

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Б.Б. Мойзес, И.В. Плотникова, Л.А. Редько

СТАТИСТИЧЕСКИЕ МЕТОДЫ КОНТРОЛЯ КАЧЕСТВА И ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

*Рекомендовано в качестве учебного пособия
Редакционно-издательским советом
Томского политехнического университета*

Издательство
Томского политехнического университета
2016

УДК 519.242(075.8)
ББК 22.172.в6я73
М74

Мойзес Б.Б.

М74 Статистические методы контроля качества и обработка экспериментальных данных : учебное пособие / Б.Б. Мойзес, И.В. Плотникова, Л.А. Редько ; Томский политехнический университет. – Томск : Изд-во Томского политехнического университета, 2016. – 119 с.

ISBN 978-5-4387-0700-4

В пособии представлены главные темы для рассмотрения студентами очной формы обучения: виды измерений и их погрешности, методы обработки экспериментальных данных. Основное внимание уделено вопросам построения и использования вероятностных и статистических моделей контроля качества, планированию эксперимента, рассмотрена интерполяция и аппроксимация функций.

Предназначено для студентов, обучающихся по направлению 120301 «Приборостроение».

УДК 519.242(075.8)
ББК 22.172.в6я73

Рецензенты

Доктор физико-математических наук, профессор ТГУ
О.Н. Чайковская

Кандидат технических наук
директор Государственного регионального центра
стандартизации, метрологии и испытаний
в Томской области
М.М. Чухланцева

ISBN 978-5-4387-0700-4

© ФГАОУ ВО НИ ТПУ, 2016
© Мойзес Б.Б., Плотникова И.В.,
Редько Л.А., 2016
© Оформление. Издательство Томского
политехнического университета, 2016

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ	5
ВВЕДЕНИЕ.....	6
1. НАУЧНЫЕ МЕТОДЫ ИССЛЕДОВАНИЯ	7
1.1. Теоретические методы	7
1.2. Эмпирические методы.....	8
1.3. Эксперимент как предмет исследования.....	9
2. ИЗМЕРЕНИЕ ФИЗИЧЕСКИХ ВЕЛИЧИН	11
2.1. Виды измерений.....	11
2.2. Погрешности измерений	14
2.3. Типы физических величин.....	16
3. КРАТКИЕ СВЕДЕНИЯ ИЗ ТЕОРИИ ВЕРОЯТНОСТИ.....	17
3.1. Основные понятия и определения.....	17
3.1.1. Случайная величина	17
3.1.2. Испытания и события	18
3.1.3. Генеральная совокупность и выборка	19
3.1.4. Абсолютная и относительная частота. Вероятность.....	19
3.2. Функция распределения и плотность вероятности	20
3.3. Числовые характеристики случайных величин	23
3.4. Законы распределения случайной величины	24
3.4.1. Законы распределения дискретной случайной величины	25
3.4.1.1. Биномиальное распределение	25
3.4.1.2. Закон распределения Пуассона.....	26
3.4.1.3. Гипергеометрическое распределение	28
3.4.2. Законы распределения непрерывной случайной величины	29
3.4.2.1. Равномерный закон распределения	29
3.4.2.2. Нормальный закон распределения	30
3.4.2.3. Логарифмически-нормальное распределение	33
3.4.2.4. Показательный закон распределения	34
3.4.2.5. Распределение Вейбулла–Гнеденко	35
4. СТАТИСТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ ДАННЫХ.....	37
4.1. Определение закона распределения случайной величины.....	38
4.2. Определение функции отклика	44
4.2.1. Корреляционный анализ	45
4.2.2. Регрессионный анализ.....	48
4.2.2.1. Метод наименьших квадратов	49
4.2.2.2. Парная линейная регрессионная модель	51
4.2.2.3. Параболическая регрессионная модель	52
4.2.2.4. Регрессия в виде показательной функции	54
4.2.2.5. Регрессия в виде степенной функции	55
4.2.3. Проверка математической модели на адекватность.....	56

4.3. Планирование эксперимента	58
4.3.1. Полный факторный эксперимент	58
4.3.1.1. Матрица планирования	59
4.3.1.2. Кодирование переменных	60
4.3.1.3. Свойства матрицы планирования	62
4.3.1.4. Расчет коэффициентов регрессии	62
4.3.2. Дробный факторный эксперимент	63
5. ИНТЕРПОЛЯЦИЯ И АППРОКСИМАЦИЯ ФУНКЦИИ	67
5.1. Интерполяция	67
5.1.1. Канонический полином	68
5.1.2. Линейная интерполяция	68
5.1.3. Интерполяционный полином Лагранжа	70
5.1.4. Интерполяционный полином Ньютона	72
5.1.4.1. Первая интерполяционная формула Ньютона	72
5.1.4.2. Вторая интерполяционная формула Ньютона	74
5.2. Аппроксимация функций	76
6. СТАТИСТИЧЕСКИЕ МЕТОДЫ КОНТРОЛЯ КАЧЕСТВА	78
6.1. Понятие и методы статистического контроля	78
6.2. Система нормативных документов, регламентирующих применение статистических методов	81
6.3. Семь простых методов контроля качества	83
6.4. Последовательность применения методов	95
6.5. Цикл PDCA и семь инструментов качества	96
6.6. Мышление в формате АЗ	99
ПРИЛОЖЕНИЕ 1. Значения функции Лапласа	103
ПРИЛОЖЕНИЕ 2. Критические значения критерия Пирсона (χ^2-критерий)	105
ПРИЛОЖЕНИЕ 3. Критические значения коэффициента Кохрена (G-критерия)	107
ПРИЛОЖЕНИЕ 4. Критические значения коэффициента Стьюдента (t-критерия)	109
ПРИЛОЖЕНИЕ 5. Значения критерия Фишера (F-критерия)	111
СПИСОК ЛИТЕРАТУРЫ	116

ПРЕДИСЛОВИЕ

Разложим название курса «Математические методы обработки экспериментальных данных» на «составляющие».

Метод – совокупность приемов или операций практического или теоретического освоения действительности, подчиненных решению конкретной задачи. Другими словами, метод – систематизированная совокупность действий, которые необходимо предпринять, чтобы решить поставленную задачу. В качестве метода могут выступать система операций при работе на определенном оборудовании, приемы научного исследования и т. д. [1].

Математические методы – совокупность методов количественного и качественного изучения и анализа состояния и (или) поведения объектов и систем, относящихся к различным отраслям науки, техники и технологий.

Научное исследование (научное познание) – один из видов познавательной деятельности, направленный на получение объективного, истинного знания, отражающего закономерности действительности. Задачи научного познания: описание, объяснение и предсказание процессов и явлений действительности.

Эксперимент (опыт) – один из методов научного исследования некоторого явления действительности в контролируемых и управляемых условиях [1, 5].

Данные – сведения о состоянии любого объекта: человека, фирмы, машины, прибора и т. д. [2], таким образом, экспериментальные данные – данные, полученные в результате эксперимента.

Обработка данных – процесс приведения данных к виду, удобному для использования. Любая система обработки данных выполняет три основные группы операций [2]:

- сбор данных;
- собственно их обработку;
- получение выходных данных.

Таким образом, под математическими методами обработки экспериментальных данных можно понимать совокупность приемов количественного и качественного изучения объектов посредством приведения сведений о его состоянии (поведении), полученных в контролируемых и управляемых условиях, к виду, удобному для использования.

В результате проведения экспериментов полученные данные требуют всевозможных обработок. При этом существует много методов и технологий обработки экспериментальных данных, выбор которых зависит от того результата, к которому стремится исследователь.

ВВЕДЕНИЕ

Познавательная деятельность, как сознательная деятельность индивидуума, направленная на приобретение сведений об объекте, кроме научного познания включает следующие виды познания [3, 4]:

- обыденно-практическое, основанное на опыте, который человек приобретает, осуществляя какую-либо деятельность. Результатом выступают привычки, убеждения, верования, знание языка, обычаев и т. д.;
- мифическое – познание мира через мифы, предания, совершение обрядов, ритуалов, мифическое мировосприятие: природа – арена усилий, действие борющихся сил; схватка этих сил проявляется в природных явлениях и т. д.;
- религиозное, в котором истиной является Бог, а способ раскрытия истины – откровение. Цель практического религиозного познания – внутреннее преобразование человека, открывающее ему доступ к истине, Богу, вере;
- философское. В философии, как и в науке, используется строгий язык понятий. Однако истина философского познания имеет этическую составляющую. Философия не ограничивается только решением познавательных задач. Одна из главнейших функций философии – духовное преобразование человека;
- художественное, носителем которого является искусство. Существующая реальность находит отражение в художественных образах. Язык художественных образов выступает и как средство общения, и как творческая игра смыслов.

Только применение целого ряда методов позволяет всесторонне изучить исследуемую проблему, все ее аспекты и параметры.

1. НАУЧНЫЕ МЕТОДЫ ИССЛЕДОВАНИЯ

Научные методы подразделяют на теоретические и эмпирические.

1.1. Теоретические методы

Индукция – метод исследования, в котором осуществляется переход от частного знания к более общему.

Индукция дает только вероятные заключения, нуждающиеся в дальнейшей проверке.

Примером умозаключения на основе применения метода индукции может стать рассуждение о формах государственного правления стран.

Аргентина – республика, Бразилия – республика, Венесуэла – республика, Эквадор – республика. Аргентина, Бразилия, Венесуэла, Эквадор – латиноамериканские государства. Все латиноамериканские государства являются республиками.

Италия – республика, Португалия – республика, Финляндия – республика, Франция – республика. Италия, Португалия, Финляндия, Франция – западноевропейские страны. Все западноевропейские страны являются республиками.

Посылки первого и второго индуктивного умозаключения истинны, т. к. все латиноамериканские государства – республики, а среди западноевропейских стран имеются монархии, например Англия, Бельгия и Испания [6].

Дедукция – метод исследования, в котором осуществляется переход от общего к частному.

Отличительная особенность дедуктивного умозаключения в том, что оно от истинных посылок всегда ведет к истинному заключению.

Примером дедукции может стать следующее умозаключение.

Все люди смертны. Все греки – люди. Следовательно, все греки – смертны [6].

Анализ – метод познания при помощи расчленения целостного предмета исследования на составляющие части.

Синтез – соединение отдельных сторон предмета в единое целое.

Абстрагирование – это мысленное отвлечение от несущественных свойств, связей, отношений предметов и выделение нескольких сторон, интересующих исследователя.

Конкретизация – наполнение какого-либо предмета частными признаками.

Обобщение – определение общего понятия, в котором находит отражение главное, основное, характеризующее объекты данного класса.

Аналогия – метод, посредством которого достигается знание о предметах и явлениях на основании того, что они имеют сходство с другими.

Сравнение – это установление различия между объектами материального мира или нахождение в них общего; осуществляется как при помощи органов чувств, так и при помощи специальных устройств.

Моделирование – создание и изучение заместителя (модели) объекта (например, компьютерное моделирование генома человека).

Формализация – отображение объекта или явления в знаковой форме какого-либо искусственного языка (математики, химии и т. д.).

Идеализация – это мысленное конструирование объектов, которые практически неосуществимы.

1.2. Эмпирические методы

Наблюдение – метод изучения объекта, находящегося в своих естественных условиях, без какого-либо вмешательства в него. Наблюдение осуществляется органами чувств, при помощи приборов, установок или автоматизированных систем наблюдения [6].

Практически любое исследование начинается с наблюдения для сбора предварительной информации.

Научное наблюдение в отличие от повседневного характеризуется:

- целенаправленностью – наблюдение ведется для решения поставленной задачи;
- плановностью – наблюдение проводится строго по плану согласно задаче исследования;
- активностью – исследователь активно ведет поиск, направленный на выделение нужных моментов в наблюдаемом явлении [7].

Сравнение – познавательный метод, выявляющий сходство/различие объектов, ступеней развития одного и того же объекта.

Описание – метод, состоящий в фиксации данных при помощи определенных систем обозначений, принятых в науке. Описание производится как путем обычного языка, так и специальными средствами, составляющими язык науки (символы, знаки, матрицы, графики и т. д.).

Измерение – совокупность действий, выполняемых при помощи определенных средств с целью нахождения числового значения измеряемой величины. Измерение – это наблюдение, дополненное процедурами измерения и статистической обработки результатов измерения [6].

При многообразии исследовательских методов основным источником получения информации об исследуемом объекте является эксперимент, а обработка экспериментальных данных осуществляется методами теории вероятности и математической статистики.

1.3. Эксперимент как предмет исследования

Особенности эксперимента [6, 8]:

- дает возможность активно участвовать исследователю, который:
 - сам вызывает изучаемое явление, а не ждет его появления;
 - может активно влиять на исследуемый процесс – менять условия его протекания, исключая отдельные его условия;
- воспроизводимость – условия эксперимента и проводимые при этом наблюдения (измерения) могут быть повторены требуемое для получения достоверных результатов количество раз;
- позволяет изучать объект в условиях, близких к идеальным, а также к экстремальным. Последнее позволяет выявить скрытые свойства объектов и глубже постичь их сущность.

Эксперимент условно можно разделить на несколько этапов [9]:

- подготовка проведения эксперимента;
- проведение эксперимента;
- подведение итогов эксперимента.

Подготовка проведения эксперимента включает в себя:

- постановку цели и задач;
- построение плана эксперимента;
- подбор методики проведения эксперимента;
- изучение начального состояния всей системы;
- выбор способов контроля протекания эксперимента, форм фиксации результатов, способов их обработки.

Проведение эксперимента включает в себя воздействие на объект и фиксацию результатов.

Подведение итогов заключается в описании эмпирических результатов и сравнении их с теоретическими выкладками.

В зависимости от способа получения данных об объекте (процессе, явлении) эксперименты можно разделить на пассивный и активный.

Пассивный эксперимент заключается в получении данных в режиме нормальной эксплуатации объекта.

Достоинство пассивного эксперимента заключается в отсутствии затрат на подготовку и проведение эксперимента.

Основной недостаток пассивного эксперимента заключается в том, что в нормальных условиях эксплуатации объекта различия в его состоянии невелики, и экспериментальные точки близки друг к другу. В этих условиях на точность описания сильное влияние могут оказать случайные погрешности. Для устранения этого влияния необходимо иметь достаточно большое количество экспериментальных данных.

Активный эксперимент проводится при целенаправленном изменении *факторов* (входных параметров объекта) – переменных, оказывающих влияние на результаты эксперимента. При наличии связи фиксируемой физической величины от изменения факторов физическая величина получила название отклика, а функция зависимости отклика от факторов – *функции отклика*.

Основное достоинство активного эксперимента – бóльшая точность получаемых данных, основной недостаток – затраты на проведение эксперимента.

Таким образом, можно сделать вывод о том, что активный эксперимент за счет целенаправленного изменения входных параметров позволяет получать необходимый объем информации при существенно меньшем числе опытов, чем при пассивном эксперименте.

2. ИЗМЕРЕНИЕ ФИЗИЧЕСКИХ ВЕЛИЧИН

2.1. Виды измерений

Измерение – нахождение значения физической величины опытным путем с помощью специальных технических средств [10].

Классифицировать измерения можно следующим образом:

- по видам измеряемых физических величин: линейные, оптические, электрические и др.;
- характеру зависимости измеряемой величины от времени измерения: статические и динамические;
- количеству измерений одной и той же величины: однократные и многократные;
- способу получения результатов: прямые и косвенные;
- степени однородности: совокупные и совместные;
- условиям измерений: равноточные и неравноточные;
- характеру результата: абсолютные и относительные;
- точности: технические и метрологические;
- степени воздействия на объект: контактные и бесконтактные;
- методу измерений: метод непосредственной оценки и метод сравнения с мерой и т. д.

Статические измерения – измерения, при которых измеряемая величина остается постоянной во времени, например, измерение габаритных размеров объекта.

Под *динамическими измерениями* понимают измерения, в процессе которых измеряемая величина является непостоянной во времени, например, измерение амплитуды колебаний работающего технологического оборудования.

Однократные измерения – измерения, выполненные один раз.

Многократные измерения выполняют одними и теми же приборами при повторении эксперимента при одних и тех же условиях.

При *прямых измерениях* искомое значение величины определяют непосредственно по устройству отображения измерительной информации применяемого средства измерений.

Прямое измерение можно условно выразить формулой

$$Y = X,$$

где X – значение, непосредственно получаемое из опыта; Y – искомое значение измеряемой величины.

Примером прямых измерений могут стать измерение массы весами, температуры – термометром и т. д.

Косвенные измерения – измерения, при которых искомое значение требуемой физической величины находят на основании известной зависимости между этой величиной и величинами, подвергаемыми прямым измерениям.

При косвенных измерениях числовое значение измеряемой величины находят путем вычисления по формуле

$$Y = f(X_1, X_2, \dots, X_n),$$

где X_n – значение, полученное прямым измерением (непосредственно из эксперимента); Y – искомое значение физической величины.

К косвенным измерениям можно отнести нахождение площади геометрической фигуры, например прямоугольника, по ранее проведенным прямым измерениям его сторон.

При *совокупных измерениях* осуществляется многократное измерение одной величины или измерение нескольких одноименных величин. Примером может стать нахождение значения массы отдельных гирь набора по известному значению массы одной из гирь и по результатам измерений (сравнений) масс различных сочетаний гирь.

Совместные измерения подразумевают измерение нескольких неоднородных величин для нахождения функциональной зависимости между ними, например зависимости габаритных размеров объекта от температуры.

Примером косвенного измерения является измерение сопротивления R по результатам измерений тока I и напряжения U

$$R = \frac{U}{I}.$$

Равноточными измерениями называют измерения физической величины, выполненные одинаковыми приборами в одних и тех же условиях.

Измерения, выполненные различными приборами и (или) при разных условиях, будут *неравноточными*.

При измерениях для отображения результатов используются разные оценочные шкалы, градуированные либо в единицах измеряемой физической величины, либо в различных относительных единицах, включая и безразмерные. В соответствии с этим принято различать *абсолютные* и *относительные* измерения.

Примерами абсолютных измерений являются: определение длины в метрах, массы в килограммах. Примером относительных измерений является измерение относительной влажности воздуха [11]:

$$\varphi = \frac{\rho}{\rho_0} \cdot 100 \%,$$

где ρ – абсолютная влажность, под которой понимается величина, численно равная массе водяного пара, содержащегося в 1 м^3 воздуха (т. е. плотность водяного пара в воздухе при данных условиях); ρ_0 – плотность насыщенного пара при данной температуре (максимально возможное количество водяного пара при данной температуре в 1 м^3 воздуха).

Технические измерения – это измерения, выполняемые обычными (не эталонными) техническими средствами измерений. Их выполняют с заранее установленной точностью, при которой погрешность измерений не должна превышать заранее установленного значения.

Метрологические измерения – это измерения, выполняемые с использованием эталонов. Их выполняют с максимально достижимой точностью, добиваясь минимальной погрешности измерения.

Метод непосредственной оценки – метод, при котором значение измеряемой физической величины определяется непосредственно по отсчетному устройству прибора прямого действия.

При *методе сравнения с мерой* прибор сравнивает измеряемую величину с аналогичной известной величиной, воспроизводимой мерой. Примерами используемых мер являются гири, концевые меры длины или угла и т. д. Данный метод реализуется как:

- дифференциальный метод;
- нулевой метод;
- метод совпадений;
- метод замещения;
- метод противопоставления.

Дифференциальный метод измерений – метод, в котором на измерительный прибор воздействует разность измеряемой величины и известной величины, воспроизводимой мерой.

Примером дифференциального метода является поверка мер длины сравнением с эталонными мерами на компараторе (приборе, предназначенном для сравнения мер) [12]. При этом производится неполное уравновешивание измеряемой величины X величиной X_m , воспроизводимой мерой, и определение их разности ΔX . Результатом измерений станет выражение

$$X = X_m + \Delta X.$$

Нулевой метод – метод, в котором результирующий эффект воздействия величин на прибор сравнения доводят до нуля.

Примером нулевого метода является взвешивание массы на весах с помощью набора гирь.

Метод совпадений – метод, в котором значение измеряемой величины оценивают, используя совпадение ее с величиной, воспроизводимой мерой (т. е. с фиксированной отметкой на шкале физической величины).

Примером данного метода является измерение длины при помощи штангенциркуля с нониусом.

В зависимости от одновременности или неодновременности воздействия на прибор сравнения измеряемой величины и величины, воспроизводимой мерой, различают методы замещения и противопоставления.

Метод замещения – метод, в котором измеряемую величину замещают известной величиной, воспроизводимой мерой, т. е. эти величины воздействуют на прибор последовательно.

Метод противопоставления – метод, в котором измеряемая величина и величина, воспроизводимая мерой, одновременно воздействуют на прибор сравнения, с помощью которого устанавливается соотношение между этими величинами.

2.2. Погрешности измерений

Погрешность измерения (ошибка измерения) – отклонение результата измерения от истинного значения измеряемой величины [10]:

- вследствие неточности измерительных приборов,
- при многократном повторении одного и того же измерения физической величины и т. д.

Различают погрешности:

- по форме представления:
 - абсолютные;
 - относительные;
 - систематические;
- характеру проявления:
 - случайные;
 - систематические;
 - грубые промахи.

Абсолютная погрешность измерения определяется как разность между истинным и измеренным значениями физической величины

$$x_0 - x = \Delta x, \quad (1)$$

где x_0 – истинное (действительное) значение измеряемой величины; x – результат измерения.

Абсолютная погрешность может быть как положительной, так и отрицательной.

Относительная погрешность – отношение абсолютной погрешности к истинному значению или к результату измерения, выраженное в процентах:

$$\delta = \pm \frac{\Delta x}{x} \cdot 100 \%, \text{ или } \delta = \pm \frac{\Delta x}{x_0} \cdot 100 \%. \quad (2)$$

Приведенная погрешность – отношение абсолютной погрешности к нормированному значению x_n , выраженное в процентах.

В качестве нормированного значения может быть взято, например, максимальное значение x_{\max} измеряемой величины $x_n = x_{\max}$.

Систематические погрешности – погрешности, сохраняющие величину и знак при повторении эксперимента в условиях равноточных измерений из-за различных факторов, таких как:

- несовершенство измерительной аппаратуры и метода измерений;
- неточная настройка измерительной аппаратуры;
- непостоянство условий эксперимента, включая влияние окружающей среды;
- ошибки экспериментатора и т. д.

Систематические погрешности могут быть выявлены при тарировке приборов, и потому они могут быть учтены при обработке результатов измерений.

Случайные погрешности – это погрешности, изменяющие свою величину или знак при повторении эксперимента в условиях равноточных измерений из-за суммарного эффекта действия таких факторов, как вибрация внешних источников, изменение окружающей среды и т. д.

Их невозможно учесть как систематические, но можно учесть их влияние на оценку истинного значения измеряемой величины. Анализ случайных погрешностей можно провести средствами теории вероятности и математической статистики.

Промахи или *грубые погрешности* – ошибочные измерения в результате небрежности при отсчете по прибору или неразборчивой записи показаний, при неправильном включении прибора или при нарушении условий, в которых должен проводиться опыт (изменение напряжения, загрязнение материала и т. д.). Данные погрешности легко выявляются при повторном проведении экспериментов.

В связи с вышесказанным, систематические и грубые погрешности, промахи считаются исключенными из результатов эксперимента.

2.3. Типы физических величин

Встречающиеся в эксперименте физические величины различают следующих основных типов [13]:

- случайная;
- постоянная;
- меняющаяся закономерно;
- нестабильная.

Случайной физической величиной называется величина, которая принимает в результате эксперимента одно значение из множества исходов, причем появление того или иного значения этой величины до ее измерения точно предсказать нельзя. Вместе с тем проведение достаточно большого количества измерений позволяет во многих случаях описать результаты измерений определенными статистическими закономерностями. Выявление этих закономерностей – одна из основных задач любого эксперимента. В качестве случайной величины можно рассматривать амплитуду колебаний технологического оборудования во время работы.

К постоянным величинам относят физические постоянные, например скорость света в вакууме. Постоянными величинами можно считать некоторые параметры конкретного объекта, находящегося при фиксированных условиях. Однако многократные измерения данных величин могут привести к разбросу значений, что делает постоянную величину случайной, а для ее обработки используются методы, характерные для обработки результатов измерения случайной величины.

Значение физической величины, меняющейся закономерно, подчиняется соответствующему закону. Примером может стать процесс затухания амплитуды свободных колебаний маятника.

