1,46	8856	1,71	9106	1,96	9837
1,22	9131	1,47	8856	1,72	9126	1,97	9877
1,23	9108	1,48	8857	1,73	9147	1,98	9917
1,24	9085	1,49	8859	1,74	9168	1,99	9958
						2,00	1,0000

2. Бета-функция. *Бета - функцией* (или *интегралом Эйлера первого рода)* называется интеграл $B(\rho;q)=\int\limits_0^1 x^{p-1}(1-x)^{q-1}dx.$ (1)

Интеграл (1) есть функция двух параметров p и q; он сходится при $p>0, q<0$. Функция B является симметричной относительно параметров, т.е. $B(p;q)=B(q;p)$.

Если сделать замену переменной интегрирования, полагая $x=\sin^2 t, dx=2\sin t\cos tdt$, причем t изменяется от 0 до $\frac{\pi}{2}$, то формула (1) примет вид $B(p;q)=2\int\limits_0^{\pi/2}\sin^{2p-1}t\cos^{2q-1}tdt$, или

$$\int\limits_0^{\pi/2}\sin^m x\cos^n xdx=\frac{1}{2}B\left(\frac{m+1}{2},\frac{n+1}{2}\right)(m>0,n>0)\ . \qquad (2)$$

К интегралам (1) и (2) приводятся многие интегралы, встречающиеся в прикладных задачах.

Для вычислении значений бета- функции пользуются следующей зависимостью между бета- и гамма- функцией:

$$B(p,q)=\frac{\Gamma(\mathrm{p})\cdot\Gamma(\mathrm{q})}{\Gamma(\mathrm{p}+\mathrm{q})}. \qquad (3)$$

Если $q=1-p$, то $B(p,1-p)=\frac{\Gamma(\mathrm{p})\cdot\Gamma(1\text{-}\mathrm{p})}{\Gamma(1)}=\frac{\pi}{\sin\pi p}(1<p<1)$.

Используя бета- функцию, легко найти значение $\Gamma\left(1/2\right)$. Пусть $p=q=1/2$; тогда $B\left(1/2,1/2\right)=\frac{\left[\Gamma\left(1/2\right)\right]^2}{\Gamma(1)}$. Так как $B\left(1/2,1/2\right)=B\left(1/2,1-1/2\right)=\pi/\sin(\pi/2)=\pi$, а $\Gamma(1)=1$, то $\Gamma(1/2)=\sqrt{\pi}$.

Список используемой литературы

1. Данко П.Е. Высшая математика в упражнениях и задачах./ П.Е. Данко, Л.Г. Попов – ч.1, П, Ш, - М.: Высшая школа, 1974,1980.
2. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников. – М.: ФИЗМАТЛИТ, 2006 – 816 с.
3. Кочергин А.Н. Основы надежности металлорежущих станков и измерительных приборов (для вузов спец. «тех. машиностр., металлорежущие станки и инструменты»./ А.Н. Кочергин и Л.Д. Ковалев – Минск: «Вышэйш. школа», 1974.
4. Математические методы в теории надежности. Основные характеристики надежности и их статистический анализ./ Б.В. Гнеденко, Ю.К. Беляев, А.Д. Соловьев. – М.: издательство «Наука», 1965 – 524 с.
5. Сборник задач по математике для втузов. Ч. 3. Теория вероятностей и математическая статистика: Учеб. пособие для втузов/ Под ред. А.В. Ефимова. – 2-е изд., перераб. и доп. – М.: Наука, 1990 – 428 с.
6. Смирнов Н.В. Курс теории вероятностей и математической статистики для технических приложений./ Н.В. Смирнов, И.В. Дунин-Барковский - М.: издательство «Наука», 1969 – 512 с.
7. Теория статистики: учебник/ Р.А. Шмойлова, В.Г. Минашкин, Н.А. Садовникова, Е.Б. Шувалова; под ред. Р.А, Шмойловой. – 4-е изд., перераб. и доп. – М.: Финансы и статистика, 2007. – 656 с.
8. Хан Г. Статистические модели в инженерных задачах. Пер. с анг. Е.Г. Коваленко. Под ред. В.В. Налимова/ Г. Хан, С. Шапиро – М.: «Мир», 1969.
9. www.quality.eup.ru
10. www.statmetKach.com (Статистические методы контроля и управления качеством: лаб. работы, учебник/ В.В. Заляжных, 2008).

Учебное издание

Ирина Викторовна Ребро
Владимир Андреевич Носенко
Неля Николаевна Короткова

Прикладная математическая статистика для технических специальностей

Учебное пособие

Редактор Е.М. Марносова

Темплан 2011 г. Поз. №
Подписано в печать .12.2010 г. Формат 60×84 1/16. Бумага газетная.
Гарнитура Times. Печать офсетная. Усл. печ. л. .
Тираж 100 экз. Заказ .

Волгоградский государственный технический университет
400131, г. Волгоград, пр. Ленина, 28, корп. 1.

Отпечатано в ИУНЛ ВолгГТУ
400131, г. Волгоград, пр. Ленина, 28, корп. 7.