Нестабильной величиной называется величина, изменение которой во времени не описывается какой-либо статистической закономерностью. Обычно это объясняется отсутствием достаточного количества информации. Для перевода данной величины в другую категорию необходимо экспериментально или теоретически установить закономерность изменения ее во времени.

3. КРАТКИЕ СВЕДЕНИЯ ИЗ ТЕОРИИ ВЕРОЯТНОСТИ

При проведении фиксации результатов экспериментов на значение измеряемой величины влияет множество факторов, не имеющих прямого отношения к изучаемому явлению (объекту), в первую очередь, случайных, не носящих закономерный характер. Поэтому многие физические величины, измеренные в ходе эксперимента, являются случайными.

Случайные погрешности не устраняются, но могут быть описаны закономерностями теории вероятностей, что позволяет при многократных измерениях определить пределы, внутри которых заключено истинное значение измеряемой величины [14].

3.1. Основные понятия и определения

В основе обработки случайных величин лежат знания вероятностных закономерностей массовых однородных случайных событий, являющихся предметом теории вероятностей. Данные знания позволяют построить закономерности изменения численных характеристик, описывающих данные события [15].

Методы теории вероятностей широко применяются в различных отраслях науки, техники и технологии: теории автоматического управления, теории надежности, теории ошибок наблюдений, теории массового обслуживания и т. д.

Закономерности теории вероятностей лежат в основе математической статистики, которая, в свою очередь, применяется при организации производства, при анализе технологических процессов, предупредительном и приемочном контроле качества продукции и т. п.

3.1.1. Случайная величина

Случайной величиной называется величина, которая при многократных равнооточных измерениях может принимать различные числовые значения. Типы случайных величин приведены ниже (рис. 1) [16].

Случайную величину принято называть *одномерной*, или *скалярной*, если в качестве результата эксперимента регистрируется одно число.

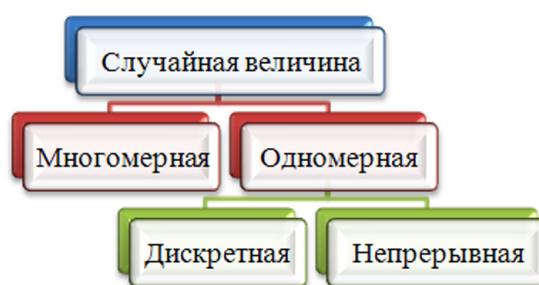


Рис. 1. Упрощенная схема классификации типов случайных величин

Если результатом эксперимента является регистрация набора характеристик, то соответствующую случайную величину называют *многомерной*, или *векторной*.

Одномерную случайную величину называют *дискретной*, или *непрерывной*, в зависимости от пространства элементарных событий, в котором она определена (в дискретном или непрерывном):

- дискретная случайная величина может принимать только определенные числовые значения;
- непрерывная случайная величина принимает непрерывный ряд значений.

3.1.2. Испытания и события

Испытанием (опытом) называется одно из сколь угодно большого числа раз воспроизводимого определенного перечня условий.

Событие – совокупность *явлений*, происходящих в результате *испытания* [17].



Рис. 2. Виды событий

Достоверное событие – событие, которое обязательно произойдет при соблюдении определенной совокупности условий. Например, если в сосуде содержится вода при нормальном атмосферном давлении и температуре 20 °С, то событие «вода в сосуде находится в жидком состоянии» – достоверное, а совокупность условий – атмосферное давление и температура воды [15].

Невозможное событие – событие, которое заведомо не может произойти при заданной совокупности условий. Так, для совокупности условий предыдущего примера событие «вода в сосуде находится в твердом состоянии» заведомо не произойдет.

Случайное событие – событие, которое при осуществлении совокупности условий может либо произойти, либо не произойти. В случае подбрасывания монеты может выпасть либо орел, либо решка. Каждое из данных событий («выпадение орла» либо «выпадение решки») – случайное.

События называют *несовместными*, если появление одного из них исключает появление других событий в одном и том же испытании. Примером может стать извлечение из ящика детали. Извлечение бракованной детали исключает извлечение небракованной. События «извлечение небракованной детали» и «извлечение бракованной детали» – несовместные.

События называют *равновозможными*, если появление одного из них не является более возможным, чем появление другого. Так, в примере с подбрасыванием монеты, предполагая, что она изготовлена из однородного материала, имеет правильную форму, а наличие чеканки не оказывает влияния на выпадение той или иной стороны монеты, появление орла или решки – равновозможные события.

Несколько событий образуют *полную группу событий*, если в результате испытания появится хотя бы одно из них, т. е. появление хотя бы одного из событий полной группы есть достоверное событие. Например, при стрельбе по цели обязательно произойдет одно из двух событий: попадание либо промах. Данные несовместные события образуют полную группу.

3.1.3. Генеральная совокупность и выборка

Генеральная совокупность – полный набор всех возможных значений, которые может принимать случайная физическая величина. Генеральная совокупность может быть конечной (число наблюдений $N = \text{const}$) или бесконечной ($N \rightarrow \infty$).

Выборкой объема называют набор n значений величин x_i , полученный из генеральной совокупности N .

Другими словами, под выборкой можно понимать реально рассматриваемую совокупность значений (x_1, x_2, \dots, x_i) случайной величины X , а под генеральной совокупностью – гипотетически существующую совокупность возможных значений.

Цель обработки набора величин x_i выборки – определение закономерностей, описывающих генеральную совокупность [18].

3.1.4. Абсолютная и относительная частота. Вероятность

Абсолютная частота случайного события A – количество m проявления данного события, зафиксированного в объеме данных n .

Относительная частота случайного события A :

$$W(A) = \frac{m_k}{n}, \quad (3)$$

где m_k – число появления данного события в серии испытаний; n – общее число проведенных одинаковых испытаний.

Наблюдения различных явлений демонстрируют, что при малом количестве испытаний в серии значения относительной частоты появления события A для разных серий испытаний различны. При большом числе испытаний в сериях относительные частоты появления события A в различных сериях отличаются друг от друга незначительно, и отличие уменьшается при увеличении числа испытаний в сериях. Другими словами, относительная частота при большом числе испытаний, как правило, перестает носить случайный характер (зависеть от серии) [19]:

$$W(A) = \frac{m_k}{n} \xrightarrow{n \rightarrow \infty} p, \quad (4)$$

где p – вероятность появления случайного события A .

Вероятностью события A называют отношение числа благоприятствующих этому событию исходов m к общему числу всех равновероятных несовместных элементарных исходов n , образующих полную группу [15]:

$$P(A) = \frac{m}{n} = p. \quad (5)$$

Из определения вероятности вытекают следующие ее свойства [15]:

- вероятность случайного события есть положительное число

$$0 \leq P(A) \leq 1;$$

- вероятность достоверного события

$$P(A) = 1;$$

- вероятность невозможного события

$$P(A) = 0.$$

3.2. Функция распределения и плотность вероятности

Функция распределения и плотность вероятности наиболее полно и наглядно описывают изменение случайной величины.

Для построения зависимостей функции распределения и плотности вероятности необходимо провести прямые многократные равноточные измерения заданной физической величины X , в результате чего получить набор значений (выборку) x_i (x_1, x_2, \dots, x_i). При этом истинное значение измеряемой величины x_0 остается неизвестным. Область значений разобьем на равные интервалы и определим количество измерений, попавших в тот или иной интервал. Обозначив через m_1, m_2, \dots, m_k количество измерений, попавших, соответственно в первый, второй, ..., k -й интервал, получим для каждого интервала значения:

- абсолютной частоты – m_1, m_2, \dots, m_k ;

- относительной частоты (3)

$$W_i = \frac{m_i}{n},$$

где i – порядковый номер интервала;

- плотности относительной частоты

$$f_W(x_i) = \frac{m_i}{\Delta x \cdot n}.$$

Откладывая по оси абсцисс интервалы Δx , а по оси ординат – значения m_i , W_i или f_W , строится гистограмма распределения (рис. 3).

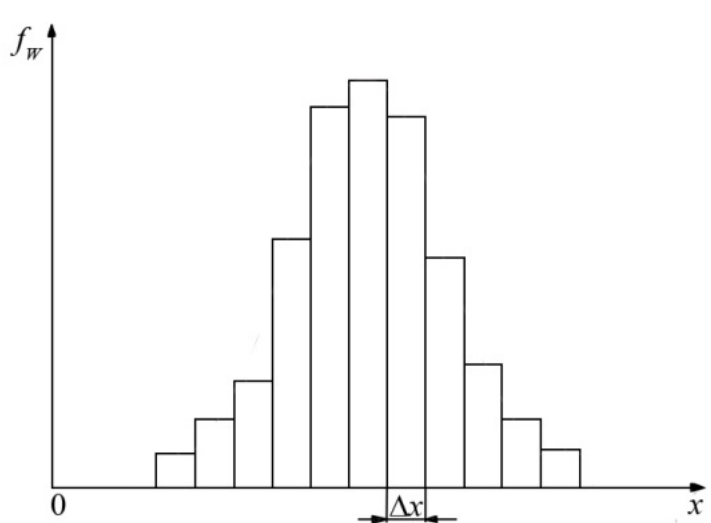


Рис. 3. Гистограмма распределения

При условии проведения большого количества испытаний (4) мы определим значения:

- вероятности того, что величина X может принимать определенные значения в каждом интервале $x_i + \Delta x$ (5)

$$P(X) = \frac{m}{n};$$

- плотности вероятности (распределения) – вероятности попадания случайной величины в заданный интервал

$$f(x) = \frac{m}{\Delta x \cdot n} = \frac{P(x)}{\Delta x}. \quad (6)$$

Под функцией распределения (вероятностей) случайной величины X понимают функцию $F(x)$, значение которой в точке x равно вероятности события $\{X \leq x\}$

$$F(x) = P(X \leq x).$$

Другими словами, функция распределения определяет вероятность появления случайной величины X , не превышающей значение x .

Для нахождения окончательной зависимости плотности вероятности необходимо увеличить число интервалов k . При этом длина интервала Δx стремится к нулю, а формула (4) примет вид:

$$f(x) = \frac{P(x)}{dx} = \frac{F(x)}{dx}.$$

Гистограмма преобразуется в плавную кривую (рис. 4).

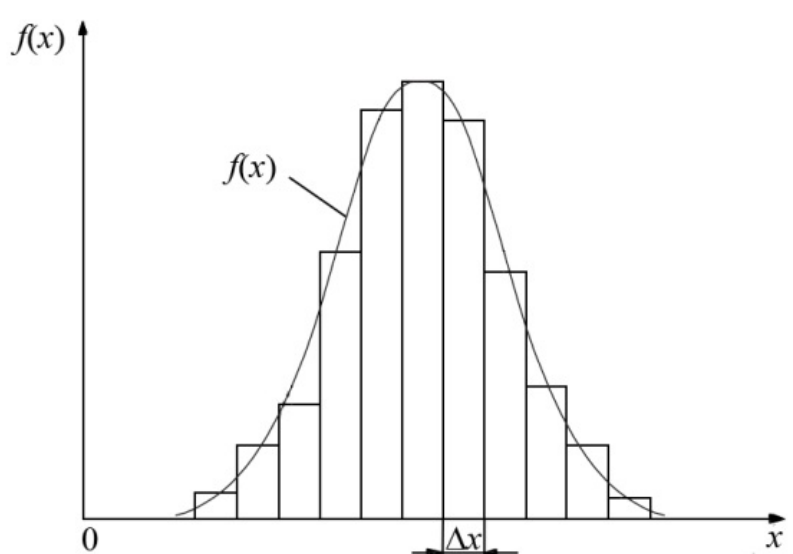


Рис. 4. Плотность вероятности

Плотность распределения, т. е. первая производная от функции распределения, позволяет вычислить, как площадь под кривой, ограниченной интервалом $x \in [x_i, x_i + dx]$, вероятность попадания случайной величины в данный интервал (рис. 5)

$$F(x) = \int_{x_i}^{x_i+dx} f(x) dx$$

и в конечном итоге определить, какие значения случайной величины наиболее вероятны.

Для интервала бесконечной длины $x \in [-\infty, +\infty]$ вероятность того, что фиксируемая случайная величина примет какое-либо значение, равна 1, т. е. это достоверное событие:

$$F(x) = \int_{-\infty}^{+\infty} f(x) dx = 1. \quad (7)$$

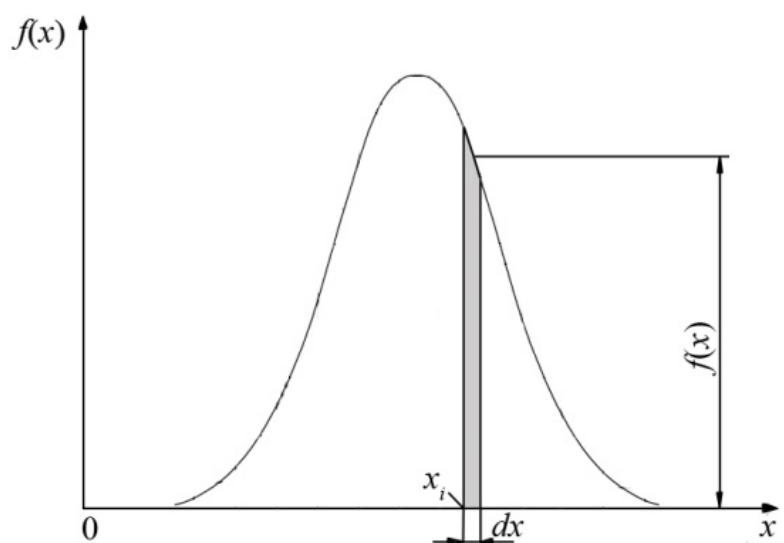


Рис. 5. Плотность вероятности

3.3. Числовые характеристики случайных величин

К основным числовым характеристикам случайных величин относят:

- среднее арифметическое значение,
- математическое ожидание,
- моду,
- медиану,
- выборочную дисперсию,
- среднеквадратичное отклонение,
- максимальное и минимальное значение,
- размах,
- квантиль,
- коэффициент вариации.

Среднее арифметическое значение фиксируемых величин в результате n измерений, если получены значения x_1, x_2, \dots, x_n :

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n) = \frac{1}{n} \sum_{i=1}^n x_i. \quad (8)$$

Математическое ожидание – характеристика центра группирования случайных величин (рис. 6, 7). В большинстве случаев математическое ожидание характеризует наиболее вероятное расположение значений случайной величины.

Математическое ожидание определяется:

- для дискретных случайных величин

$$M(x) = \sum_{i=1}^n x_i p_i; \quad (9)$$

- непрерывных случайных величин

$$M(x) = \int_{-\infty}^{+\infty} x f(x) dx. \quad (10)$$

Модой непрерывной случайной величины называется то ее значение, в котором плотность вероятности наибольшая. *Модой* дискретной случайной величины называется ее наиболее часто встречающееся значение.

Медианой случайной величины X называется такое ее значение, при котором площади под кривой справа и слева от медианы равны.

Дисперсией называют меру отклонения случайной величины X от ее математического ожидания $M(x)$.

Дисперсия определяется:

- для дискретных случайных величин

$$D(x) = \sum_{i=1}^n (x_i - M(x))^2 \cdot p_i; \quad (11)$$

- непрерывных случайных величин

$$D(x) = \int_{-\infty}^{+\infty} (x - M(x))^2 f(x) dx. \quad (12)$$

Среднеквадратичным отклонением называется величина, равная

$$\sigma(x) = \sqrt{D(x)}. \quad (13)$$

Размах – разница между максимальным и минимальным значениями выборки

$$R = x_{\max} - x_{\min}. \quad (14)$$

Квантиль – значение случайной величины, соответствующее заданной вероятности.

Коэффициент вариации характеризует относительную меру отклонения измеренных значений от среднего

$$v(x) = \frac{\sigma(x)}{M(x)}.$$

3.4. Законы распределения случайной величины

Законом распределения случайной величины называется функциональная зависимость между возможными значениями случайной величины и соответствующими им вероятностями [17]. Законы распределения наиболее полно описываются числовыми характеристиками, функцией распределения и плотностью вероятности.

Рассмотрим основные законы, встречающиеся при описании отказов технических систем и их элементов (в теории надежности).

3.4.1. Законы распределения дискретной случайной величины

Изменение дискретной величины X задается в табличной форме: определенному значению x_i соответствует вероятность p_i (табл. 1).

Таблица 1

Соответствие значений случайной величины вероятностям

x_i	x_1	x_2	x_3	x_{n-1}	x_n
p_i	p_1	p_2	p_3	p_{n-1}	p_n

Наиболее часто используют три закона распределения – биномиальное, Пуассона и гипергеометрическое [20].

3.4.1.1. Биномиальное распределение

Биномиальный закон распределения используется для определения вероятности наступления некоторого события X целочисленных m раз в n независимых испытаниях, выполняемых в одинаковых условиях

$$P_n(X) = C_n^m \cdot p^m \cdot q^{n-m}, \quad (15)$$

где p – вероятность того, что фиксируемое событие произойдет; q – вероятность того, что фиксируемое событие не произойдет; C_n^m – число сочетаний из n элементов по m

$$C_n^m = \frac{n!}{m!(n-m)!}. \quad (16)$$

Закон распределения считается заданным, если известны параметры n и p (табл. 2), т. к. согласно теории суммы вероятностей сумма вероятностей противоположных событий равна единице [15]

$$q = 1 - p.$$

Таблица 2

Варианты значений вероятностей

x_i	0	1	2	...	m	...	x_n
p_i	q^n	$n \cdot p \cdot q^{n-1}$	$C_n^2 \cdot p^2 \cdot q^{n-2}$...	$C_n^m \cdot p^m \cdot q^{n-m}$...	p^n

Математическое ожидание и дисперсия определяются следующим образом:

$$M(x) = np; \quad (17)$$

$$D(x) = npq. \quad (18)$$

Биномиальный закон распределения широко используется в теории и практике статистического контроля качества продукции, при описании функционирования систем массового обслуживания, в теории стрельбы, других областях [16].

Пример

В партии однотипных деталей стандартные составляют 95 %. Наугад из партии берут 500 деталей. Определить математическое ожидание $M(x)$, дисперсию $D(x)$ и среднеквадратичное отклонение $\sigma(x)$ для дискретной случайной величины X (появление числа стандартных деталей среди 500).

Решение

Случайная величина X имеет биномиальный закон распределения вероятностей и может принимать значения $x = m = 0, 1, \dots, 500$. Вероятности возможных значений для данной задачи определяются по формулам (15), (16):

$$P_n(X = m) = C_{500}^m \cdot p^m \cdot q^{500-m}, \quad C_n^m = \frac{500!}{m!(500-m)!},$$

где p – вероятность появления стандартной детали, $p = 0,95$; где q – вероятность появления нестандартной детали, $q = 0,05$.

Согласно формулам (13, 17, 18) определяем нужные величины:

$$M(x) = np = 500 \cdot 0,95 = 475;$$

$$D(x) = npq = 500 \cdot 0,95 \cdot 0,05 = 23,75;$$

$$\sigma(x) = \sqrt{500 \cdot 0,95 \cdot 0,05} = 4,87.$$

3.4.1.2. Закон распределения Пуассона

Закон распределения Пуассона моделирует случайную величину, представляющую собой число событий, произошедших за фиксированное время с некоторым фиксированным значением параметра распределения. Вероятность успешных испытаний

$$P(X = m) = \frac{\lambda^m}{m!} e^{-\lambda}, \quad (19)$$

где e – основание натурального логарифма, равное 2,71828; m – количество успехов в единицу времени; λ – параметр распределения (ожидаемое количество успехов)

$$\lambda = np,$$

где n – общее количество испытаний; p – вероятность появления успеха (ожидаемого события).

Табличная форма закона распределения приведена ниже (табл. 3).

Математическое ожидание и дисперсия:

$$M(x) = \lambda; \quad (20)$$

$$D(x) = \lambda. \quad (21)$$

Таблица 3

Варианты значений вероятностей

x_i	0	1	2	...	x_n
p_i	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{1}{2!} \lambda^2 e^{-\lambda}$...	$\frac{1}{n!} \lambda^n e^{-\lambda}$

Распределение Пуассона используется при определении числа сбоев на автоматической линии, числа отказов сложной системы, анализе результатов маркетинговых исследований потребителей, расчете характеристик планов статистического приемочного контроля в случае малых значений приемочного уровня дефектности¹, для описания числа разладок статистически управляемого технологического процесса, статистических закономерностей несчастных случаев и т. д. [20].

Пример

В результате проверки 1000 партий одинаковых изделий получено распределение количества бракованных изделий в партии (табл. 4).

Таблица 4

Результаты статической обработки

Число бракованных деталей в партии, x_i	0	1	2	3	4	Итого
Количество партий, содержащих данное число бракованных деталей, f_i	604	306	77	12	1	1000

Определить математическое ожидание $M(x)$, дисперсию $D(x)$ и среднеквадратичное отклонение $\sigma(x)$, для дискретной случайной величины X (появление бракованных деталей).

Решение

Определим параметр распределения – среднее число бракованных изделий в партиях:

$$\lambda = \frac{0 \cdot 604 + 1 \cdot 306 + 2 \cdot 77 + 3 \cdot 12 + 4 \cdot 1}{1000} = 0,5.$$

По формулам (13, 20, 21) найдем соответствующие величины:

$$M(x) = 0,5; \quad D(x) = 0,5; \quad \sigma(x) = 0,707.$$

¹ Максимальный уровень дефектности для одиночных партий или средний уровень для последовательности партий, который для целей приемки продукции рассматривается как удовлетворительный.

3.4.1.3. Гипергеометрическое распределение

Данное распределение типично для выборочного контроля качества продукции по альтернативному признаку. Генеральной совокупностью в этом случае является контролируемая партия продукции, в которой m единиц продукции не соответствуют определенному показателю. Вероятность того, что в выборке объемом n будет обнаружено ровно m несоответствующих изделий, определяется соотношением (22)

$$P(X = m) = \frac{C_M^m \cdot C_{N-M}^{n-m}}{C_N^n}, \quad (22)$$

где X – дискретная величина со значениями $0, 1, 2, \dots, m, \dots, M$; N и n – объем партии и выборки; M и m – количество дефектных единиц продукции в партии и выборке; C_M^m – число сочетаний из M по m

$$C_M^m = \frac{M!}{m!(M-m)!}.$$

Математическое ожидание и дисперсия вычисляются по приведенным ниже формулам:

$$M(x) = n \frac{M}{N}; \quad (23)$$

$$D(x) = \frac{nM(N-M)(N-n)}{N^2(N-1)}. \quad (24)$$

Данное распределение широко используется в практике статистического приемочного контроля качества промышленной продукции, в задачах, связанных с организацией выборочного обследования, и др.

Пример

В ящике содержится 50 деталей из двух серий, из них 25 – из второй серии. Наугад из ящика берут 5 деталей. Какова вероятность того, что среди них 4 детали окажутся из второй серии?

Решение

Для выбора пяти деталей начальными условиями станут:

$$n = 5; \quad N = 50; \quad m = 5; \quad M = 25; \quad N - M = 25.$$

Согласно (22) решение примет вид

$$P(X = 4) = \frac{C_{25}^4 \cdot C_{50-25}^{5-4}}{C_{50}^5} = \frac{C_{25}^4 \cdot C_{25}^{5-4}}{C_{50}^5} = \frac{\frac{25!}{4! \cdot 21!} \cdot \frac{25!}{1! \cdot 24!}}{\frac{50!}{5! \cdot 45!}} \approx 0,149.$$

3.4.2. Законы распределения непрерывной случайной величины

Во многих практических задачах приходится сталкиваться с определенными законами распределения непрерывных случайных величин. Часто встречаются законы равномерного, нормального, показательного распределения вероятностей.

3.4.2.1. Равномерный закон распределения

Встречаются случайные величины, о которых заранее известно, что они могут принять какое-либо значение в строго определенных границах, причем в этих границах все значения случайной величины имеют одинаковую вероятность (плотность вероятности).

Функция распределения (рис. 6)

$$F(x) = \frac{x-a}{b-a}, \quad a < x < b. \quad (25)$$

Плотность вероятности (рис. 6)

$$f(x) = \begin{cases} 0, & x < a; \\ \frac{1}{b-a}, & a < x < b; \\ 0, & b < x. \end{cases} \quad (26)$$

Математическое ожидание и дисперсия вычисляются по приведенным ниже формулам:

$$M(x) = \frac{a+b}{2}; \quad (27)$$

$$D(x) = \frac{(b-a)^2}{12}. \quad (28)$$

Равномерному закону распределения подчиняются, например, погрешности округления, возникающие в результате отбрасывания у чисел одной или нескольких цифр. С этим довольно часто приходится встречаться экспериментаторам в различных областях измерительной практики. Данное свойство используется в некоторых генераторах псевдослучайных чисел [21].

Равномерному закону подчиняются случайные погрешности, например, от трения в опорах стрелочных приборов, от квантования по значению в цифровых приборах и ряд других [22].

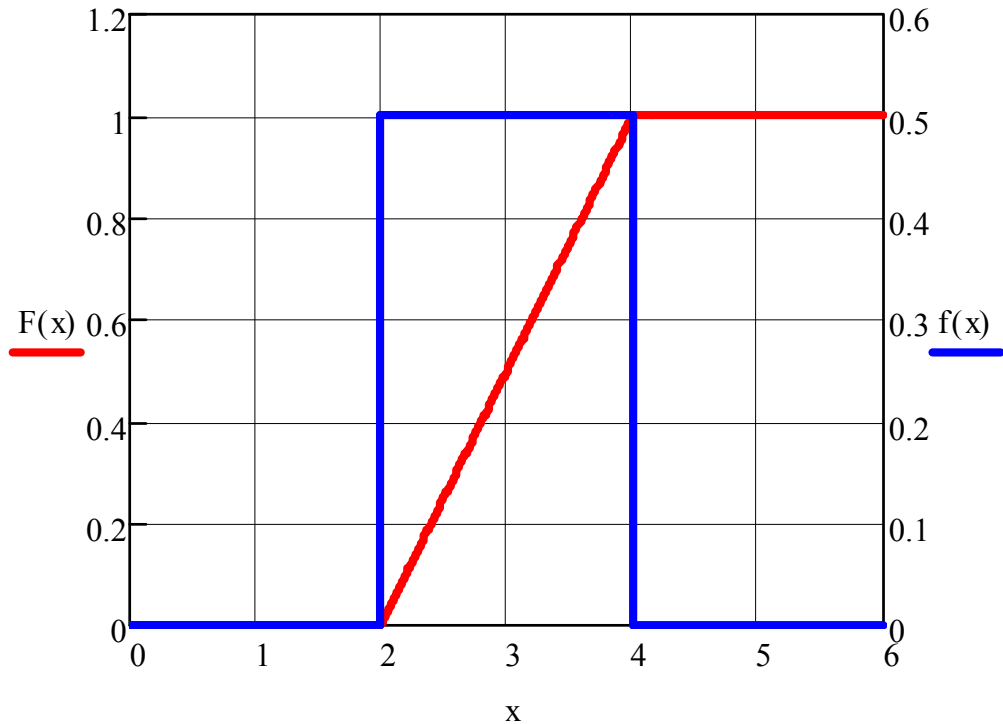


Рис. 6. Графики функции распределения $F(x)$ и плотности вероятности $f(x)$ при $a = 2$, $b = 4$

3.4.2.2. Нормальный закон распределения

Нормальный закон распределения называют фундаментальным законом в теории вероятности, т. к. его применение часто встречается при изучении природных и социально-экономических явлений.

Особенность закона состоит в том, что он является предельным законом, к которому приближаются другие законы распределения при весьма часто встречающихся типичных условиях.

Распределение плотности вероятности случайной величины называется нормальным, если плотность вероятности описывается функцией следующего вида (рис. 7):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad (29)$$

где m – математическое ожидание, мода, медиана; σ – стандартное отклонение.

Функция распределения (рис. 7)

$$F(x) = \int_a^b f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{(x-m)^2}{2\sigma^2}} dx, \quad (30)$$

где a и b – пределы изменения значений случайной величины X .

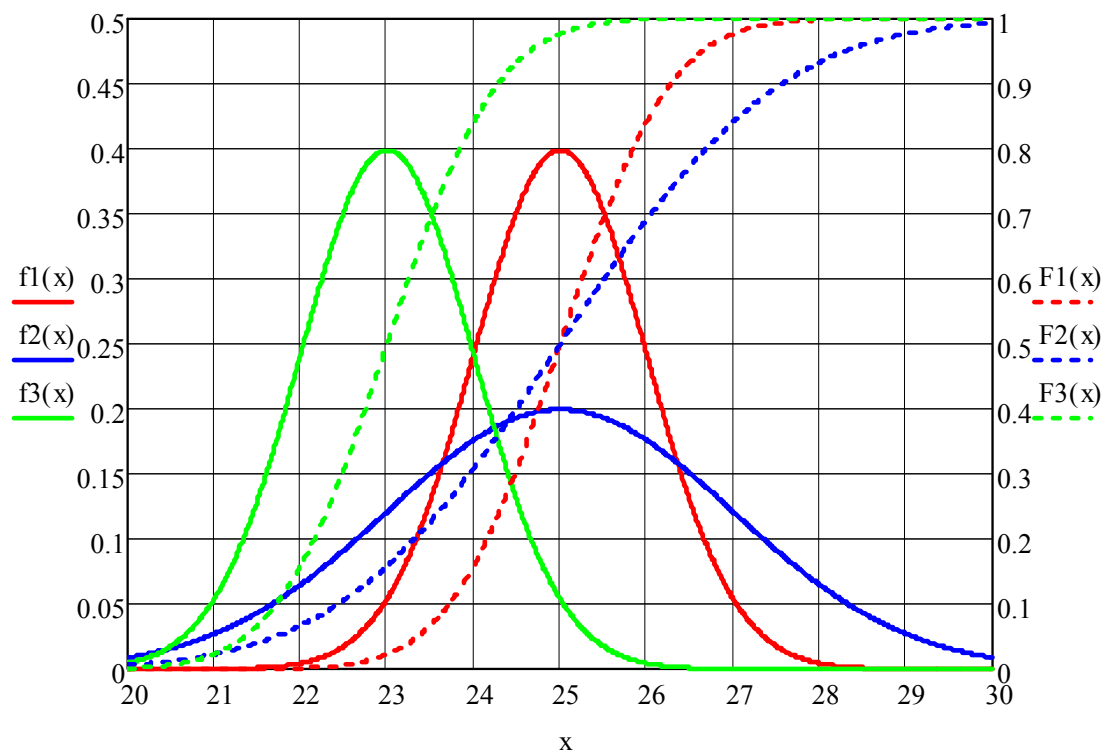


Рис. 6а. Графики функции распределения $F(x)$ и плотности вероятности $f(x)$: $m=1, \sigma=1(f_1(x), F_1(x))$; $m=25, \sigma=2(f_2(x), F_2(x))$; $m=23, \sigma=1(f_3(x), F_3(x))$

Вероятность попадания случайной непрерывной величины X в заданный интервал $[a, b]$

$$P(a < X < b) = \int_a^b f(x) dx. \quad (31)$$

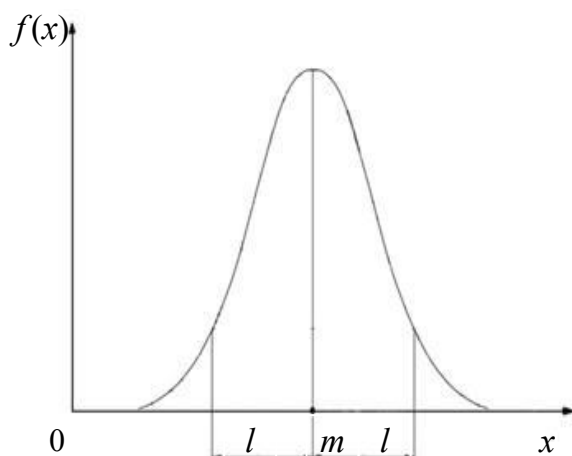


Рис. 7. График плотности распределения случайной величины X

Математическое ожидание и дисперсия вычисляются по приведенным формулам:

$$M(x) = \int_{-\infty}^{\infty} x f(x) dx; \quad (32)$$

$$D(x) = \int_{-\infty}^{\infty} (M(x) - x)^2 f(x) dx. \quad (33)$$

Часто приходится вычислять вероятность того, что значение случайной величины X попадает в интервал $(m-l, m+l)$ (рис. 7).

Вероятность, с которой в условиях данного эксперимента полученные экспериментальные данные можно считать надежными либо достоверными, называют *доверительной вероятностью*, или *надежностью*, а интервал $x \in [-l, +l]$, соответствующий доверительной вероятности, называется *доверительным интервалом*.

Вероятность того, что значение случайной величины X попадет в интервал $(m - l, m + l)$, выраженный через среднеквадратичное отклонение σ ($l = \sigma, 2\sigma, 3\sigma$), приведена в табл. 5.

Таблица 5

Правило трех сигм

Доверительный интервал	Вероятность, %
$m - \sigma \leq x \leq m + \sigma$	68,3
$m - 2\sigma \leq x \leq m + 2\sigma$	95,5
$m - 3\sigma \leq x \leq m + 3\sigma$	99,7

По правилу трех сигм вероятность того, что случайная величина отклонится от своего математического ожидания на величину, большую утроенного среднеквадратичного отклонения, практически равна нулю.

Часто, применяя закон нормального распределения, пользуются понятием функции Лапласа. Для ее получения в выражении (33) производят следующую замену

$$t = \frac{x - m}{\sigma}, \quad (34)$$

следовательно,

$$x = \sigma t + m, \quad dx = d(\sigma t + m) = \sigma dt, \quad (35)$$

а пределы интегрирования по введенной переменной t , соответственно верхний и нижний:

$$\frac{a - m}{\sigma}, \quad \frac{b - m}{\sigma}. \quad (36)$$

Подставив (38), (39) и (40) в (34) и проведя преобразования, получим

$$P(a < X < b) = \frac{1}{\sqrt{2\pi}} \int_0^{\frac{b-m}{\sigma}} e^{-\frac{z^2}{2}} dz - \frac{1}{\sqrt{2\pi}} \int_0^{\frac{a-m}{\sigma}} e^{-\frac{z^2}{2}} dz = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right).$$

Выражение

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz \quad (37)$$

получило название функции Лапласа.

Обычно функция Лапласа используется для вычисления вероятности попадания в заданный интервал нормально распределенной случайной величины X :

$$P(a < X < b) = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right). \quad (38)$$

Другими словами, сначала нормализуются величины a и b (40), а затем вычисляется либо определяется по таблице (прил. 1) от них функция Лапласа

$$\Phi\left(\frac{b-m}{\sigma}\right) \text{ и } \Phi\left(\frac{a-m}{\sigma}\right).$$

Пример

Найти вероятность того, что случайная величина X , имеющая нормальное распределение с математическим ожиданием $m=15$ и среднеквадратичное отклонение $\sigma=5$, примет значение, принадлежащее интервалу $[10, 20]$.

Решение

Вероятность найдем по выражению (38) и прил. 1:

$$\begin{aligned} P(10 < X < 20) &= \Phi\left(\frac{20-15}{5}\right) - \Phi\left(\frac{10-15}{5}\right) = \Phi(1) - \Phi(-1) = \\ &= \Phi(1) + \Phi(1) = 2 \cdot 0,341 = 0,682. \end{aligned}$$

3.4.2.3. Логарифмически-нормальное распределение

В логарифмически-нормальном распределении нормальному закону распределения с параметрами m и σ подчиняется изменение величины Y , значение которой равно логарифму значения случайной физической величины X ($X = e^Y$)

$$y = \ln x. \quad (39)$$

Плотность вероятности задается формулой

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - m)^2}{2\sigma^2}}, \quad (40)$$

а функция распределения –

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b \frac{e^{-\frac{(\ln x - m)^2}{2\sigma^2}}}{x} dx. \quad (41)$$

Математическое ожидание и дисперсию определяют по формулам:

$$M(x) = e^{\frac{m-\sigma^2}{2}}; \quad (42)$$

$$D(x) = (e^{\sigma^2} - 1)e^{2m+\sigma^2}. \quad (43)$$

Графики плотности вероятности и функции распределения приведены ниже (рис. 8).

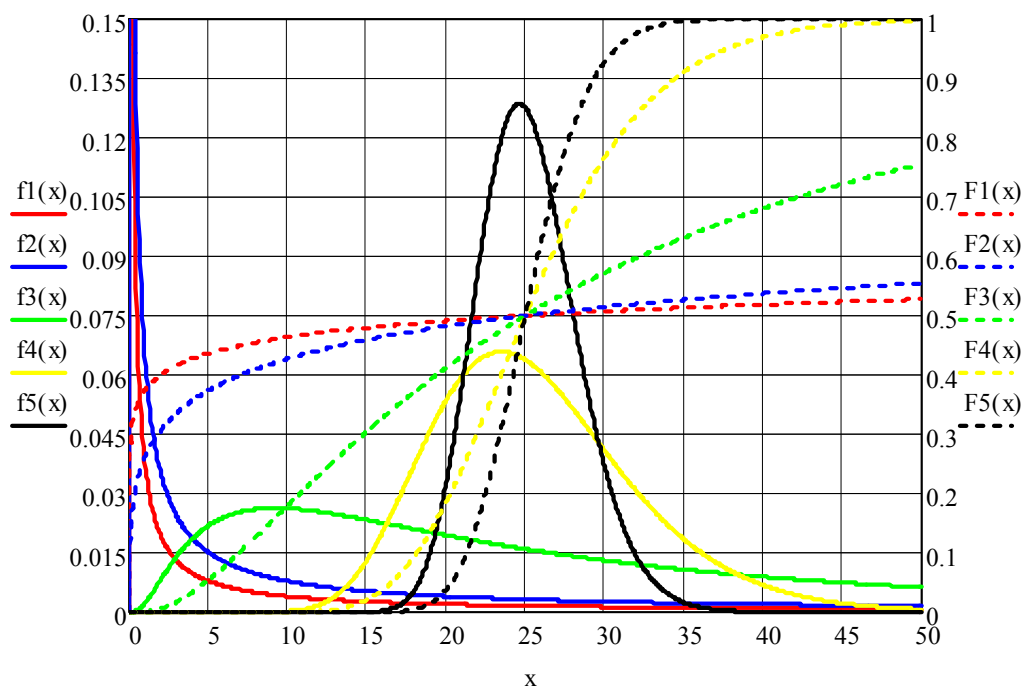


Рис. 8. Графики плотности вероятности $f(x)$ и функции распределения

$F(x)$: $m = 25, \sigma = 10(f1(x), F1(x))$; $m = 25, \sigma = 5(f2(x), F2(x))$;

$m = 25, \sigma = 1(f3(x), F3(x))$; $m = 25, \sigma = 0,25(f4(x), F4(x))$;

$m = 25, \sigma = 0,125(f5(x), F5(x))$

Логарифмически-нормальное распределение встречается в ряде задач [20], например, оно дает распределение размеров частиц при дроблении, содержание элементов в минералах в изверженных горных породах, выявляет численность рыб в море и т. д.

3.4.2.4. Показательный закон распределения

Плотность вероятности и функция распределения показательного (экспоненциального) закона распределения (рис. 9) запишутся как:

$$f(x) = \begin{cases} 0, & x < 0; \\ \lambda e^{-\lambda x}, & x \geq 0; \end{cases} \quad (44)$$

$$F(x) = \begin{cases} 0, & x < 0; \\ 1 - e^{-\lambda x}, & x \geq 0; \end{cases} \quad (45)$$

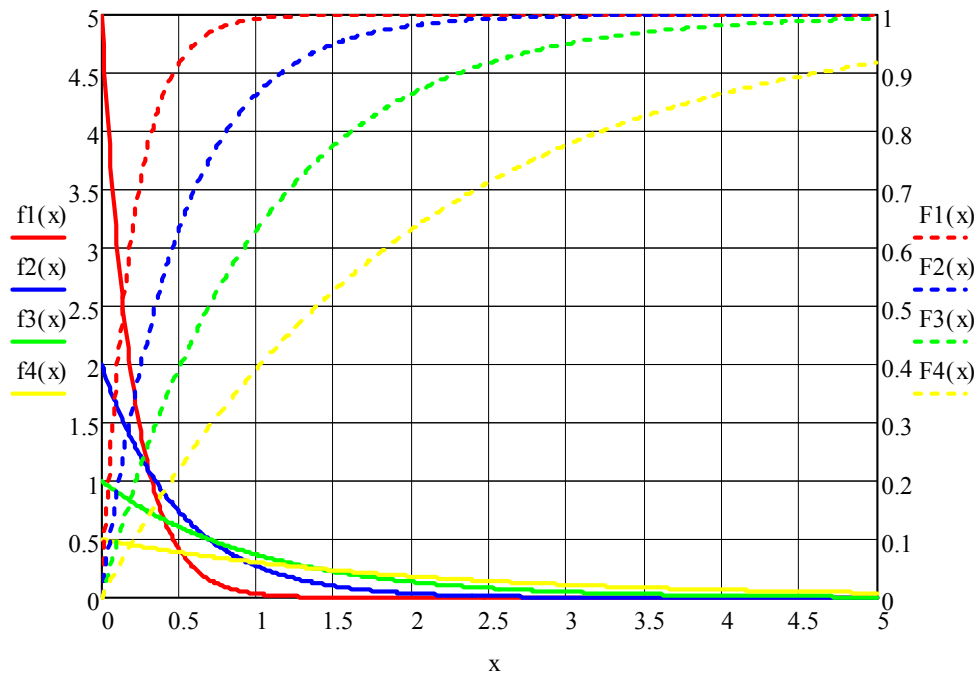


Рис. 9. Графики плотности вероятности $f(x)$ и функции распределения $F(x)$: $\lambda = 5(f1(x), F1(x))$; $\lambda = 2(f2(x), F2(x))$; $\lambda = 1(f3(x), F3(x))$; $\lambda = 0,5(f4(x), F4(x))$

Математическое ожидание и дисперсию определяют по формулам:

$$M(x) = \frac{1}{\lambda}; \quad (46)$$

$$D(x) = \frac{1}{\lambda^2}. \quad (47)$$

Статистический смысл параметра λ заключается в том, что λ – среднее число событий на единицу времени.

Другими словами, $1/\lambda$ – средний промежуток времени между двумя последовательными событиями.

Данное распределение часто встречается в теории надежности, в которой X – срок службы аппаратуры или отдельных ее узлов.

3.4.2.5. Распределение Вейбулла–Гнеденко

Функция распределения и плотность вероятности данного закона задаются формулами

$$F(x) = 1 - e^{-\lambda x^\alpha}, \quad (48)$$

$$f(x) = \lambda \alpha x^{\alpha-1} - e^{-\lambda x^\alpha}, \quad (49)$$

где α – параметр формы (определяется подбором в результате обработки экспериментальных данных); λ – параметр масштаба.

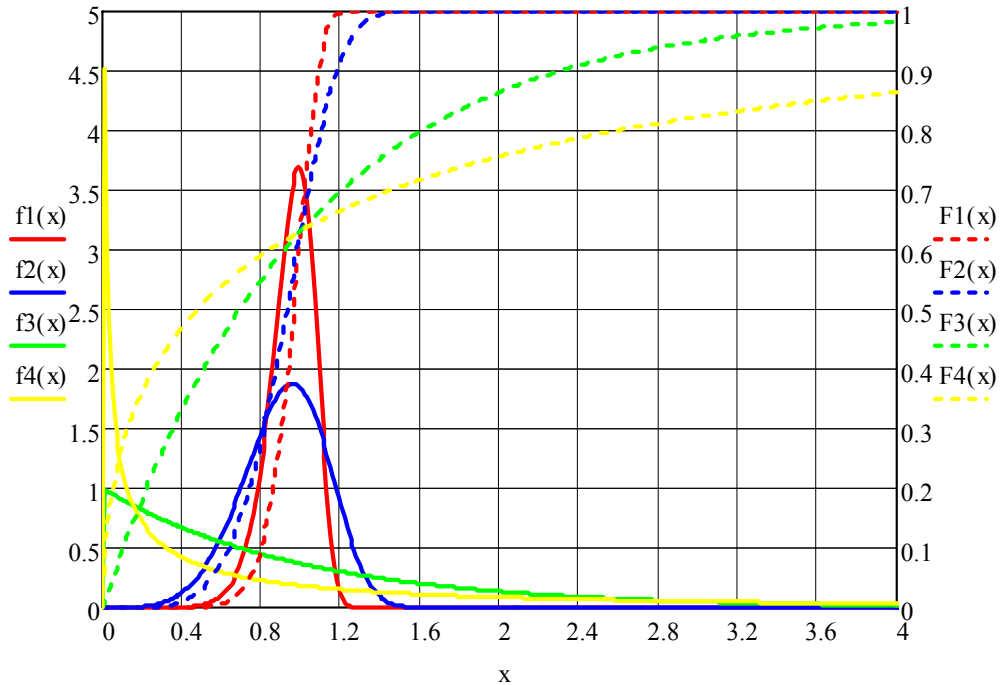


Рис. 10. Графики плотности вероятности $f(x)$ и функции распределения $F(x)$: $\lambda = 10(f1(x), F1(x))$; $\lambda = 5(f2(x), F2(x))$; $\lambda = 1(f3(x), F3(x))$; $\lambda = 0,5(f4(x), F4(x))$

Распределение Вейбулла–Гнеденко традиционно используется для описания времени жизни образцов, поставленных на испытание, при построении вероятностных моделей ситуаций, в которых поведение объекта определяется «наиболее слабым звеном», как в случае с цепью, работоспособность которой определяется звеном, которое имеет наименьшую прочность.

4. СТАТИСТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ ДАННЫХ

Математические законы теории вероятностей не являются беспредметными абстракциями, лишенными физического содержания, а представляют собой математическое выражение реальных закономерностей, существующих в массовых случайных явлениях [23]. Теория вероятностей дает возможность определить вероятности событий, законы распределения и числовые характеристики случайных величин, но для того, чтобы провести обработку и представить данные, необходимо провести эксперимент, применить (разработать) метод фиксации данных.

Раздел математики, изучающий методы сбора, систематизации и обработки результатов наблюдений с целью выявления статистических закономерностей, получил название *математическая статистика* (рис. 11).



Рис. 11. Направления математической статистики

В связи с этим основной задачей, решаемой математической статистикой, является [23] определение закона распределения случайной величины и нахождение числовых характеристик распределения либо вида функции отклика и ее параметров.

Методы математической статистики используются при планировании организации производства, анализе технологических процессов, для контроля качества продукции и многих других целей.

4.1. Определение закона распределения случайной величины

Для определения закона распределения, описывающего заданную выборку (набор экспериментальных данных), необходимо сравнить две выборки данных: экспериментальную и теоретическую, для чего требуется выполнить следующие шаги.

1. Группировка данных и построение гистограммы

Для построения гистограммы данные группируют, для чего диапазон изменения значений выборки разбивают на несколько равных интервалов k шириной

$$\Delta x = \frac{x_{\max} - x_{\min}}{k}, \quad (50)$$

где x_{\min} и x_{\max} – соответственно минимальное и максимальное значения в выборке $\{x_1, \dots, x_N\}$. Количество интервалов k согласно различным рекомендациям [24] может получиться разным, а при больших объемах выборок n достаточно большим. Можно воспользоваться рекомендациями Всероссийского научно-исследовательского института метрологии (табл. 6) [24, 25].

Таблица 6

Рекомендации по выбору количества интервалов

n	k
40...100	7...9
100...500	8...12
500...1000	10...16
1000...10000	12...22

Далее, для каждого интервала подсчитывается количество попадающих в него значений m . Затем для каждого интервала вычисляется относительная частота (3).

По полученным данным строится гистограмма (см. рис. 3).

2. Проведение анализа формы гистограммы и выдвижение гипотезы о законе распределения физической величины

Во многих случаях при изучении статистических данных (выборки) необходимо знать закон распределения генеральной совокупности. Для этого проводится сравнение формы гистограммы с внешним видом закономерностей законов распределения. Делается предположение о законе изменения физической величины, т. е. выдвигается статистическая гипотеза о том, что генеральная совокупность распределена по выбранному закону. Данная гипотеза является основной (нулевой) гипотезой H_0 .

Одновременно с основной гипотезой H_0 выдвигается альтернативная гипотеза H_1 , являющаяся логическим отрицанием гипотезы H_0 и принимаемая при отвержении гипотезы H_0 .

После выдвижения гипотезы вычисляются значения функции распределения (плотности вероятности и т. д.) для экспериментальной выборки в ряде точек по принятому закону.

3. Проверка правдоподобия гипотезы о законе распределения

Проверка статистической гипотезы означает проверку согласования исходных выборочных данных (выборки) с выдвинутой основной гипотезой.

При этом возможны две ситуации – основная гипотеза:

- подтверждается;
- опровергается.

Таким образом, при проверке статистических гипотез существует вероятность допустить две ошибки: или, соответственно:

- опровергнуть верную гипотезу – ошибка первого рода α ;
- принять ложную гипотезу – ошибка второго рода β .

Вероятность совершения ошибки первого рода называется уровнем значимости. Чаще принимаются уровни значимости $\alpha = 0,05; 0,01; 0,001$, которые называют пятипроцентным, однопроцентным и 0,1%-м [26].

При проверке гипотезы возможны четыре варианта исходов (табл. 7).

Таблица 7

Варианты исходов

Гипотеза H_0	Решение	Вероятность	Название
Верна	Принимается	$1-\alpha$	Доверительная вероятность
	Отвергается	α	Вероятность ошибки первого рода
Неверна	Принимается	β	Вероятность ошибки второго рода
	Отвергается	$1-\beta$	Мощность критерия

Для проверки гипотезы H_0 используют специально подобранную случайную величину – статистический критерий.

Различают следующие статистические критерии:

- значимости,
- согласия,
- проверки на однородность.

Проверка гипотезы на значимость предполагает проверку гипотезы о равенстве численных значений принятого закона распределения двух выборок из генеральных совокупностей случайных величин X и Y .

Проверка на однородность – это проверка гипотезы о том, что две (или более) выборки взяты из одной генеральной совокупности.

Проверка на согласие – проверка предположения о том, что исследуемая случайная величина подчиняется принятому закону распределения.

Существует несколько критериев сравнения законов распределения выборок, так называемых критериев согласия, в частности критерий χ^2 (Пирсона), Колмогорова, Смирнова и др.

Выбирается критерий для сравнения законов, вычисляется его значение и сравнивается с его критическим значением. Если значение критерия превосходит критическое, то изменение данных в выборке не соответствует принятому закону. Критические значения для каждого критерия проще всего определять по таблицам.

Критерий Пирсона является универсальным, т. к. применим для любых видов функции $F(x)$, даже при неизвестных значениях их параметров, что имеет место при анализе результатов механических испытаний. Значение критерия Пирсона для выборки определяется как:

$$\chi^2 = \sum_{i=1}^k \frac{(m_i - np_{i0})^2}{np_{i0}} = n \sum_{i=1}^k \frac{(W_i - p_{i0})^2}{p_{i0}},$$

где n – объем выборки; k – количество интервалов; m_i – количество значений физической величины X , попавших в тот или иной интервал (абсолютная частота); W_i – относительная частота; p_{i0} – теоретическое значение вероятности, вычисленное по формулам принятого закона распределения.

При $n \rightarrow \infty$ ($W_i \rightarrow p_i$):

$$\chi^2 = n \sum_{i=1}^k \frac{(p_i - p_{i0})^2}{p_{i0}}, \quad (51)$$

p_i – значение вероятности, определенное по эмпирическим данным.

Пример

Определить возможность описания данных выборки $n = 50$ определенным законом распределения случайной физической величины при доверительной вероятности 95 %. Данные выборки: 19,70; 18,72; 20,24; 21,28; 21,20; 21,73; 17,82; 19,77; 21,10; 18,91; 19,31; 20,31; 18,15; 19,02; 19,23; 21,08; 19,43; 20,60; 20,13; 20,23; 20,17; 19,63; 20,34; 19,91; 19,81; 19,49; 21,97; 20,87; 22,38; 19,35; 20,66; 21,83; 20,54; 20,90; 19,92; 19,92; 20,18; 20,68; 19,62; 20,76; 18,56; 19,15; 18,48; 19,64; 19,97; 20,03; 19,68; 20,49; 20,66; 19,26.

Решение

1. Определяем минимальное и максимальное значения выборки: $x_{\min} = 17,82$, $x_{\max} = 22,38$.

Принимаем количество интервалов $k = 8$ (см. табл. 6) и определяем ширину интервала (50).

$$\Delta x = \frac{22,38 - 17,82}{8} = 0,57.$$

Определим диапазон каждого интервала, абсолютную и относительную частоты (табл. 7).

Таблица 7

Параметры выборки

Номер интервала	Диапазон интервала	Среднее значение \bar{x}_i	Частота	
			абсолютная m_i	относительная W_i
1	17,82...18,39	18,11	2	0,040
2	18,39...18,96	18,68	4	0,080
3	18,96...19,53	19,25	8	0,160
4	19,53...20,1	19,82	12	0,240
5	20,1...20,67	20,39	12	0,240
6	20,67...21,24	20,96	7	0,140
7	21,24...21,81	21,53	3	0,060
8	21,81...22,38	22,10	2	0,040

По данным в таблице строим гистограмму (рис. 12).

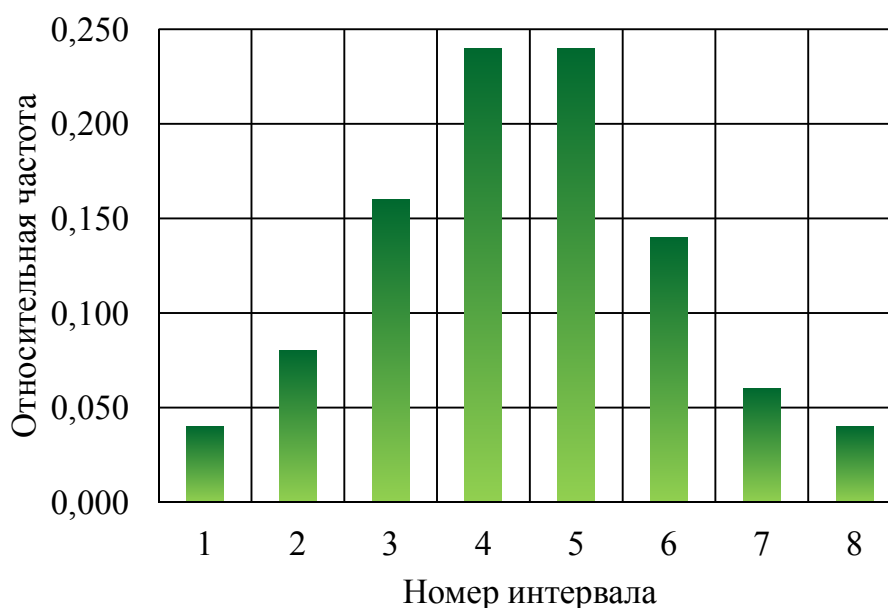


Рис. 12. Гистограмма распределения

2. Выдвижение предположения о законе распределения физической величины

По внешнему виду гистограммы (рис. 12) выдвигаем предположение о нормальном законе распределения физической величины, согласно

которому (29), (30) изменение физической величины зависит от двух параметров: математического ожидания $M(x)$ и среднеквадратического отклонения $\sigma(x)$.

Определим математическое ожидание и среднеквадратичное отклонение.

Данные параметры можно определить, воспользовавшись формулами (9), (11) и (13), сделав допущение о том, что $W_i \rightarrow p_i$, а $x_i \approx \bar{x}_i$ для интервала:

$$M(x) = \sum_{i=1}^k \bar{x}_i W_i \approx 20,05;$$

$$D(x) = \sum_{i=1}^n (\bar{x}_i - M(x))^2 \cdot W_i = 0,859;$$

$$\sigma(x) = \sqrt{D(x)} = 0,927.$$

По формуле (38) и табличным данным (прил. 1) определяем расчетные значения вероятности попадания значений физической величины в каждый интервал:

$$\begin{aligned} P(17,82 < X < 18,39) &= \Phi\left(\frac{18,39 - 20,05}{0,927}\right) - \Phi\left(\frac{17,82 - 20,05}{0,927}\right) = \\ &= \Phi(-1,79) - \Phi(-2,41) = \Phi(2,41) - \Phi(1,79) = 0,491 - 0,463 \approx 0,028; \end{aligned}$$

$$\begin{aligned} P(18,39 < X < 18,96) &= \Phi\left(\frac{18,96 - 20,05}{0,927}\right) - \Phi\left(\frac{18,39 - 20,05}{0,927}\right) = \\ &= \Phi(-1,18) - \Phi(-1,79) = \Phi(1,79) - \Phi(1,18) = 0,463 - 0,38 = 0,083; \end{aligned}$$

$$\begin{aligned} P(18,96 < X < 19,53) &= \Phi\left(\frac{19,53 - 20,05}{0,927}\right) - \Phi\left(\frac{18,96 - 20,05}{0,927}\right) = \\ &= \Phi(-0,56) - \Phi(-1,18) = \Phi(1,18) - \Phi(0,56) = 0,381 - 0,212 = 0,169; \end{aligned}$$

$$\begin{aligned} P(19,53 < X < 20,1) &= \Phi\left(\frac{20,1 - 20,05}{0,927}\right) - \Phi\left(\frac{19,53 - 20,05}{0,927}\right) = \\ &= \Phi(0,05) - \Phi(-0,56) = 0,02 + 0,212 = 0,232; \end{aligned}$$

$$\begin{aligned} P(20,1 < X < 20,67) &= \Phi\left(\frac{20,67 - 20,05}{0,927}\right) - \Phi\left(\frac{20,1 - 20,05}{0,927}\right) = \\ &= \Phi(0,67) - \Phi(0,05) = 0,249 - 0,02 = 0,229; \end{aligned}$$

$$P(20,67 < X < 21,24) = \Phi\left(\frac{21,24 - 20,05}{0,927}\right) - \Phi\left(\frac{20,67 - 20,05}{0,927}\right) = \\ = \Phi(1,28) - \Phi(0,67) = 0,4 - 0,249 = 0,151;$$

$$P(21,24 < X < 21,81) = \Phi\left(\frac{21,81 - 20,05}{0,927}\right) - \Phi\left(\frac{21,24 - 20,05}{0,927}\right) = \\ = \Phi(1,9) - \Phi(1,28) = 0,471 - 0,4 = 0,071;$$

$$P(21,81 < X < 22,38) = \Phi\left(\frac{22,38 - 20,05}{0,927}\right) - \Phi\left(\frac{21,81 - 20,05}{0,927}\right) = \\ = \Phi(2,51) - \Phi(1,9) = 0,494 - 0,471 = 0,023.$$

Рассчитав погрешность (2), можно сделать вывод о близости статистического распределения к нормальному закону (табл. 8).

Таблица 8

Сравнение расчетных и эмпирических вероятностей

Номер интервала	Среднее значение \bar{x}_i	Вероятность		Погрешность $ \delta $
		расчетная	эмпирическая	
1	18,11	0,028	0,040	30
2	18,68	0,083	0,080	3,75
3	19,25	0,169	0,160	5,63
4	19,82	0,232	0,240	3,33
5	20,39	0,229	0,240	4,58
6	20,96	0,151	0,140	7,86
7	21,53	0,071	0,060	18,33
8	22,10	0,023	0,040	42,5

Справедливость выдвинутого предположения проверим с помощью критерия согласия Пирсона.

3. Проверка правдоподобия гипотезы

Для проверки гипотезы рассчитаем эмпирическое значение коэффициента Пирсона (51)

$$\chi^2_3 = 1,08.$$

Для определения критического значения коэффициента Пирсона (прил. 2) найдем число степеней свободы df [27]:

$$df = k - 1 - r, \quad (52)$$

где k – количество интервалов выборки (количество варьируемых параметров); r – число параметров предполагаемого распределения, которые оценены по данным выборки:

- $r = 2$ – для нормального закона распределения (математическое ожидание и среднеквадратичное отклонение), $df = k - 3$;
- $r = 1$ – для показательного закона распределения (параметр распределения λ), $df = k - 2$. Исходя из числа степеней свободы $df = 5$, по табличным данным (прил. 2) определяем критическое значение критерия Пирсона:

$$\chi_{кр}^2 = 11,07.$$

Так как $\chi_{\text{э}}^2 < \chi_{кр}^2$, то гипотезу о нормальном распределении данных в выборке можно принять.

4.2. Определение функции отклика

При изучении любых объектов необходимо знание закономерностей протекания в них процессов. В тех случаях, когда механизм протекающих в объекте процессов неизвестен, изучают зависимость выходного параметра объекта от изменения факторов, т. е. стараются определить функцию отклика или, другими словами, разрабатывают эмпирическую модель. Изучаемый объект в этом случае представляется в виде «черного ящика» (рис. 12).

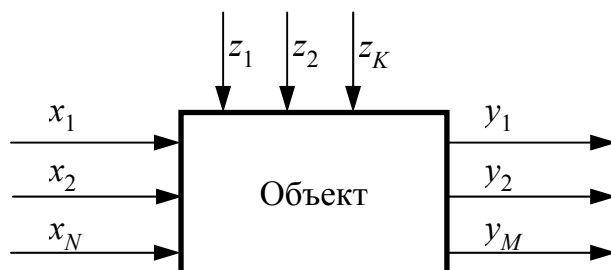


Рис. 12. Схематичное представление исследуемого объекта:
 x_i – входные параметры (факторы); y_i – выходные параметры; z_i – помехи

Для получения эмпирических зависимостей собирают статистические данные о состоянии (поведении) объекта, поэтому модели называются статистическими.

Математической моделью объекта будет функция отклика вида:

$$y = f(x_1, x_2, \dots, x_N, b_1, b_2, \dots, b_N), \quad (53)$$

где y – выходной параметр процесса; x_1, \dots, x_N – независимые переменные (факторы), варьируемые при постановке эксперимента; b_1, \dots, b_N – коэффициенты эмпирической модели.

При неполном знании механизма изучаемых явлений, аналитическое выражение функции отклика неизвестно, поэтому математическая модель представляется в виде полинома – уравнения регрессии:

$$y = \beta_0 + \sum_{i=1}^n \beta_i x_i + \sum_{\substack{i,j=1 \\ i \neq j}}^n \beta_{ij} x_i x_j + \sum_{i=1}^n \beta_{ii} \cdot x_i^2 + \dots, \quad (54)$$

где β_i , β_{ij} , β_{ii} – теоретические коэффициенты (коэффициенты регрессии), характеризующие соответственно линейные эффекты, эффекты взаимодействия и квадратичные эффекты.

Так как результат эксперимента величина случайная и значения выходных параметров отличаются друг от друга, определить можно только так называемые выборочные коэффициенты регрессии: b_0 , b_i , b_{ij} , b_{ii} , ..., которые являются лишь оценками для теоретических коэффициентов регрессии, которые можно было бы получить для генеральной совокупности.

В результате пользуются приближенным уравнением регрессии для выборки экспериментальных данных:

$$\hat{y} = b_0 + \sum_{i=1}^n b_i x_i + \sum_{i,j=1}^n b_{ij} x_i x_j + \sum_{i=1}^n b_{ii} x_i^2 + \dots, \quad (55)$$

где \hat{y} – выборочная оценка для y (предсказанное значение выходного параметра); b_0 – свободный член уравнения регрессии; b_i , b_{ij} , b_{ii} – коэффициенты регрессии.

Для получения конкретного вида эмпирической модели (53) необходимо выполнить следующее:

- определить степень тесноты взаимосвязи между x и y (корреляционный анализ);
- найти конкретный вид функции (54, 55) и определить значения коэффициентов регрессии b_i (регрессионный анализ);
- выполнить статистический анализ полученных результатов.

4.2.1. Корреляционный анализ

Корреляционный анализ применяется для выявления наличия взаимосвязи между входными величинами (факторами) и выходным параметром, т. е. при изменении одной величины по определенному закону меняется другая.

Для количественной оценки тесноты связи служит выборочный коэффициент корреляции. Выделяют три типа коэффициента корреляции:

- парной корреляции;

- частной корреляции;
- множественной корреляции.

1. Коэффициент парной корреляции определяет тесноту зависимости между двумя переменными x или y :

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1) \cdot \sigma_x \cdot \sigma_y}, \quad (56)$$

где \bar{x} , \bar{y} – среднеарифметические значения переменных x , y ; n – число опытов; σ_x , σ_y – среднеквадратические отклонения случайных величин:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}; \quad \sigma_y = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}}. \quad (57)$$

Значение коэффициента корреляции меняется в пределах от +1 до –1 (табл. 9).

Таблица 9

Значение коэффициента корреляции

Значение	Интерпретация
$ r \leq 0,2$	Очень слабая корреляция
$0,2 < r \leq 0,5$	Слабая корреляция
$0,5 < r \leq 0,7$	Средняя корреляция
$0,7 < r \leq 0,9$	Высокая корреляция
$0,9 < r \leq 1,0$	Очень высокая корреляция

При линейной корреляции случайные величины x и y связаны точной линейной функциональной зависимостью:

$$\hat{y} = b_0 + b_1 x,$$

а коэффициент корреляции принимает значение $r_{xy} = \pm 1$, при этом знак коэффициента корреляции соответствует знаку коэффициента b_1 .

О наличии (отсутствии) корреляции между величинами качественно можно судить по виду поля корреляции (рис. 13).

Прямая линия или кривая, проведенная через область значений (рис. 13, а, б, в), получила название аппроксимирующей.

2. Коэффициент частной корреляции определяет линейную зависимость между двумя переменными после устранения части зависимости.

При исследовании зависимости y от x_1 и x_2 наличие корреляции между x_1 и x_2 и между y и x_2 будет влиять на корреляцию между y и x_1 . Для устранения влияния x_2 необходимо измерить корреляцию между y и x_1 при $x_2 = \text{const}$.

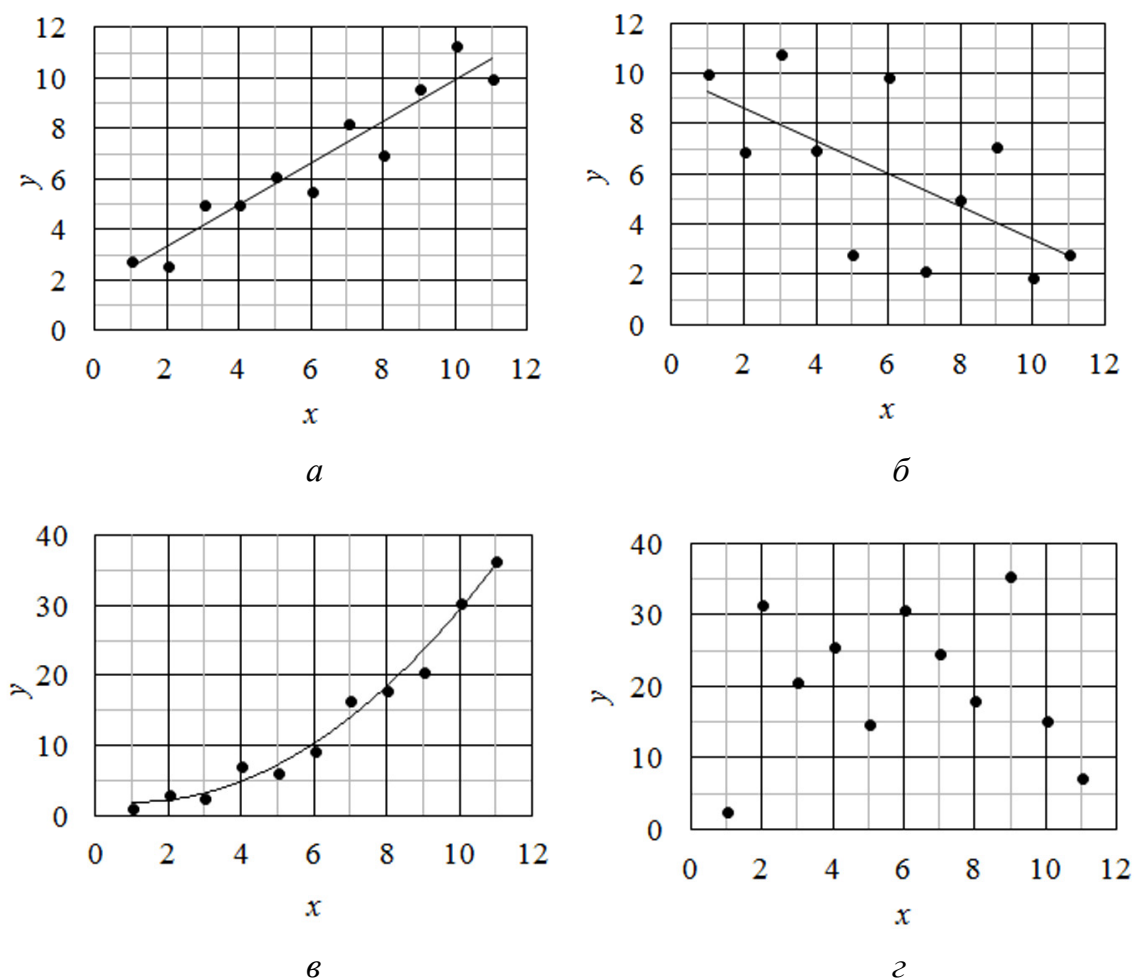


Рис. 13. Поля корреляции:

а – сильная положительная линейная корреляция; б – слабая отрицательная корреляция; в – сильная нелинейная корреляция; г – корреляции нет

Частный коэффициент $r_{yx_1x_2}$ оценивает степень влияния фактора x_1 на y при условии, что влияние x_2 на y исключено:

$$r_{yx_1x_2} = \frac{r_{yx_1} - r_{yx_2} \cdot r_{x_1x_2}}{\sqrt{(1 - r_{yx_2}^2)(1 - r_{x_1x_2}^2)}}. \quad (58)$$

Частный коэффициент $r_{yx_2x_1}$ оценивает степень влияния фактора x_2 на y при условии, что влияние x_1 на y исключено:

$$r_{yx_2x_1} = \frac{r_{yx_2} - r_{yx_1} \cdot r_{x_1x_2}}{\sqrt{(1 - r_{yx_1}^2)(1 - r_{x_1x_2}^2)}}. \quad (59)$$

3. Множественный коэффициент корреляции определяет тесноту связи одной переменной от нескольких:

$$r = \sqrt{1 - \frac{D_{y_{\text{ост}}}}{D_y}}, \quad (60)$$

где D_y – общая дисперсия эмпирического ряда; $D_{y_{\text{ост}}}$ – остаточная дисперсия в ряду y .

Коэффициент множественной корреляции принимает только положительные значения в пределах от 0 до 1.

Коэффициент корреляции демонстрирует существование или отсутствие связи между величинами, но не дает вид функциональной зависимости (53). Для ее определения требуется провести регрессионный анализ, т. е. определить вид уравнения регрессии (55).

4.2.2. Регрессионный анализ

Регрессионный анализ проводится после того, как установлена связь между двумя величинами, с целью нахождения уравнения регрессии (55) – вида статистической математической модели. Сложность полинома оказывается достаточной для моделирования большого числа реальных объектов исследования.

В случаях, когда степень полинома заранее предсказать нельзя, пользуются пошаговым поиском, т. е. сначала пытаются описать объект линейной моделью, проверяют ее на адекватность, если модель оказывается неадекватной, то увеличивают число членов полинома, повышают его степень, затем вводят нелинейную связь и т. д.

Вид уравнения регрессии зависит от выбираемого метода приближения. Существует несколько методов:

- опорных точек,
- медианных центров,
- наименьших квадратов и т. д.

Метод опорных точек состоит в том, что строят диаграмму разброса (рассеяния), выбирают, по мнению исследователя, наиболее характерные точки, через которые проводят аппроксимирующую кривую (рис. 14). Точки, отстоящие далеко от данной кривой, рассматриваются как грубые промахи и при анализе данных не рассматриваются.

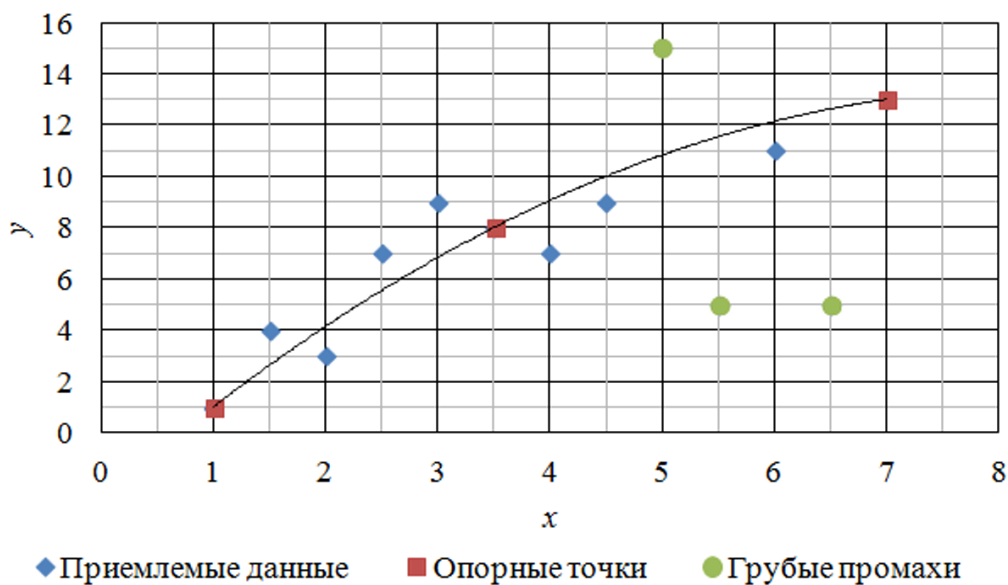


Рис. 14. Диаграмма разброса экспериментальных данных

По методу медианных центров область значений делится на равные интервалы (рис. 15), количество которых определяется исходя из принятого числа коэффициентов регрессии. Медианный центр интервала определяется как точка пересечения горизонтальной и вертикальной линий, по обе стороны от которых остается равное количество точек. Аппроксимирующую кривую строят, ориентируясь на медианные центры.

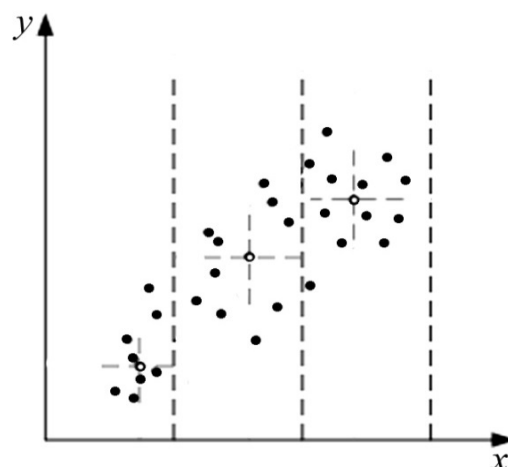


Рис. 15. Метод медианных центров

Чаще пользуются методом наименьших квадратов.

4.2.2.1. Метод наименьших квадратов

По методу наименьших квадратов уравнение, определяющее связь между входными и выходными параметрами, ищется в виде полинома:

$$P_N(x) = b_0 + b_1x + b_2x^2 + \dots + b_Nx^N, \quad (61)$$

где b_i – корни полинома; N – степень полинома.

Коэффициенты полинома (61) определяются из условия, по которому сумма квадратов отклонений $e_i = y_i - f_i(x)$ (рис. 16) стремится к минимуму

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - f(x_i))^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \rightarrow \min, \quad (62)$$

где y_i – экспериментальные значения; $f(x_i)$, \hat{y}_i – расчетные значения.

Так как каждое значение x_i в общем случае «сопровождается» соответствующим коэффициентом b_i ($i = 0, 1, 2, \dots, N$), то задача сводится к нахождению данных коэффициентов.

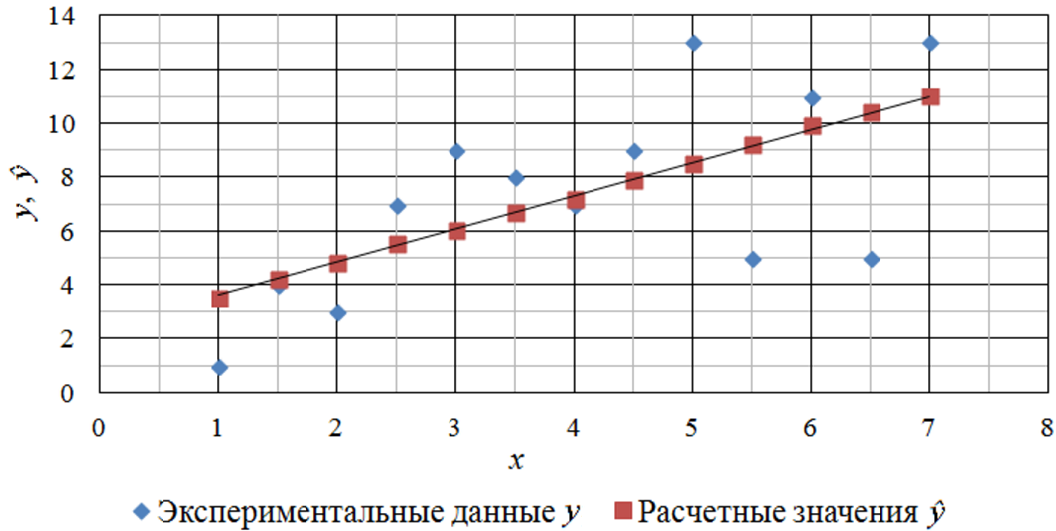


Рис. 16. Пример применения метода наименьших квадратов для линейной корреляции

В связи с тем, что погрешность e_i зависит от x_i и a_i , можно ввести обозначение функции

$$F(x_i, b_i) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (63)$$

Коэффициенты b_i определяются из условия равенства нулю производных функции $F(x_i, b_i)$, что соответствует минимальному значению функции, для чего составляется система уравнений:

$$\begin{cases} \frac{dF}{db_0} = 0; \\ \frac{dF}{db_1} = 0; \\ \dots \\ \frac{dF}{db_N} = 0. \end{cases} \quad (64)$$

В результате для полинома степени N система уравнений (64) примет вид:

[illegible]

Существенным недостатком метода является громоздкость вычислений при большом количестве коэффициентов, вследствие чего к нему прибегают при необходимости получения очень точных значений функции, а на практике чаще пользуются более простыми моделями.

4.2.2.2. Парная линейная регрессионная модель

Уравнение парной линейной регрессии (полиномиальная второй степени):

$$\hat{y} = b_0 + b_1 x. \quad (66)$$

Для ее получения пользуются методом наименьших квадратов (63, 65)

$$F(x_i, b_i) = \sum_{i=1}^n (y_i - b_1 x_i - b_0)^2 \rightarrow \min; \quad (67)$$

$$\begin{cases} \frac{dF}{db_0} = 0; \\ \frac{dF}{db_1} = 0. \end{cases} \quad (68)$$

Подставляя (67) в формулу (66), получаем

$$\begin{cases} \frac{dF}{db_0} = -2 \cdot \sum_{i=1}^n (y_i - b_1 x_i - b_0) \cdot 1 = 0; \\ \frac{dF}{db_1} = -2 \cdot \sum_{i=1}^n (y_i - b_1 x_i - b_0) \cdot x_i = 0; \\ \begin{cases} b_0 n + b_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i; \\ b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n (x_i y_i). \end{cases} \end{cases} \quad (69)$$

Решая систему (69) методом подстановки, получим формулы для нахождения коэффициентов:

$$b_1 = \frac{n \cdot \sum_{i=1}^n (x_i \cdot y_i) - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{n \cdot \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}, \quad (70)$$

$$b_0 = \frac{\sum_{i=1}^n y_i - b_1 \sum_{i=1}^n x_i}{n} = \frac{\sum_{i=1}^n y_i \cdot \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n (x_i \cdot y_i)}{n \cdot \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}. \quad (71)$$

Ниже приведена диаграмма разброса экспериментальных данных с аппроксимирующей прямой.

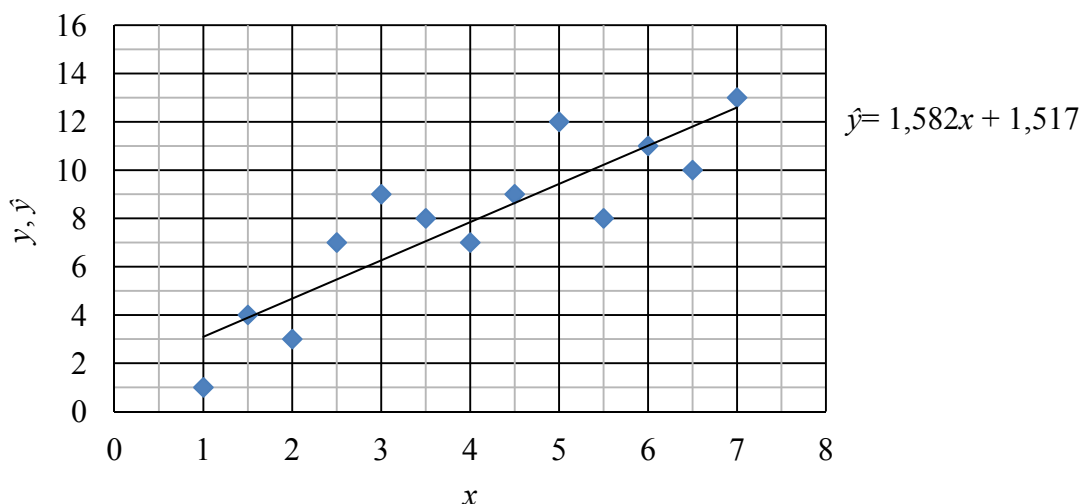


Рис. 17. Пример парной линейной регрессии (аппроксимации)

При невозможности описать связь между x и y уравнением парной линейной регрессии применяют уравнения с нелинейными связями, повышают степень полинома.

4.2.2.3. Параболическая регрессионная модель

Параболическая регрессионная модель – уравнение полинома второй степени:

$$\hat{y} = b_0 + b_1x + b_2x^2. \quad (72)$$

Коэффициент b_i определяется по методу наименьших квадратов

$$F(x_i, b_i) = \sum_{i=1}^n (y_i - b_0 - b_1x_i - b_2x_i^2)^2 \rightarrow \min. \quad (73)$$

Согласно (64), (72) и (73) составляем систему уравнений:

$$\begin{cases} \frac{dF}{db_0} = -2 \cdot \sum_{i=1}^n (y_i - b_0 - b_1 x_i - b_2 x_i^2) \cdot 1 = 0; \\ \frac{dF}{db_1} = -2 \cdot \sum_{i=1}^n (y_i - b_0 - b_1 x_i - b_2 x_i^2) \cdot x_i = 0; \\ \frac{dF}{db_2} = -2 \cdot \sum_{i=1}^n (y_i - b_0 - b_1 x_i - b_2 x_i^2) \cdot x_i^2 = 0. \end{cases} \quad (74)$$

Преобразовывая (74), получим систему с неизвестными b_0, b_1, b_2 :

$$\begin{cases} b_0 n + b_1 \sum_{i=1}^n x_i + b_2 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i, \\ b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2 + b_2 \sum_{i=1}^n x_i^3 = \sum_{i=1}^n (x_i \cdot y_i), \\ b_0 \sum_{i=1}^n x_i^2 + b_1 \sum_{i=1}^n x_i^3 + b_2 \sum_{i=1}^n x_i^4 = \sum_{i=1}^n (x_i^2 \cdot y_i). \end{cases} \quad (75)$$

Введем обозначения:

$$\begin{aligned} S_1 &= \sum_{i=1}^n x_i; \quad S_2 = \sum_{i=1}^n x_i^2; \quad S_3 = \sum_{i=1}^n x_i^3; \quad S_4 = \sum_{i=1}^n x_i^4; \\ S_5 &= \sum_{i=1}^n y_i; \quad S_6 = \sum_{i=1}^n (x_i \cdot y_i); \quad S_7 = \sum_{i=1}^n (x_i^2 \cdot y_i). \end{aligned}$$

С учетом принятых обозначений система (75) примет вид:

$$\begin{cases} b_0 n + b_1 S_1 + b_2 S_2 = S_5, \\ b_0 S_1 + b_1 S_2 + b_2 S_3 = S_6, \\ b_0 S_2 + b_1 S_3 + b_2 S_4 = S_7. \end{cases} \quad (76)$$

Коэффициенты b_0, b_1, b_2 системы (76) найдутся методом Крамера:

$$b_0 = \frac{\Delta_0}{\Delta}, \quad b_1 = \frac{\Delta_1}{\Delta}, \quad b_2 = \frac{\Delta_2}{\Delta},$$

где

$$\Delta = \begin{vmatrix} n & S_1 & S_2 \\ S_1 & S_2 & S_3 \\ S_2 & S_3 & S_4 \end{vmatrix}; \quad \Delta_0 = \begin{vmatrix} S_5 & S_1 & S_2 \\ S_6 & S_2 & S_3 \\ S_7 & S_3 & S_4 \end{vmatrix}; \quad \Delta_1 = \begin{vmatrix} n & S_5 & S_2 \\ S_1 & S_6 & S_3 \\ S_2 & S_7 & S_4 \end{vmatrix}; \quad \Delta_2 = \begin{vmatrix} n & S_1 & S_5 \\ S_1 & S_2 & S_6 \\ S_2 & S_3 & S_7 \end{vmatrix}.$$

Пример параболической аппроксимации приведен ниже (рис. 18).

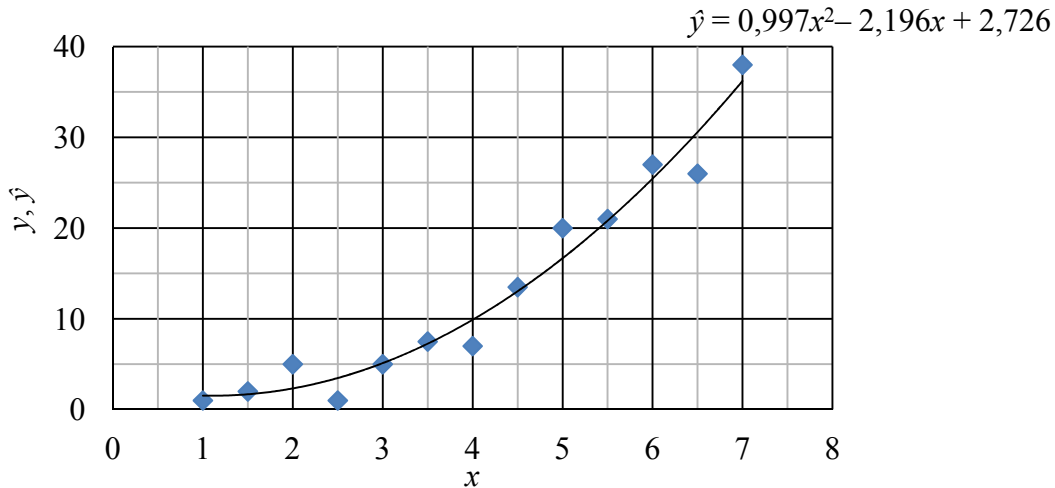


Рис. 18. Пример параболической регрессии (аппроксимации)

4.2.2.4. Регрессия в виде показательной функции

В ряде случаев возникает необходимость воспользоваться зависимостью вида

$$\hat{y} = b_0 e^{b_1 x}. \quad (77)$$

Прологарифмировав уравнение (77), получим

$$\ln \hat{y} = \ln b_0 + b_1 x.$$

Введя обозначения:

$$\hat{Y} = \ln \hat{y}, \quad B_0 = \ln b_0, \quad B_1 = b_1,$$

получим линейный полином первой степени:

$$\hat{Y} = B_0 + B_1 x.$$

Далее уравнение решается по методу наименьших квадратов

$$F = \sum_{i=1}^n (Y_i - (B_0 + B_1 x_i))^2 \rightarrow \min.$$

Формулы для вычисления коэффициентов B_0 и B_1 аналогичны как для случая линейной аппроксимации (70), (71):

$$B_1 = \frac{n \sum_{i=1}^n (x_i Y_i) - \sum_{i=1}^n x_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}, \quad B_0 = \frac{\sum_{i=1}^n Y_i - B_1 \sum_{i=1}^n x_i}{n}.$$

Коэффициенты определяются как: $b_1 = e^{B_1}$, $b_0 = B_0$.

4.2.2.5. Регрессия в виде степенной функции

Степенная функция имеет вид

$$\hat{y} = b_0 x^{b_1}. \quad (78)$$

Логарифмируя уравнение (78), получим

$$\lg \hat{y} = \lg b_0 + b_1 \lg x.$$

Введя обозначения:

$$Y = \lg \hat{y}; \quad B_0 = \lg b_0; \quad B_1 = b_1; \quad X = \lg x,$$

получим линейный полином первой степени

$$\hat{Y} = B_0 + B_1 x.$$

Коэффициенты находятся по методу наименьших квадратов:

$$B_1 = \frac{n \sum_{i=1}^n (X_i Y_i) - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2}, \quad B_0 = \frac{\sum_{i=1}^n Y_i - B_1 \sum_{i=1}^n X_i}{n}.$$

Определим коэффициенты:

$$b_0 = 10^{B_0}, \quad b_1 = B_1.$$

Пример регрессии в виде степенной функции приведен ниже (рис. 18).

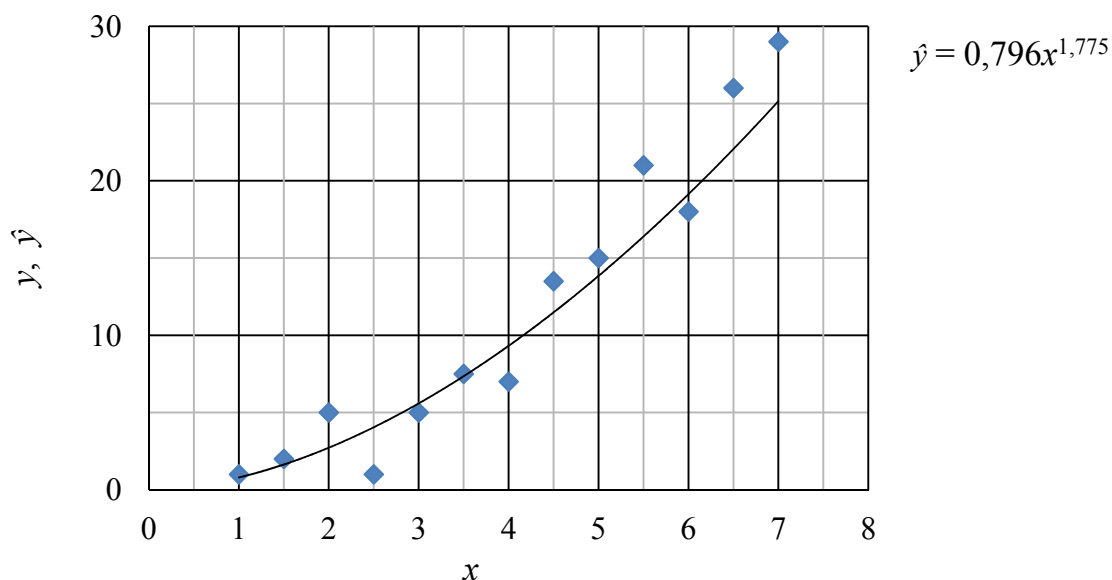


Рис. 19. Регрессия в виде степенной функции

4.2.3. Проверка математической модели на адекватность

Проверка модели на адекватность производится в несколько шагов, которые рассмотрим на примере парной линейной регрессии (66).

1. При наличии параллельных опытов считается среднее значение

$$\bar{y}_i = \frac{\sum_{u=1}^m y_{iu}}{m}, \quad (79)$$

где m – число параллельных опытов; u – порядковый номер параллельного опыта; i – порядковый номер опыта в выборке размером n , $i = 1 \dots n$.

2. Проверка однородности дисперсий по критерию Кохрена (G -критерий).

Определение выборочных дисперсий (для $i = 1 \dots n$):

$$D_i = \frac{\sum_{u=1}^m (y_{iu} - \bar{y}_i)^2}{m - 1}. \quad (80)$$

Определение значения коэффициента Кохрена

$$G = \frac{D_{\max}}{\sum_{i=1}^n D_i}, \quad (81)$$

где D_{\max} – максимальное значение выборочной дисперсии (80).

Принимается уровень значимости α , и определяются числа степени свободы соответственно числителя df_1 и знаменателя df_2 :

$$df_1 = m - 1; \quad df_2 = n. \quad (82)$$

Расчетное значение коэффициента (81) сравнивается с табличным (прил. 3). При $G < G_{\text{табл}}$ выборочные дисперсии D_i однородны, и результаты параллельных опытов можно считать выборками из одной генеральной совокупности.

3. Проверка коэффициентов уравнения регрессии на значимость по критерию Стьюдента (t -критерий).

Определение дисперсии воспроизводимости

$$D_{\text{воспр}} = \frac{\sum_{i=1}^n D_i}{n(m - 1)}, \quad (83)$$

где $n(m - 1)$ – число степеней свободы для одинакового числа опытов m .

Определение критерия Стьюдента

$$t_{b_i} = \frac{|b_i|}{\sigma_{b_i}}, \quad (84)$$

где b_i – i -й коэффициент уравнения регрессии; σ_{b_i} – среднеквадратичное отклонение i -го коэффициента.

Для уравнения парной линейной регрессии среднеквадратичное отклонение определится по нижеприведенным формулам:

$$\sigma_{b_0} = \sqrt{\frac{D_{\text{воспр}} \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}}; \quad (85)$$

$$\sigma_{b_1} = \sqrt{\frac{n \cdot D_{\text{воспр}}}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}}. \quad (86)$$

Рассчитанное значение t -критерия (84) сравнивается с его критическим значением (прил. 4). Если $t_{b_i} > t_{\text{кр}}$, то коэффициент b_i значим, в противном случае коэффициент не значим и данным слагаемым можно пренебречь.

4. Проверка уравнения регрессии на адекватность по критерию Фишера (F -критерий).

В данном случае полагают, что уравнение регрессии адекватно описывает исследуемый процесс, т. е. статически значимо, и связь между y и x существенна, если выполняется условие

$$F = \frac{D_{\text{ост}}}{D_{\text{воспр}}} > F_{\text{кр}}, \quad (87)$$

где $F_{\text{кр}}$ – критическое значение критерия Фишера; $D_{\text{ост}}$ – остаточная дисперсия, показывающая вариацию результата под влиянием факторов, не учтенных регрессией:

$$D_{\text{ост}} = \frac{\sum_{i=1}^n (\bar{y}_i - \hat{y}_i)^2}{n - l}, \quad (88)$$

где \bar{y}_i – среднее значение выходного параметра по результатам параллельных опытов (79); \hat{y}_i – расчетное значение выходного параметра.

Если $F \leq F_{кр}$, то делается вывод, что уравнение регрессии статистически незначимо, т. к. при заданном уровне надежности существует риск сделать неправильный вывод о наличии связи между x и y .

Если при проведении эксперимента параллельные опыты не были предусмотрены, то выражение примет вид:

$$D_{ост} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - l}, \quad (89)$$

где y_i – экспериментальное значение выходного параметра.

Число степеней свободы для выбора $F_{кр}$ (прил. 5) определяется по зависимостям:

$$df_1 = n - l; \quad df_2 = n(m - 1),$$

где n – общее количество опытов; $l = N + 1$ – число членов аппроксимирующего полинома (число коэффициентов регрессии, включая свободный член); N – количество факторов (x_1, x_2, \dots).

4.3. Планирование эксперимента

Если пассивный эксперимент заключается в наблюдении за объектом исследования, фиксировании результатов с последующей обработкой полученных данных методами теории вероятности и математической статистики, то при проведении активного эксперимента необходимо составить его план, т. е. выбрать число и условия проведения опытов, достаточных для решения поставленной задачи с требуемой точностью.

В зависимости от степени уравнения регрессии, которым в конечном итоге опишется состояние объекта, составляются планы первого, второго и третьего порядков.

При проведении активного эксперимента необходимо определиться с видом проводимого эксперимента: полный или дробный факторный.

В учебных целях рассмотрим планы первого порядка, при составлении которых получают статистическую математическую модель вида:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \dots + b_Nx_N. \quad (90)$$

4.3.1. Полный факторный эксперимент

При планировании эксперимента по схеме полного факторного эксперимента реализуются все возможные комбинации факторов на всех выбранных для исследования уровнях факторов.

Уровни факторов – это границы исследуемой области по данному технологическому параметру.

Необходимое количество опытов n при полном факторном эксперименте определяется по формуле:

$$n = l^N, \quad (91)$$

где N – число факторов; l – число уровней, на которых варьируются факторы.

Обычно применяется планирование на двух уровнях, т. е. $l = 2$, тогда, например, при $N = 2$, число опытов будет

$$n = 2^2 = 4.$$

Основной (нулевой) уровень (центр плана эксперимента) – это некоторое начальное значение фактора при составлении математической модели, это точка с координатами (x_1^0, \dots, x_N^0) .

Интервал варьирования – часть области определения фактора, симметричная относительно его основного уровня.

4.3.1.1. Матрица планирования

Составление матрицы планирования рассмотрим на примере.

Пусть на процесс влияют два фактора: x_1 и x_2 . Известно, фактор x_1 изменяется в интервале $100 \dots 200$, а фактор x_2 – в интервале $50 \dots 100$.

Для двух факторов без учета их взаимодействия соответствующая модель (90) может быть записана в виде:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2. \quad (92)$$

Центр плана эксперимента определится:

$$x_1^0 = \frac{x_1^{\min} + x_1^{\max}}{2} = \frac{200 + 100}{2} = 150;$$

$$x_2^0 = \frac{x_2^{\min} + x_2^{\max}}{2} = \frac{100 + 50}{2} = 75.$$

Интервалы варьирования:

$$\Delta x_1 = \frac{x_1^{\max} - x_1^{\min}}{2} = \frac{200 - 100}{2} = 50;$$

$$\Delta x_2 = \frac{x_2^{\max} - x_2^{\min}}{2} = \frac{100 - 50}{2} = 25.$$

Определенные условия эксперимента сведем в таблицу (табл. 10).

Графически условия эксперимента можно представить в виде схемы полного факторного эксперимента 2^2 (рис. 20). Схема эксперимента указывает расположение в пространстве опытных точек независимых переменных или условия всех опытов, которые необходимо провести.

Таблица 10

Условия эксперимента

Уровни факторов	Факторы	x_1	x_2
Верхний уровень		200	30
Нижний уровень		100	20
Основной уровень		150	25

При полном факторном эксперименте эксперимент ставится только на границе области с центром в точке A . В большинстве случаев эксперимент задается в виде *матрицы планирования* (плана эксперимента), каждая строчка которой представляет собой условия опыта, а каждый столбец матрицы соответствует значениям переменных в различных опытах (табл. 11).

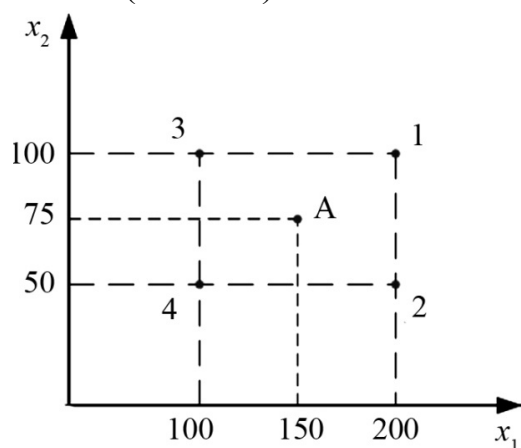


Рис. 20. Схема расположения опытных точек в натуральном пространстве

Таблица 11

Матрица планирования

№ опыта	Факторы эксперимента		Выходной параметр
	x_1	x_2	y
1	200	100	y_1
2	200	100	y_2
3	100	50	y_3
4	100	50	y_4

4.3.1.2. Кодирование переменных

Для упрощения записи условий эксперимента и удобства расчетов от натуральных единиц измерения переходят к безразмерным, т. е. производят кодирование факторов:

$$X_i^{B/H} = \frac{x_i^{B/H} - x_i^0}{\Delta x_i}, \quad (93)$$

где $x_i^{B/H}$ – значения верхнего или нижнего уровня переменной в натуральных единицах; x_i^0 – основной уровень натуральной переменной; Δx_i – интервал варьирования натуральной переменной; $X_i^{B/H}$ – кодированное значение i -го фактора (на верхнем или нижнем уровне).

Перейдем в нашем примере от натуральных переменных к кодированным.

Для фактора x_1 :

$$X_1^B = \frac{200 - 150}{50} = +1; \quad X_1^H = \frac{100 - 150}{50} = -1;$$

для фактора x_2 :

$$X_2^B = \frac{100 - 75}{25} = +1; \quad X_2^H = \frac{50 - 75}{25} = -1.$$

Представим матрицу планирования в кодированных единицах в двух вариантах (табл. 12) с введением в нее фиктивной переменной x_0 , необходимой для расчета свободного коэффициента уравнения b_0 (92).

Таблица 12

Матрицу планирования в кодированных единицах

№ опыта	x_0		x_1		x_2	
1	+1	+	+1	+	+1	+
2	+1	+	+1	+	-1	-
3	+1	+	-1	-	+1	+
4	+1	+	-1	-	-1	-

При этом натуральное пространство (рис. 21) преобразуется в факторное (рис. 22).

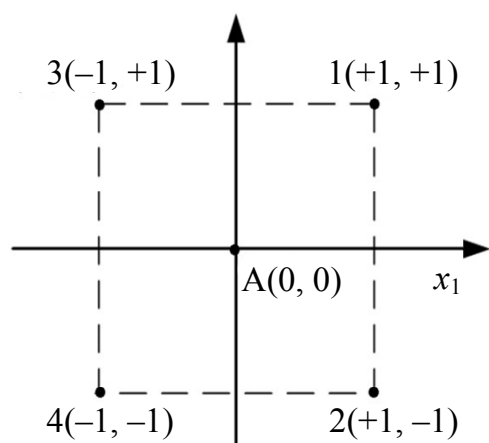


Рис. 21. Схема расположения опытных точек в факторном пространстве

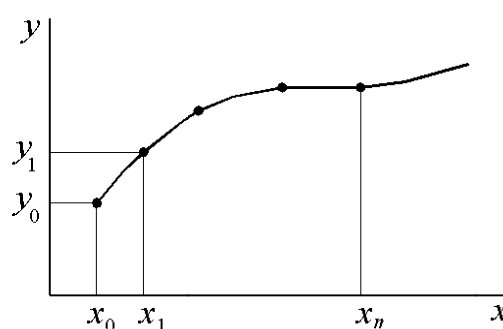


Рис. 22. Вид интерполирующей функции

4.3.1.3. Свойства матрицы планирования

Матрица планирования в безразмерных единицах обладает следующими свойствами:

- ортогональность – скалярное произведение двух любых столбцов матрицы равно нулю:

$$\sum_{i=1}^n x_{ui} \cdot x_{ji} = 0, \quad (94)$$

где i – номер опыта, $i = 1, \dots, n$; n – количество опытов; u, j – номер фактора, $u, j = 1, \dots, N$, $u \neq j$, N – количество факторов;

- симметричность – сумма элементов всех столбцов матрицы кроме первого равна нулю:

$$\sum_{i=1}^n x_{ui} = 0; \quad (95)$$

- нормировка – сумма квадратов элементов каждого столбца равна числу опытов

$$\sum_{i=1}^n x_{ui}^2 = n; \quad (96)$$

- свойство ротабельности – точки в матрице планирования выбраны таким образом, что точность предсказания значений функции отклика одинакова на равных расстояниях от центра эксперимента.

На основании перечисленных свойств значительно упрощается расчет коэффициентов регрессии.

4.3.1.4. Расчет коэффициентов регрессии

На основании метода наименьших квадратов (62) с учетом свойств матрицы планирования (84–96) коэффициенты уравнения регрессии определяются по формулам:

$$b_0 = \frac{1}{n} \sum_{i=1}^n x_{0i} \cdot y_i; \quad (97)$$

$$b_i = \frac{1}{n} \sum_{i=1}^n x_{ui} \cdot y_i; \quad (98)$$

$$b_{ij} = \frac{1}{n} \sum_{i=1}^n x_{ui} \cdot x_{ji} \cdot y_i. \quad (99)$$

Пример

Вычислить коэффициенты уравнения регрессии для двухфакторного эксперимента при условии наличия корреляционной связи между факторами на основании экспериментальных данных (табл. 13).

Экспериментальные данные

№ опыта	x_0	x_1	x_2	x_1x_2	y
1	+1	1	1	1	38,2
2	+1	1	-1	-1	32,5
3	+1	-1	1	-1	26,8
4	+1	-1	-1	1	20,7

Уравнение регрессии (90) для условий эксперимента примет вид

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2.$$

Решение

Коэффициенты регрессии определим по формулам (97)–(99)

$$b_0 = \frac{+1 \cdot 38,2 + 1 \cdot 32,5 + 1 \cdot 26,8 + 1 \cdot 20,7}{4} = 29,55;$$

$$b_1 = \frac{+1 \cdot 38,2 + 1 \cdot 32,5 - 1 \cdot 26,8 - 1 \cdot 20,7}{4} = 5,8;$$

$$b_2 = \frac{+1 \cdot 38,2 - 1 \cdot 32,5 + 1 \cdot 26,8 - 1 \cdot 20,7}{4} = 2,95;$$

$$b_{12} = \frac{+1 \cdot 38,2 - 1 \cdot 32,5 - 1 \cdot 26,8 + 1 \cdot 20,7}{4} = -0,1.$$

С учетом полученных коэффициентов запишем уравнение:

$$\hat{y} = 29,55 + 5,8x_1 + 2,95x_2 - 0,1x_1x_2.$$

После вычисления коэффициентов регрессии приступают к проверке уравнения регрессии на адекватность (см. п. 4.2.3).

4.3.2. Дробный факторный эксперимент

При помощи полного факторного эксперимента эффективно разрабатываются математические модели исследуемого объекта. Недостаток данного вида эксперимента – его громоздкость, т. е. при увеличении числа факторов возрастает количество опытов. Так, при двухуровневом эксперименте и числе факторов, равном семи, количество опытов составит:

$$n = 2^7.$$

Для сокращения количества опытов пользуются дробным факторным экспериментом, идея которого заключается в сокращении числа опытов при сохранении матрицей планирования своих свойств (см. п. 4.3.1.3).

Дробный факторный эксперимент – система опытов, представляющая собой определенную часть полного факторного эксперимента,

а именно: $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$ и т. д. Эти системы опытов получили название *дробные реплики*.

При этом в качестве реплики берут ближайший полный факторный эксперимент, а число опытов должно быть больше или равно числу неизвестных коэффициентов в уравнении регрессии.

Сокращение числа опытов приводит к появлению корреляции между оценками коэффициентов, что не позволяет отдельно оценивать эффекты факторов и эффекты взаимодействия – получаются так называемые смешанные оценки.

Для дробных реплик используются специальные алгебраические соотношения, облегчающие выявление смешанных эффектов, – *генерирующие соотношения*. Генерирующие соотношения демонстрируют, какое из взаимодействий факторов принято незначимым в аспекте влияния на выходной параметр и может быть заменено в матрице планирования новой независимой переменной.

Рассмотрим на примере построение дробной реплики для полного факторного эксперимента $n = 2^3 = 8$, который гипотетически описывается линейным уравнением

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3.$$

Составим матрицу планирования для полного факторного эксперимента $n = 2^3$ (табл. 14).

Таблица 14

Матрица планирования

№ опыта	x_0	x_1	x_2	x_3
1	+	+	+	+
2	+	–	+	+
3	+	+	–	+
4	+	–	–	+
5	+	+	+	–
6	+	–	+	–
7	+	+	–	–
8	+	–	–	–

Допустим, что необходимо сократить число опытов при обязательном выполнении следующих условий:

- матрица планирования сохраняет свои свойства;
- число опытов не должно быть меньше числа коэффициентов линейной модели, т. е. четырех.

Ближайший в данном случае полный факторный эксперимент:

$$n = 2^2 = 4.$$

Полагают, что взаимодействие между факторами x_1 и x_2 в нем отсутствует, в связи с чем в качестве генерирующего соотношения (новой независимой переменной x_3) можно выбрать взаимодействие $+x_1x_2$:

$$x_3 = x_1x_2. \quad (100)$$

Число опытов в дробном факторном эксперименте определяется по формуле:

$$n = 2^{N-p}, \quad (101)$$

где n – общее число факторов; p – число факторов, приравненных к произведению факторов (число генерирующих соотношений).

Если при числе факторов, равном трем, фактор x_3 приравнен произведению x_1x_2 , то число опытов:

$$n = 2^{3-1} = 4.$$

Получается дробная реплика (полуреплика) от полного факторного эксперимента $n = 2^3$, сохраняющая все необходимые свойства (табл. 15).

Таблица 15

Матрица планирования первой полуреплики

№ опыта	x_0	x_1	x_2	$x_3 (x_1x_2)$
1	+	+	+	+
2	+	–	+	–
3	+	+	–	–
4	+	–	–	+

Если приравнять $x_3 = -x_1x_2$, то получим вторую половину матрицы полного факторного эксперимента (табл. 16).

Таблица 16

Матрица планирования второй полуреплики

№ опыта	x_0	x_1	x_2	$x_3 (x_1x_2)$
1	+	+	+	–
2	+	–	+	+
3	+	+	–	+
4	+	–	–	–

С генерирующими соотношениями можно производить алгебраические операции, а именно умножать левую и правую части на линейные эффекты, определенные взаимодействия. При этом если фактор

входит в уравнение в четной степени, то он заменяется единицей. Так, умножив обе части генерирующего соотношения (100) на x_3 , получим:

$$x_3^2 = x_1 x_2 x_3$$

или

$$1 = x_1 x_2 x_3. \quad (102)$$

Выражение (102) получило название *определяющий контраст*.

Зная определяющий контраст, можно получить систему смешанных оценок для данной дробной реплики, для чего определяющий контраст (102) умножается на каждый фактор и взаимодействие факторов.

В нашем примере

$$x_1 = x_1^2 x_2 x_3 = x_2 x_3, \quad x_2 = x_1 x_2^2 x_3 = x_1 x_3, \quad x_3 = x_1 x_2 x_3^2 = x_1 x_2,$$

что соответствует оценкам

$$b_1 \rightarrow \beta_1 + \beta_{23}, \quad b_2 \rightarrow \beta_2 + \beta_{13}, \quad b_3 \rightarrow \beta_3 + \beta_{12},$$

где β_1 – истинные коэффициенты; b_1 – оценки коэффициентов, вычисленных по данным выборки.

По плану дробного факторного эксперимента можно вычислить коэффициенты b_i , однако они будут смешаны с парными и тройными взаимодействиями. Таким образом, сокращение числа опытов приводит к корреляции между столбцами матрицы данного вида эксперимента.

При условии постулирования линейной модели все парные взаимодействия незначимы, что позволяет минимизировать число опытов: вместо 8 опытов для изучения трех факторов можно поставить четыре.

Пример

По экспериментальным данным (табл. 18) найти значение функции при $x=1$ и $x=3,2$. Решение проиллюстрировать графически.

Таблица 18

Экспериментальные данные

x	0	2	3	3,5
y	-1	0,2	0,5	0,8

Решение

1. Точка $x=1$ принадлежит первому локальному отрезку $[0, 2]$, т. е. $i=1$ и, следовательно, по вышеприведенным формулам (1, 2):

$$a_1 = \frac{y_1 - y_0}{x_1 - x_0} = \frac{0,2 - (-1)}{2 - 0} = 0,6; \quad b_1 = y_0 - a_1 x_0 = -1 - 0,6 \cdot 0 = -1;$$

$$y = a_1 x + b_1 = 0,6 \cdot 1 - 1 = -0,4.$$

Точка $x=3,2$ принадлежит третьему интервалу $[3, 3,5]$, т. е. $i=3$ и, следовательно, по формулам (1, 2):

$$a_3 = \frac{y_3 - y_2}{x_3 - x_2} = \frac{0,8 - 0,5}{3,5 - 3} = 0,6; \quad b_3 = y_2 - a_3 x_2 = 0,5 - 0,6 \cdot 3 = -1,3;$$

$$y = a_3 x + b_3 = 0,6 \cdot 3,2 - 1,3 = 0,62.$$

2. По данным табл. 18 строим график (рис. 23).

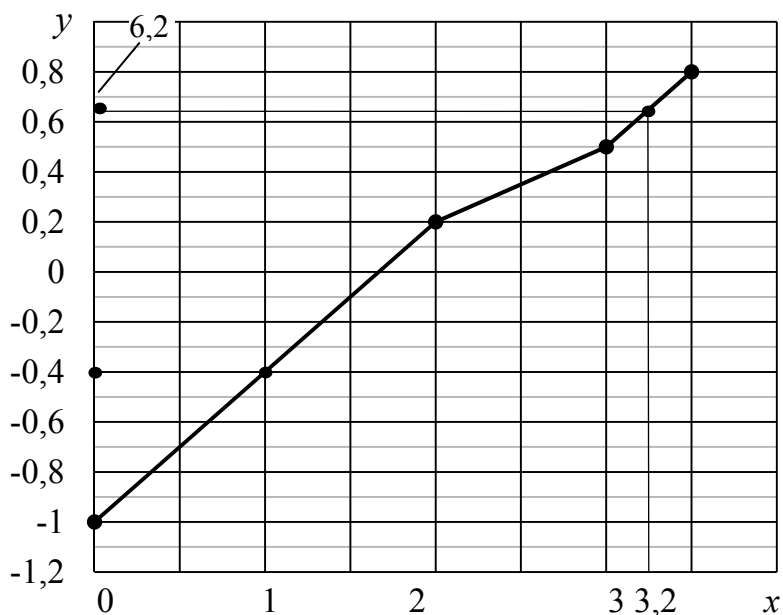


Рис. 23. Графическое решение поставленной задачи

5.1.3. Интерполяционный полином Лагранжа

Интерполяционный полином Лагранжа имеет вид:

$$P_n(x) = \sum_{i=0}^n y_i \cdot L_n(x), \quad (108)$$

где $L_n(x)$ – множитель Лагранжа

$$L_n(x) = \frac{(x-x_0)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_n)}{(x_i-x_0)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_n)} = \prod_{\substack{k=0 \\ k \neq i}}^n \frac{(x-x_k)}{(x_i-x_k)}.$$

Следовательно,

$$P_n(x) = \sum_{i=0}^n y_i \left(\prod_{\substack{k=0 \\ k \neq i}}^n \frac{x-x_k}{x_i-x_k} \right).$$

Числитель и знаменатель не включают в себя значения $x = x_i$, т. к. в этом случае результат будет равен нулю.

В развернутом виде формулу Лагранжа можно записать:

$$\begin{aligned} P_n(x) = & y_0 \frac{(x-x_1)(x-x_2)\dots(x-x_n)}{(x_0-x_1)(x_0-x_2)\dots(x_0-x_n)} + \\ & + y_1 \frac{(x-x_0)(x-x_2)\dots(x-x_n)}{(x_1-x_0)(x_1-x_2)\dots(x_1-x_n)} + \dots + \\ & + y_n \frac{(x-x_0)(x-x_1)(x-x_2)\dots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)(x_n-x_2)\dots(x_n-x_{n-1})}. \end{aligned} \quad (109)$$

Интерполяционный полином Лагранжа обычно применяется в теоретических исследованиях при доказательстве теорем, аналитическом решении задач и т. п.

Пример

Для функции $y = \sin \pi x$ найти выражение интерполяционного полинома Лагранжа, выбрав узлы $x_0 = 0$, $x_1 = \frac{1}{6}$, $x_2 = \frac{1}{2}$.

Найти значения полинома Лагранжа для значений x : $\frac{1}{4}$ и $\frac{1}{3}$.

Определить абсолютную и относительную погрешности вычислений.

Решение

Таблица 19

1. Вычислим соответствующие значения функции в узлах:

$$y_0 = 0; \quad y_1 = \sin \frac{\pi}{6} = \frac{1}{2};$$

$$y_2 = \sin \frac{\pi}{2} = 1 \text{ (табл. 19).}$$

Узлы интерполяции

Индекс	0	1	2
x	0	$\frac{1}{6}$	0,5
y	0	0,5	1

Применяя формулу (101), получим

$$P_n(x) = \frac{\left(x - \frac{1}{6}\right)\left(x - \frac{1}{2}\right)}{\left(0 - \frac{1}{6}\right)\left(0 - \frac{1}{2}\right)} \cdot 0 + \frac{(x-0)\left(x - \frac{1}{2}\right)}{\left(\frac{1}{6} - 0\right)\left(\frac{1}{6} - \frac{1}{2}\right)} \cdot \frac{1}{2} + \frac{(x-0)\left(x - \frac{1}{6}\right)}{\left(\frac{1}{2} - 0\right)\left(\frac{1}{2} - \frac{1}{6}\right)} \cdot 1;$$

$$P_n(x) = 3,5x - 3x^2.$$

2. Определим значения полинома Лагранжа для значений x : $\frac{1}{4}$ и $\frac{1}{3}$:

$$P_n\left(\frac{1}{4}\right) = \frac{7}{2} \cdot \frac{1}{4} - 3 \cdot \frac{1}{16} \approx 0,688 \text{ и } P_n\left(\frac{1}{3}\right) = \frac{7}{2} \cdot \frac{1}{3} - 3 \cdot \frac{1}{9} \approx 0,833.$$

3. Определим погрешности вычислений.

Для этого найдем значения функции $y = \sin \pi x$ при заданных значениях x , составив соответствующую таблицу (табл. 20).

Таблица 20

Таблица данных

x	0	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$
y	0	0,5	0,71	0,87	1
$P_n(x)$	0	0,5	0,69	0,83	1

Абсолютная погрешность измерения, определяемая как разность между истинным и измеренным значениями физической величины:

$$\Delta_{1/4} = 0,71 - 0,69 = 0,02 \text{ и } \Delta_{1/3} = 0,87 - 0,83 = 0,04.$$

Относительная погрешность находится как отношение абсолютной погрешности к истинному значению или к результату измерения:

$$\delta_{1/4} = \frac{0,02}{0,71} \cdot 100 \approx 2,82, \text{ или } \delta_{1/4} = \frac{0,02}{0,69} \cdot 100 \approx 2,9;$$

$$\delta_{1/3} = \frac{0,04}{0,87} \cdot 100 \approx 4,6, \text{ или } \delta_{1/3} = \frac{0,04}{0,83} \cdot 100 \approx 4,82.$$

5.1.4. Интерполяционный полином Ньютона

Интерполяционный многочлен можно записать в форме, предложенной Ньютоном, если узлы интерполяции равноотстоящие по величине, т. е.

$$x_{i+1} - x_i = h = \text{const}, \quad (110)$$

где h – шаг интерполяции, а $x_i = x_0 + nh$.

Интерполяционные полиномы Ньютона используются в случае, когда точка интерполирования находится в начале или конце таблицы, соответственно, первая и вторая интерполяционная формула Ньютона.

5.1.4.1. Первая интерполяционная формула Ньютона

Вид интерполирующего полинома в данном случае:

$$P_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \dots + a_n(x - x_0) \dots (x - x_{n-1}). \quad (111)$$

Решение многочлена сводится к определению коэффициентов a_i .

При нахождении коэффициентов пользуются конечными разностями.

Конечные разности первого порядка запишутся в виде:

$$\begin{aligned} \Delta y_0 &= y_1 - y_0; \\ \Delta y_1 &= y_2 - y_1; \\ &\dots\dots\dots \\ \Delta y_{n-1} &= y_n - y_{n-1}, \end{aligned} \quad (112)$$

где y_i – значения функции при соответствующих значениях x_i .

Конечные разности второго порядка:

$$\begin{aligned} \Delta^2 y_0 &= \Delta y_1 - \Delta y_0; \\ \Delta^2 y_1 &= \Delta y_2 - \Delta y_1; \\ &\dots\dots\dots \\ \Delta^2 y_{n-2} &= \Delta y_{n-1} - \Delta y_{n-2}. \end{aligned} \quad (113)$$

Конечные разности высших порядков найдутся аналогично:

$$\begin{aligned} \Delta^k y_0 &= \Delta^{k-1} y_1 - \Delta^{k-1} y_0; \\ \Delta^k y_1 &= \Delta^{k-1} y_2 - \Delta^{k-1} y_1; \\ &\dots\dots\dots \\ \Delta^k y_{n-2} &= \Delta^{k-1} y_{n-1} - \Delta^{k-1} y_{n-2}. \end{aligned}$$

Коэффициенты a_0, a_1, \dots, a_n находятся из условия (103).

Коэффициент a_0 находим при $x = x_0$. Выражение (110) примет вид

$$a_0 = P(x_0) = y_0.$$

Коэффициент a_1 находим при $x = x_1$ при подстановке в выражение (111) значения коэффициента $a_0 = y_0$:

$$P_n(x_1) = y_1 = y_0 + a_1(x_1 - x_0).$$

Используя формулы (111) и (112):

$$a_1 = \frac{y_1 - y_0}{x_1 - x_0} = \frac{\Delta y_0}{h}.$$

Для определения a_2 , полагая $x = x_2$, на основе формул (111), (112) и (113) получим

$$\begin{aligned} P_n(x_2) = y_2 &= y_0 + (x_2 - x_0) + a_2(x_2 - x_0)(x_2 - x_1) = y_0 + 2\Delta y_0 + a_2 2h^2; \\ a_2 &= \frac{y_2 - y_0 - 2\Delta y_0}{2h^2} = \frac{y_2 - y_0 - 2y_1 + 2y_0}{2h^2} = \frac{y_2 - y_0 - 2y_1 + 2y_0}{2h^2} = \\ &= \frac{(y_2 - y_1) - (y_1 - y_0)}{2h^2} = \frac{\Delta y_1 - \Delta y_0}{2h^2} = \frac{\Delta^2 y_0}{2!h^2}. \end{aligned}$$

Общая формула для нахождения всех коэффициентов имеет вид

$$a_i = \frac{\Delta^i y_0}{i!h^i}.$$

В результате (111) примет вид

$$\begin{aligned} P_n(x) = y_0 + \frac{\Delta y_0}{1! \cdot h}(x - x_0) + \frac{\Delta^2 y_0}{2! \cdot h^2}(x - x_0)(x - x_1) + \dots + \\ + \frac{\Delta^n y_0}{n! \cdot h^n}(x - x_0) \dots (x - x_{n-1}). \end{aligned} \quad (114)$$

Пример

По полученным данным (табл. 21) определить значение вязкости жидкости при температуре $T = 12^\circ\text{C}$.

Таблица 21

Зависимость вязкости воды от температуры

$T, ^\circ\text{C}$	0	25	50	75	100
$\rho, \text{кг/м}^3$	1000	997	988	975	960

Решение

Для упрощения расчетов составим таблицу конечных разностей (табл. 22).

В связи с тем, что требуется найти значение в начале таблицы, воспользуемся формулой для первого интерполяционного полинома Ньютона (106):

$$\begin{aligned}
P_4(T) &= \rho_0 + \frac{\Delta\rho_0}{h}(T-T_0) + \frac{\Delta^2\rho_0}{2!h^2}(T-T_0)(T-T_1) + \\
&+ \frac{\Delta^3\rho_0}{3!h^3}(T-T_0)(T-T_1)(T-T_2) + \frac{\Delta^4\rho_0}{4!h^4}(T-T_0)(T-T_1)(T-T_2)(T-T_3) = \\
&= 1000 - \frac{3}{25}(T-0) - \frac{6}{25^2 \cdot 2}(T-0)(T-25) + \frac{2}{25^3 \cdot 6}(T-0)(T-25)(T-50) = \\
&= 0,0000213 \cdot T^3 - 0,0064 \cdot T^2 + 0,0267 \cdot T + 1000.
\end{aligned}$$

Таблица 22

Таблица конечных разностей

Индекс	T	ρ	$\Delta\rho$	$\Delta^2\rho$	$\Delta^3\rho$	$\Delta^4\rho$
0	0	1000	-3	-6	2	0
1	25	997	-9	-4	2	
2	50	988	-13	-2		
3	75	975	-15			
4	100	960				

Подставив в формулу полученного полинома значение $T = 12^\circ\text{C}$, найдем значение плотности $\rho = 999,35 \text{ кг/м}^3$.

5.1.4.2. Вторая интерполяционная формула Ньютона

Для нахождения значений функции в конце интервала интерполирования интерполяционный полином запишется в виде

$$\begin{aligned}
P_n(x) &= a_0 + a_1(x-x_n) + a_2(x-x_n)(x-x_{n-1}) + \dots + \\
&+ a_n(x-x_n)(x-x_{n-1})\dots(x-x_1).
\end{aligned} \tag{115}$$

Алгоритм определения коэффициентов a_0, a_1, \dots, a_n аналогичен рассмотренному.

Из условия (103), подставляя в (115) $x = x_n$, найдем коэффициент a_0 :

$$P_n(x_n) = y_n = a_0.$$

Для $x = x_{n-1}$:

$$P_n(x_{n-1}) = y_{n-1} = y_n + a_1(x_{n-1} - x_n), \quad a_1 = \frac{y_n - y_{n-1}}{h} = \frac{\Delta y_{n-1}}{h}.$$

Для $x = x_{n-2}$:

$$\begin{aligned}
P_n(x_{n-2}) &= y_{n-2} = y_n + \frac{\Delta y_{n-1}}{h}(x_{n-2} - x_n) + a_2(x_{n-2} - x_n)(x_{n-2} - x_{n-1}) = \\
&= y_n + \frac{\Delta y_{n-1}}{h}(-2h) + a_2 \cdot 2h^2 = y_n - 2\Delta y_{n-1} + a_2 2h^2;
\end{aligned}$$

$$a_2 = \frac{\Delta^2 y_{n-2}}{2!h^2}.$$

Формула для нахождения всех коэффициентов запишется как:

$$a_i = \frac{\Delta^i y_{n-1}}{i!h^i}.$$

Подставив выражения для определения коэффициентов a_i в формулу (115), получим вторую интерполяционную формулу Ньютона:

$$\begin{aligned} P_n(x) = & y_n + \frac{\Delta y_{n-1}}{h}(x - x_n) + \frac{\Delta^2 y_{n-2}}{2!h^2}(x - x_n)(x - x_{n-1}) + \\ & + \frac{\Delta^3 y_{n-3}}{3!h^3}(x - x_n)(x - x_{n-1})(x - x_{n-2}) + \\ & + \dots + \frac{\Delta^n y_0}{n!h^n}(x - x_n)(x - x_{n-1}) \dots (x - x_1). \end{aligned} \quad (116)$$

Пример

Дана таблица значений (см. табл. 21) $\rho = f(T)$.

1. Построить интерполяционный многочлен Ньютона.
2. Определить значение полинома для температуры $T = 90^\circ\text{C}$.

Решение

Для построения полинома воспользуемся формулой (115) и табл. 22:

$$\begin{aligned} P_4(x) = & \rho_4 + \frac{\Delta \rho_3}{h}(T - T_4) + \frac{\Delta^2 \rho_2}{2!h^2}(T - T_4)(T - T_3) + \\ & + \frac{\Delta^3 \rho_1}{3!h^3}(T - T_4)(T - T_3)(T - T_2) + \frac{\Delta^4 \rho_0}{4!h^4}(T - T_4)(T - T_3)(T - T_2)(T - T_1) = \\ & = 960 - \frac{15(T - 100)}{25} - \frac{2(T - 100)(T - 75)}{2! \cdot 25^2} + \\ & + \frac{2(T - 100)(T - 75)(T - 50)}{3! \cdot 25^3} = \\ & = 0,0000213 \cdot T^3 - 0,0064 \cdot T^2 - 0,2933 \cdot T + 1028. \end{aligned}$$

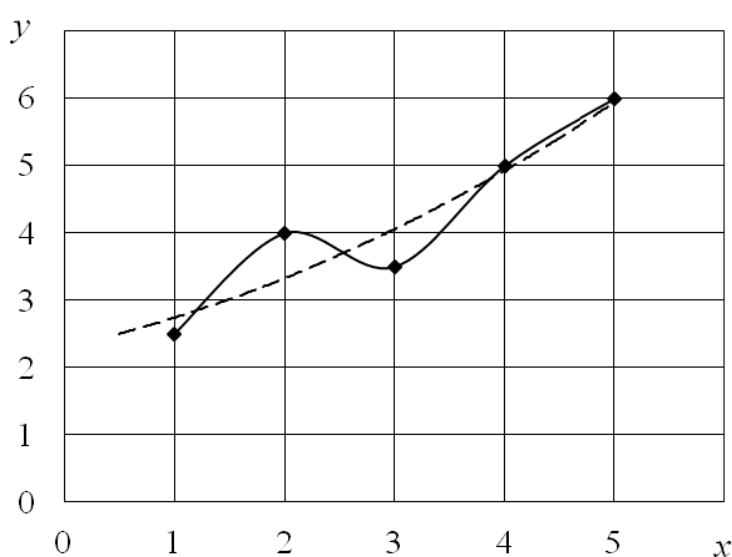
Подставив в формулу полученного полинома значение $T = 90^\circ\text{C}$, найдем значение плотности $\rho = 965,29 \text{ кг/м}^3$.

5.2. Аппроксимация функций

Аппроксимация, как и интерполяция, – это метод замены одних математических объектов другими, близкими к исходным.

При интерполировании искомая функция строго проходит через узловые точки таблицы и количество коэффициентов в интерполирующей функции равно количеству табличных значений.

Аппроксимация – метод приближения, при котором для нахождения дополнительных значений, отличных от табличных данных, приближенная функция проходит не через узлы интерполяции, а между ними (рис. 24).



— — интерполирующая функция --- — аппроксимирующая функция

Рис. 24. Интерполирующая и аппроксимирующая функции

Интерполяцией данные описываются более точно, чем при аппроксимации, но в ряде случаев обосновано применение аппроксимации:

- при значительном количестве табличных данных – интерполирующая функция становится громоздкой;
- интерполирующей функцией невозможно описать данные при повторении эксперимента в одних тех же начальных условиях – требуется статистическая обработка;
- для сглаживания погрешностей эксперимента. Данные x_i и y_i обычно содержат ошибки, поэтому интерполяционная формула повторяет эти ошибки. Из рис. 25 видно, что значение y постоянно и равномерно увеличивается при росте x , а разброс данных относительно аппроксимирующей функции можно объяснить погрешностью эксперимента.

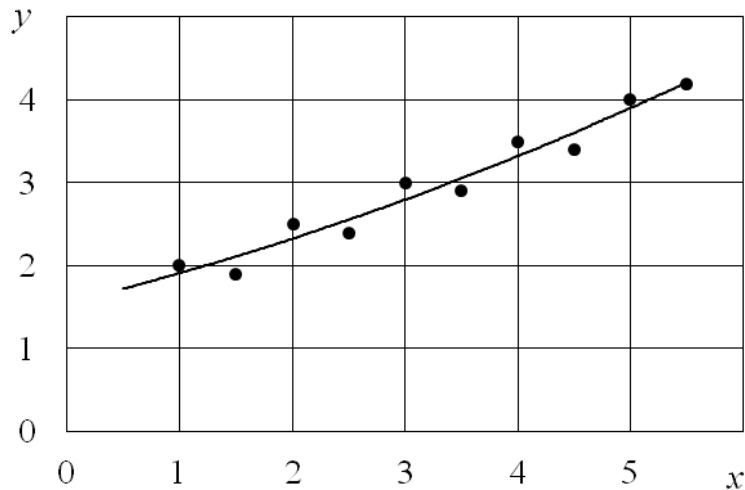


Рис. 25. Пример построения аппроксимирующей функции

Геометрически задача построения функции $f(x)$ по эмпирической формуле состоит в проведении усредненной кривой – кривой, проходящей через середину области значений (рис. 24, 25).

При построении аппроксимирующей зависимости определяют:

- аналитический характер эмпирической формулы. Предпочтение отдается простым формулам, обладающим хорошей точностью;
- коэффициенты эмпирической зависимости.

Для нахождения аппроксимирующей функции пользуются теми же методами приближения, что при регрессионном анализе (см. п. 4.2.2), задачей которого и является подбор математических зависимостей, наилучшим образом описывающих изменение экспериментальных данных. В связи с этим различают аппроксимацию:

- линейную – $P(x) = b_0 + b_1 \cdot x$ (см. п. 4.2.2.2);
- полиномиальную – $P_N(x) = b_0 + b_1x + b_2x^2 + \dots + b_Nx^N$ (см. п. 4.2.2.1);
- в виде определенных функций, в частности:
 - показательной – $y = b_0e^{b_1x}$ (см. п. 4.2.2.4);
 - степенной – $y = b_0x^{b_1}$ (см. п. 4.2.2.5).

6. СТАТИСТИЧЕСКИЕ МЕТОДЫ КОНТРОЛЯ КАЧЕСТВА

6.1. Понятие и методы статистического контроля

Прежде чем говорить о методах управления качеством, тем более статистических, определимся с понятием «качество».

Качеством называют совокупность характерных свойств, формы, внешнего вида и условий применения, которыми должны быть наделены товары для соответствия своему назначению или совокупность характерных особенностей, формы, внешнего вида и условий эксплуатации, определяющих пригодность товаров. Современная трактовка понятия «качество» – совокупность свойств товаров и услуг, определяющих их способность удовлетворять реальные и потенциальные потребности потребителей.

Одним из главных средств, повышающих соответствие изделий техническим требованиям и одновременно предусматривающих сдерживание роста затрат по управлению предприятием наряду с модернизацией оборудования, усовершенствованием технологии и т. п., является использование статистических методов.

Статистические методы, основанные на использовании математической статистики, являются эффективным инструментом сбора, анализа и интерпретации информации о качестве. Применение этих методов, не требуя больших затрат, позволяет с заданной степенью точности и достоверности судить о состоянии исследуемых явлений (объектов, процессов) в системе менеджмента качества, прогнозировать и решать проблемы на всех этапах жизненного цикла продукции и на основе этого вырабатывать оптимальные управленческие решения. В соответствии с положениями стандартов ИСО серии 9000, статистические методы рассматриваются как единый набор высокоэффективных средств обеспечения и улучшения качества на основе объективно полученных и интерпретированных фактов. Стандарты ориентируют на разработку сквозного механизма применения статистических методов, начиная с исследования требований рынка к качеству продукции и заканчивая ее утилизацией после использования.

На сегодняшний день известны следующие категории статистических методов, которые классифицируют по степени сложности:

1. Элементарные статистические методы, включающие так называемые «7 инструментов». Из них:

1.1. Карта Парето.

1.2. Причинно-следственный анализ (диаграмма Исикава).

1.3. Группировка данных по общим признакам (диаграмма сродства).

- 1.4. Контрольный лист.
- 1.5. Гистограмма.
- 1.6. Диаграмма разброса.
- 1.7. Контрольная карта.
- 2. Промежуточные статистические методы:
 - 2.1. Теория выборочных исследований.
 - 2.2. Статистический выборочный контроль.
 - 2.3. Методы проведения статистических оценок и определения критериев.
 - 2.4. Методы применения сенсорных проверок (экспертные оценки).
 - 2.5. Методы планирования и расчета экспериментов.
 - 2.6. Корреляционный и регрессионный анализы.
- 3. Передовые статистические методы:
 - 3.1. Передовые методы планирования и расчета экспериментов.
 - 3.2. Многофакторный (дисперсионный) анализ.
 - 3.3. Методы исследования операций.

В настоящее время эти методы, получив дальнейшее развитие, стандартизированы и рекомендуются для использования в работе по повышению качества.

Первое применение научных методов статистического контроля было зафиксировано в 1924 году, когда В. Шухарт использовал для определения доли брака продукции контрольные карты. Одновременно с Шухартом, в той же фирме, в середине 20-х годов инженером Г.Ф. Доджем была предложена теория приемочного контроля, получившая вскоре мировую известность. Большой вклад в систему обеспечения качества контроля в середине 20-го века внесли американские ученые Д. Нойман, Э. Пирсон, Е. Фишер. Среди их разработок наибольшую известность получила теория проверки статистических гипотез. Большое влияние на становление статистических методов контроля, как философии качества, оказал Эдвард Деминг (США). В начале 50-х годов Деминг проводил широкомасштабное обучение японских специалистов новым методам обеспечения качества, особое внимание при этом обращая на статистические методы управления качеством. Американское научное влияние на совершенствование систем обеспечения качества привело к созданию японской научной школы в области качества, среди представителей которых следует, прежде всего, отметить К. Исикаву и Г. Тагути, внесших большой вклад в развитие статистических методов в управлении качеством. Внесли свой научный вклад в развитие статистических методов и советские ученые: В.И. Романовский, Е.Е. Слуцкий, Н.В. Смирнов, Ю.В. Линник и др. Так, например,

Смирнов заложил основы теории непараметрических рядов, а Слуцкий опубликовал несколько важных работ по статистике связанных стационарных рядов. Особенно интенсивно в СССР разрабатывались статистические методы исследования и контроля качества в массовом производстве, методы планирования эксперимента (Ю.П. Адлер и др.). В последние годы можно отметить работы российского ученого в области качества В.А. Лapidуса. Им опубликован ряд трудов по теории и практике управления качеством с учетом вариаций и неопределенности, в которых изложен «принцип распределения приоритетов», позволяющий оптимально выстроить отношения поставщика и потребителя с позиции обеспечения качества.

По статистическим методам имеется обширная литература и пакеты прикладных компьютерных программ, по разработке которых отечественные научные школы по теории вероятностей занимают ведущее место в мире.

Статистические методы обеспечения качества широко применяются в США, Японии, Великобритании, Германии, Франции, Италии, Голландии, Дании и других странах.

Несмотря на всемирную известность представителей отечественной школы математической статистики (А.И. Колмогорова, Н.В. Смирнова, А.Я. Хинчина, Я.Б. Шора и др.), Россия пока отстает от промышленно развитых стран в области массового применения статистических методов. Интенсивное распространение этих методов в нашей стране приходится на 40-50-е годы. Но впоследствии статистические методы, несмотря на их научную обоснованность и прогрессивный характер, практически не используются, за исключением отдельных предприятий. Основными причинами низкого применения статистических методов являются:

- низкая технологическая дисциплина;
- пренебрежение нормами конструкторской и технологической документации;
- несогласованность конструкторских норм с технологическими и метрологическими возможностями;
- нехватка квалифицированных специалистов в области статистических методов;
- усложненность многих методических пособий по статистическим методам, их перегруженностью математикой;
- отсутствие экономической заинтересованности предприятий во внедрении этих методов.

В соответствии с положениями стандартов ISO серии 9000 статистические методы рассматриваются как одно из высокоэффективных средств обеспечения качества. Стандарты ориентируют на разработку механизма применения статистических методов на всех этапах жизненного цикла продукции, начиная с исследования требований рынка к качеству продукции и заканчивая ее утилизацией после использования. Это означает, что работа по внедрению статистических методов должна быть направлена на создание гарантий непрерывности процесса обеспечения качества в соответствии с требованиями заказчика. Меняется и характер данной работы: от локального, случайного, – к общему, системному. Статистические методы становятся основой для информационной технологии обеспечения качества.

Среди специалистов по статистике бытует мнение, что статистические методы – если не единственное, то, по крайней мере, самое главное средство решения проблемы обеспечения качества. Такой подход крайне опасен, поскольку решение столь сложной проблемы не может быть результатом применения какого-то одного, пусть даже высокоэффективного средства. Здесь следует напомнить предупреждение Э. Деминга о том, что повышения качества, производительности труда и конкурентоспособности продукции нельзя добиться исключительно за счет массированного применения контрольных карт и других статистических методов. Статистические методы являются лишь одним из многочисленных средств обеспечения качества, и успех в этой области определяется правильным сочетанием всех имеющихся средств в зависимости от конкретных условий. Вместе с тем применение статистических методов при внедрении стандартов ИСО серии 9000 приобретает особую значимость, так как именно с их помощью возможно объективное подтверждение стабильности процессов и качества продукции предприятия.

6.2. Система нормативных документов, регламентирующих применение статистических методов

Необходимость упорядочения проводимых разными странами работ в области статистических методов привела к созданию в составе ИСО специального ТК 69: «Применение статистических методов». Разработкой рекомендаций по применению статистических методов занимается также один из комитетов ЕОК.

Комитет ISO/TC 69 основан в 1948 году. Ответственный член ИСО – Франция. На сегодняшний день комитетом опубликовано 95 стандартов. Комитет состоит из шести подкомитетов, ответственных:

- за терминологию и условные обозначения;
- применение статистических методов в менеджменте процессов;
- выборочный приемочный контроль;
- методы и результаты измерений;
- применение статистических и сопутствующих методов для внедрения Six Sigma;
- применение статистических и сопутствующих методов для создания новых технологий и продукции.

Каждый подкомитет несет ответственность за разработку стандартов, входящих в их область действия.

Для международных стандартов в области статистического контроля установлены стратегические требования, закрепленные в бизнес-плане комитета ИСО ISO/TC 69:

- описывать передовой опыт, достигаемый за счет достижения международного соглашения;
- указывать проверяемые требования;
- быть легко понятными для целевой аудитории;
- быть читаемым, избегая употребления длинных предложений, необычных слов, неопределенностей и т. д.;
- соответствовать деятельности или процессу;
- требовать низкий уровень математической сложности от пользователя;
- своевременно выпускаться;
- соответствовать потребностям рынка;
- соответствовать стратегии ТК 69;
- иметь необходимую экспертизу в ТК 69.

В России разработкой стандартов в области статистического контроля занимается Федеральный орган по техническому регулированию и метрологии. Всю нормативную документацию, регламентирующую применение статистических методов, можно условно разделить на следующие группы (рис. 26).



Рис. 26. Классификация нормативных документов, регламентирующих применение статистических методов

6.3. Семь простых методов контроля качества

Назначение методов

Применяются как непосредственно в производстве, так и на различных стадиях жизненного цикла продукции.

Цель методов

Выявление проблем, подлежащих первоочередному решению, на основе контроля действующего процесса, сбора, обработки и анализа полученных фактов (статистического материала) для последующего улучшения качества процесса.

Суть методов

Контроль качества (сравнение запланированного показателя качества с действительным его значением) – это одна из основных функций в процессе управления качеством, а сбор, обработка и анализ фактов – важнейший этап этого процесса.

Научной основой современного технического контроля является математико-статистические методы.

Из множества статистических методов для широкого применения выбраны только семь, которые понятны и могут легко применяться специалистами различного профиля. Они позволяют вовремя выявить и отобразить проблемы, установить основные факторы, с которых нужно начинать действовать, и распределить усилия с целью эффективного разрешения этих проблем.

План действий

Внедрение семи методов должно начинаться с обучения этим методам всех участников процесса.

Последовательность применения методов может быть различной в зависимости от поставленной цели.

Эти методы можно рассматривать и как отдельные инструменты, и как систему методов. Каждый метод может находить свое самостоятельное применение в зависимости от того, к какому классу относится задача.

Особенности методов

Семь основных инструментов контроля качества – набор инструментов, позволяющих облегчить задачу контроля протекающих процессов и предоставить различного рода факты для анализа, корректировки и улучшения качества процессов.

Для анализа данных используются семь так называемых статистических методов, или инструментов контроля качества:

1. Контрольный листок.
2. Гистограмма.

3. Диаграмма Парето.
4. Метод стратификации.
5. Диаграмма разброса.
6. Диаграмма Исикавы.
7. Контрольная карта.

Дополнительная информация:

1. Семь простых статистических методов – инструменты познания, а не управления.
2. Способность рассматривать события с точки зрения статистики важнее, чем знание самих методов.
3. На передовых зарубежных фирмах абсолютно все работники обязаны владеть семью простыми статистическими методами.
4. Данные необходимо собирать так, чтобы облегчить их последующую обработку. Нужно понимать, для каких целей осуществляется сбор и обработка данных.

Обычно цели сбора данных в процессе контроля качества состоят в следующем:

- контроль и регулирование процесса;
- анализ отклонений от установленных требований;
- контроль выхода процесса.

Достоинства метода

Наглядность, простота освоения и применения.

Недостатки метода

Низкая эффективность при проведении анализа сложных процессов.

Ожидаемый результат

Решение до 95 % всех проблем, возникающих на производстве.

Контрольный листок

Контрольный листок (таблица накопленных частот) – инструмент для сбора данных и их автоматического упорядочения для облегчения дальнейшего использования собранной информации. Контрольный листок составляется для построения гистограммы.

Таблица 23

Пример контрольного листа

№ интервала	Измеренные значения	Частота	Накопленная частота	Накопленная относительная частота

На основании контрольного листка строится гистограмма или при большом количестве измерений – кривая распределения плотности вероятностей.

Гистограмма

Гистограмма – инструмент, позволяющий зрительно оценить распределение статистических данных, сгруппированных по частоте попадания данных в определенный (заранее заданный) интервал.

Один из наиболее распространенных методов, помогающих интерпретировать данные по исследуемой проблеме.

Гистограмма – один из вариантов столбиковой диаграммы, позволяющий зрительно оценить распределение статистических данных, сгруппированных по частоте попадания в определенный (заранее заданный) интервал.

Применяется везде, где требуется проведение анализа точности и стабильности процесса, наблюдение за качеством продукции, отслеживание существенных показателей производства.

Целью метода является контроль действующего процесса и выявление проблем, подлежащих первоочередному решению.

Порядок построения гистограммы:

- собрать данные, выявить максимальное и минимальное значения и определить диапазон (размах) гистограммы;
- полученный диапазон разделить на интервалы, предварительно определив их число (обычно 5–20 в зависимости от числа показателей), и определить ширину интервала;
- все данные распределить по интервалам в порядке возрастания: левая граница первого интервала должна быть меньше наименьшего из имеющихся значений;
- подсчитать частоту каждого интервала;
- вычислить относительную частоту попадания данных в каждый из интервалов;
- по полученным данным построить гистограмму – столбчатую диаграмму, высота столбиков которой соответствует частоте или относительной частоте попадания данных в каждый из интервалов;
- наносится горизонтальная ось, выбирается масштаб и откладываются соответствующие интервалы;
- затем строится вертикальная ось, на которой также выбирается масштаб в соответствии с максимальным значением частот.

Достоинства метода:

- наглядность, простота освоения и применения;
- управление с помощью фактов, а не мнений;
- возможность лучше понять вариабельность, присущую процессу, глубже взглянуть на проблему и облегчить нахождение путей ее решения.

Недостаток метода – интерпретация гистограммы, построенная по малым выборкам, не позволяет сделать правильные выводы.

Характерные типы гистограмм показаны на рис. 27.

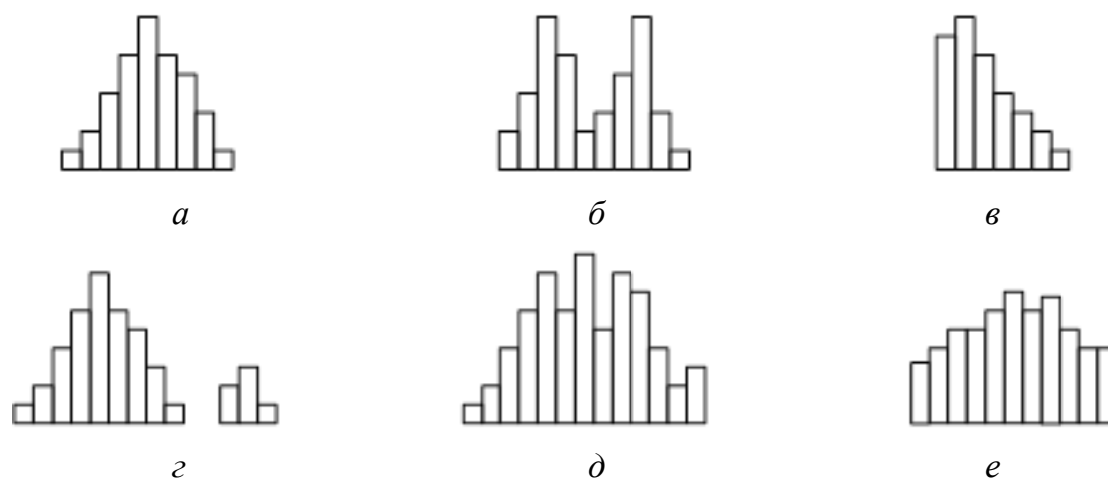


Рис. 27. Характерные типы гистограмм

На рис. 27 показаны следующие типы гистограмм:

а – обычный тип гистограммы с двусторонней симметрией, что указывает на стабильность процесса;

б – двугорбая гистограмма, т. к. в распределении имеется два пика. Такая гистограмма получается при объединении двух распределений, например, в случае двух видов сырья, изменения настройки процесса или объединения в одну партию изделий, обработанных на двух разных станках. Требуется расслоение продукции;

в – гистограмма с обрывом, получается, когда невозможно получить значение ниже (или выше) некоторой величины. Подобное распределение имеет место также, когда из партии исключены все изделия с показателем ниже (и/или выше) нормы, т. е. изначально это была партия с большим количеством дефектных изделий. Такое же распределение получается, когда измерительные приборы были неисправны;

г – гистограмма с островком, получается при ошибках в измерениях или когда некоторое количество дефектных изделий перемешано с доброкачественными;

д – гистограмма с прогалами («гребенка»), получается, когда ширина интервала не кратна единице измерения или при ошибках оператора;

е – гистограмма в форме плато, получается, когда объединяются несколько распределений при небольшой разнице средних значений. В этом случае требуется расслоение.

При исследовании гистограммы можно выяснить, в удовлетворительном ли состоянии находятся партия изделий и технологический процесс. Рассматривают следующие вопросы:

- какова ширина распределения по отношению к ширине допуска;
- каков центр распределения по отношению к центру поля допуска;
- какова форма распределения.

Диаграмма Парето

Диаграмма Парето – инструмент, позволяющий объективно представить и выявить основные факторы, влияющие на исследуемую проблему, и распределить усилия для ее эффективного разрешения.

Схема, построенная на основе группирования по дискретным признакам, ранжированная в порядке убывания (например, по частоте появления) и показывающая кумулятивную (накопленную) частоту, называется диаграммой Парето.

Парето – итальянский экономист и социолог, использовавший свою диаграмму для анализа богатств Италии.

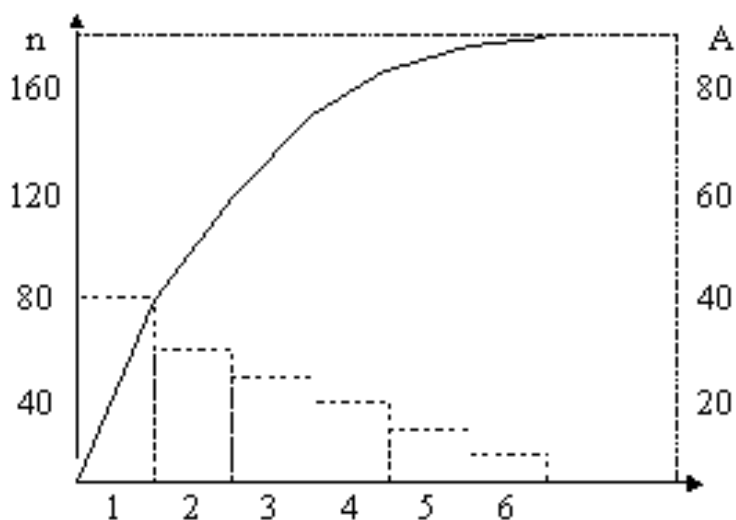


Рис. 28. Пример диаграммы Парето:

1 – ошибки в процессе производства; 2 – некачественное сырье; 3 – некачественные орудия труда; 4 – некачественные шаблоны; 5 – некачественные чертежи; 6 – прочее; A – относительная кумулятивная (накопленная) частота, %; n – число бракованных единиц продукции

Приведенная диаграмма построена на основе группирования бракованной продукции по видам брака и расположения в порядке убывания числа единиц бракованной продукции каждого вида. Диаграмму Парето можно использовать очень широко. С ее помощью можно оценить эффективность принятых мер по улучшению качества продукции, построив ее до и после внесения изменений.

Метод стратификации

Метод стратификации (расслаивания данных) – инструмент, позволяющий произвести разделение данных на подгруппы по определенному признаку.

При разделении данных на группы в соответствии с их особенностями группы именуют слоями (стратами), а сам процесс разделения – расслаиванием (стратификацией). Желательно, чтобы различия внутри слоя были как можно меньше, а между слоями – как можно больше.

В результатах измерений всегда есть больший или меньший разброс параметров. Если осуществлять стратификацию по факторам, порождающим этот разброс, легко выявить главную причину его появления, уменьшить его и добиться повышения качества продукции.

Применение различных способов расслаивания зависит от конкретных задач. В производстве часто используется способ, называемый 6М, учитывающий факторы, зависящие: от человека (*man*); машины (*machine*); материала (*material*); метода (*method*); контроля (*measurement*); среды (*mother nature*).

Другими словами, расслаивание можно осуществить:

- по исполнителям (по полу, стажу работы, квалификации и т. д.);
- машинам и оборудованию (по новому или старому, марке, типу и т. д.);
- материалу (по месту производства, партии, виду, качеству сырья и т. д.);
- способу производства (по температуре, технологическому приему и т. д.).

В торговле может быть расслаивание по районам, фирмам, продавцам, видам товара, сезонам.

Метод расслаивания в чистом виде применяется при расчете стоимости изделия, когда требуется оценка прямых и косвенных расходов отдельно по изделиям и партиям, при оценке прибыли от продажи изделий отдельно по клиентам и по изделиям и т. д.

Расслаивание также используется в случае применения других статистических методов: при построении причинно-следственных диаграмм, диаграмм Парето, гистограмм и контрольных карт.

Диаграмма разброса

Диаграмма разброса (рассеивания) – инструмент, позволяющий определить вид и тесноту связи между парами соответствующих переменных. Анализ корреляции через определение медианы. В отдельных случаях применяется вероятностная номограмма для биномиального распределения.

Применяется для выявления зависимости (корреляции) одних показателей от других или для определения степени корреляции между n парами данных для переменных x и y :

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n).$$

Эти данные наносятся на график (диаграмму разброса), и для них вычисляется коэффициент корреляции по формуле

$$r = \frac{\delta_{xy}}{\delta_x \cdot \delta_y},$$

$$r = \frac{\sum_{i=1}^n x_i \cdot y_i / n - \bar{x} \cdot \bar{y}}{\sqrt{\sum_{i=1}^n x_i^2 / n - \bar{x}^2} \sqrt{\sum_{i=1}^n y_i^2 / n - \bar{y}^2}},$$

$$\delta_x = \sqrt{\sum_{i=1}^n x_i^2 / n - \bar{x}^2}, \quad \delta_y = \sqrt{\sum_{i=1}^n y_i^2 / n - \bar{y}^2},$$

где δ_{xy} – ковариация; δ_x , δ_y – стандартные отклонения случайных переменных x и y ; n – размер выборки (количество пар данных – x_i и y_i); \bar{x} и \bar{y} – среднеарифметические значения x_i и y_i соответственно.

Рассмотрим различные варианты диаграмм разброса (или полей корреляции) на рис. 29.

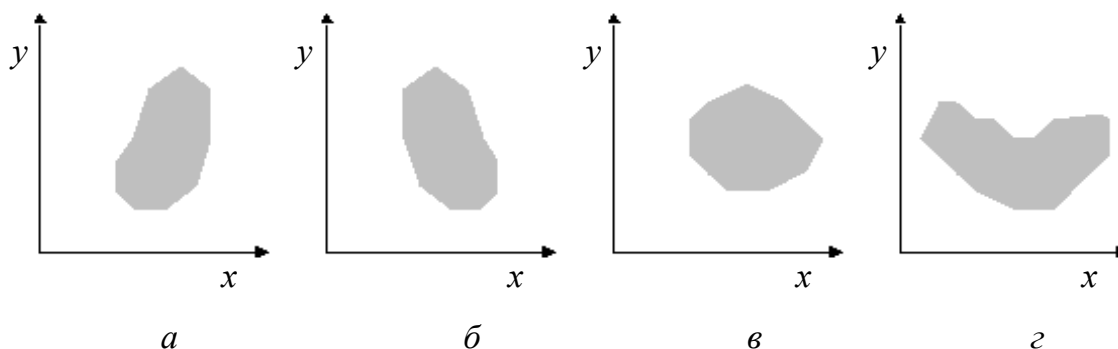


Рис. 29. Варианты диаграмм разброса

В случае:

a – можно говорить о положительной корреляции (с ростом x увеличивается y);

$б$ – проявляется отрицательная корреляция (с ростом x уменьшается y);

$в$ – при росте x , y может как расти, так и уменьшаться, говорят об отсутствии корреляции. Но это не означает, что между ними нет зависимости, между ними нет линейной зависимости. Очевидная нели-

нейная (экспоненциальная) зависимость представлена и на диаграмме разброса z .

Коэффициент корреляции всегда принимает значения в интервале $-1 \leq r \leq 1$, т. е. при $r > 0$ – положительная корреляция, при $r = 0$ – нет корреляции, при $r < 0$ – отрицательная корреляция.

Для тех же n пар данных $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ можно установить зависимость между x и y . Формула, выражающая эту зависимость, называется уравнением регрессии (или линией регрессии), и ее представляют в общем виде функцией

$$y = a + bx.$$

Для определения линии регрессии (рис. 30) необходимо статистически оценить коэффициент регрессии b и постоянную a . Для этого должны быть выполнены следующие условия:

- 1) линия регрессии должна проходить через точки (x, y) средних значений x и y ;
- 2) сумма квадратов отклонений от линии регрессии значений y по всем точкам должна быть наименьшей;
- 3) для расчета коэффициентов a и b используются формулы

$$a = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}, \quad b = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}.$$

Таким образом, уравнением регрессии можно аппроксимировать реальные данные.

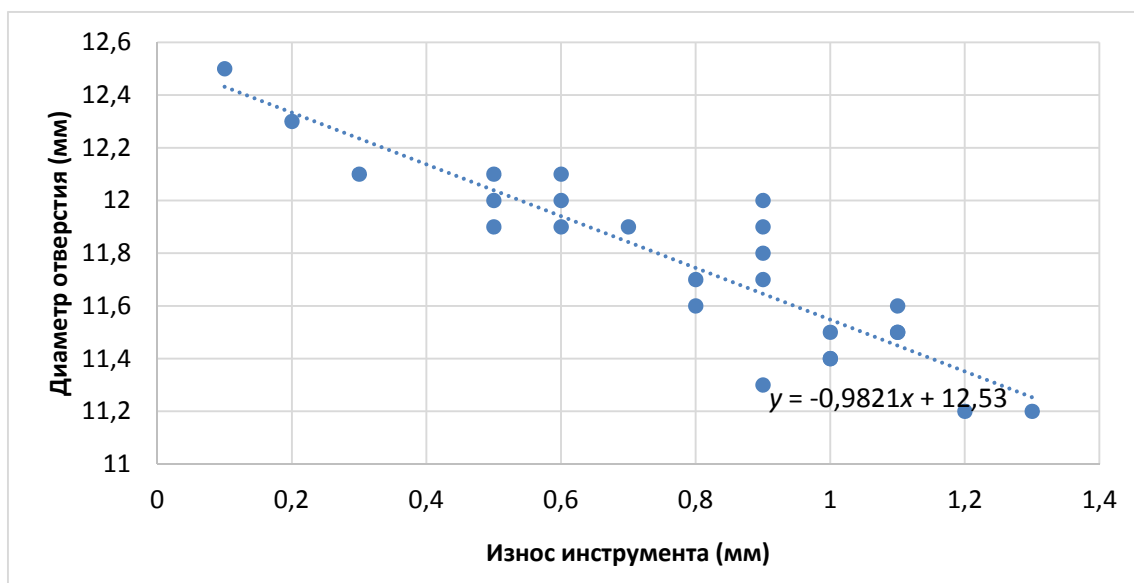


Рис. 30. Пример линии регрессии

Диаграмма Исикавы

Диаграмма Исикавы (причинно-следственная диаграмма) – инструмент, который позволяет выявить наиболее существенные факторы (причины), влияющие на конечный результат (следствие).

Причинно-следственная диаграмма используется, когда требуется исследовать и изобразить возможные причины определенной проблемы. Ее применение позволяет выявить и сгруппировать условия и факторы, влияющие на данную проблему.

Рассмотрим форму причинно-следственной диаграммы на рис. 31 (она называется еще «рыбий скелет», или диаграмма Исикавы).

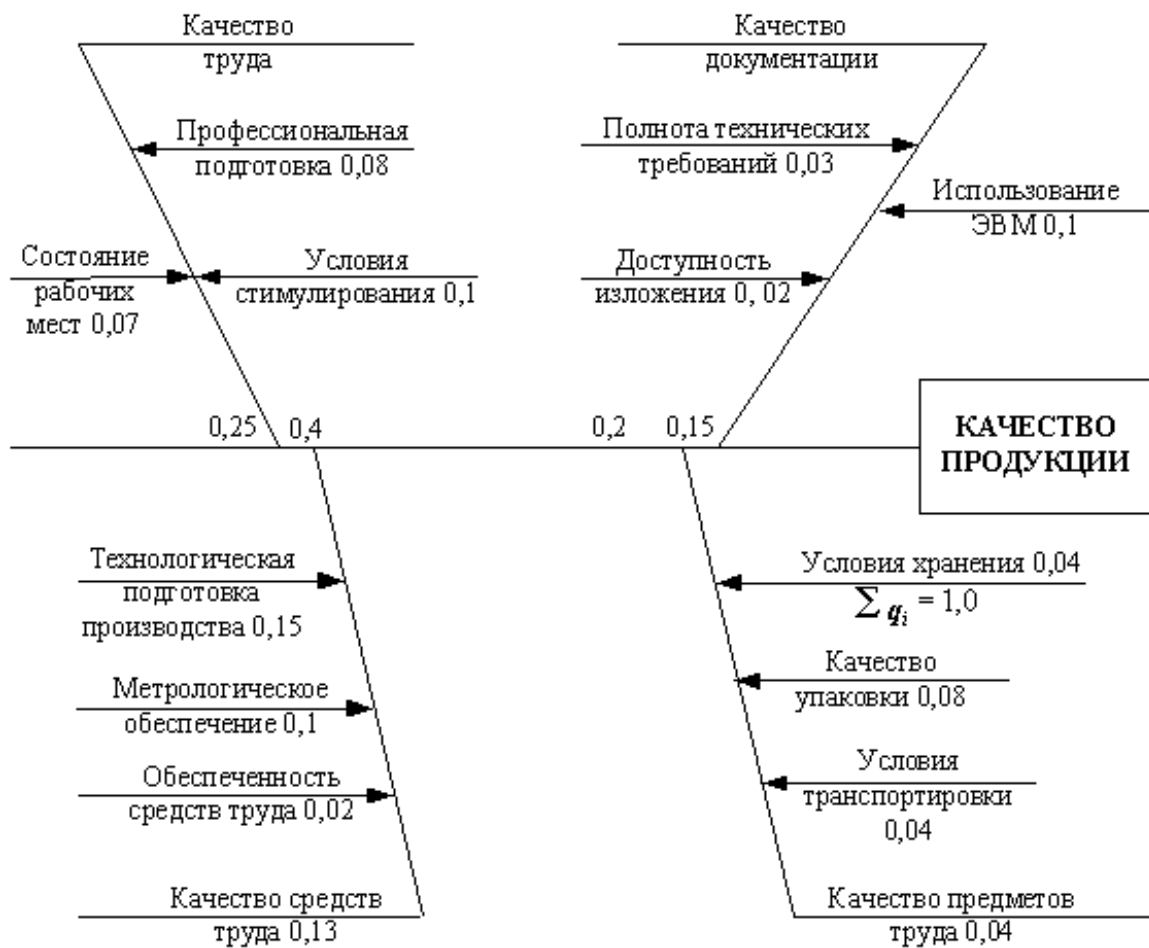


Рис. 31. Пример причинно-следственной диаграммы факторов, влияющих на качество продукции

Порядок составления диаграммы:

1. Выбирается проблема для решения – «хребет».
2. Выявляются наиболее существенные факторы и условия, влияющие на проблему, – причины первого порядка.

3. Выявляется совокупность причин, влияющих на существенные факторы и условия (причины 2-го, 3-го и последующих порядков).

4. Анализируется диаграмма: факторы и условия расставляются по значимости, устанавливаются те причины, которые в данный момент поддаются корректировке.

5. Составляется план дальнейших действий.

Контрольная карта

Контрольная карта – инструмент, позволяющий отслеживать ход протекания процесса и воздействовать на него (с помощью соответствующей обратной связи), предупреждая его отклонения от предъявленных к процессу требований.

Одним из способов достижения удовлетворительного качества и поддержания его на этом уровне является применение контрольных карт. Для управления качеством технологического процесса необходимо иметь возможность контролировать те моменты, когда выпускаемая продукция отклоняется от заданных техническими условиями допусков. Рассмотрим простой пример. Проследим за работой токарного станка в течение определенного времени и будем измерять диаметр детали, изготавливаемой на нем (за смену, час). По полученным результатам построим график и получим простейшую контрольную карту.

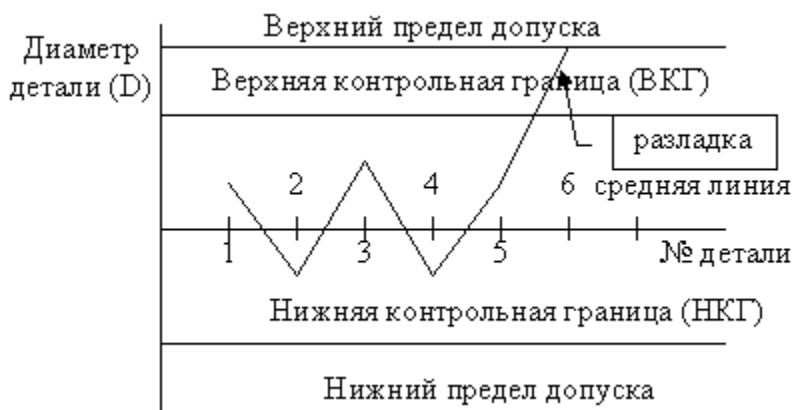


Рис. 32. Пример контрольной карты

В точке 6 произошла разладка технологического процесса, необходимо его регулирование. Положение ВКГ и НКГ определяется аналитически либо по специальным таблицам и зависит от объема выборки. При достаточно большом объеме выборки пределы ВКГ и НКГ определяют по формулам

$$\text{ВКГ} = +3\delta, \quad \text{НКГ} = -3\delta.$$

ВКГ и НКГ служат для предупреждения разладки процесса, когда изделия еще соответствуют техническим требованиям.

Контрольные карты применяются, когда требуется установить характер неисправностей и дать оценку стабильности процесса; когда необходимо установить, нуждается ли процесс в регулировании или его необходимо оставить таким, каков он есть.

Контрольной картой можно также подтвердить улучшение процесса.

Контрольная карта является средством распознавания отклонений из-за неслучайных или особых причин от вероятных изменений, присущих процессу. Вероятные изменения редко повторяются в прогнозируемых пределах. Отклонения из-за неслучайных или особых причин сигнализируют о том, что некоторые факторы, влияющие на процесс, необходимо идентифицировать, расследовать и поставить под контроль.

Контрольные карты основываются на математической статистике. Они используют рабочие данные для установления пределов, в рамках которых будут ожидаться предстоящие исследования, если процесс останется неэффективным из-за неслучайных или особых причин.

Наибольшее распространение получили контрольные карты среднего значения \bar{X} и контрольные карты размаха R , которые используются совместно или раздельно. Контролироваться должны естественные колебания между пределами контроля. Нужно убедиться, что выбран правильный тип контрольной карты для определенного типа данных. Данные должны быть взяты точно в той последовательности, в какой собраны, иначе они теряют смысл. Не следует вносить изменения в процесс в период сбора данных. Данные должны отражать, как процесс идет естественным образом.

Контрольная карта может указать на наличие потенциальных проблем до того, как начнется выпуск дефектной продукции.

Принято говорить, что процесс вышел из-под контроля, если одна или более точек вышли за пределы контроля.

Существуют два основных типа контрольных карт: для качественных (годен – негоден) и для количественных признаков. Для качественных признаков возможны четыре вида контрольных карт: число дефектов на единицу продукции; число дефектов в выборке; доля дефектных изделий в выборке; число дефектных изделий в выборке. При этом в первом и третьем случаях объем выборки будет переменным, а во втором и четвертом – постоянным.

Таким образом, целями применения контрольных карт могут быть:

- выявление неуправляемого процесса;
- контроль за управляемым процессом;
- оценивание возможностей процесса.

Обычно подлежит изучению следующая переменная величина (параметр процесса) или характеристика:

- известная важная или важнейшая;
- предположительная ненадежная;
- по которой нужно получить информацию о возможностях процесса;
- эксплуатационная, имеющая значение при маркетинге.

При этом не следует контролировать все величины одновременно. Контрольные карты стоят денег, поэтому нужно использовать их разумно: тщательно выбирать характеристики; прекращать работу с картами при достижении цели: продолжать вести карты только тогда, когда процессы и технические требования сдерживают друг друга.

Необходимо иметь в виду, что процесс может быть в состоянии статистического регулирования и давать 100 % брака. И наоборот, может быть неуправляемым и давать продукцию, на 100 % отвечающую техническим требованиям.

Контрольные карты позволяют проводить анализ возможностей процесса. Возможности процесса – это способность функционировать должным образом. Как правило, под возможностями процесса понимают способность удовлетворять техническим требованиям.

Существуют следующие виды контрольных карт:

1. Контрольные карты для регулирования по количественным признакам (измеренные величины выражаются количественными значениями):

а) контрольная карта $\bar{x} - R$ состоит из контрольной карты \bar{x} , отражающей контроль за изменением среднего арифметического, и контрольной карты R , служащей для контроля изменений рассеивания значений показателей качества. Применяется при измерении таких показателей, как длина, масса, диаметр, время, предел прочности при растяжении, шероховатость, прибыль и т. д.;

б) контрольная карта $\tilde{x} - R$ состоит из контрольной карты \tilde{x} , осуществляющей контроль за изменением значения медианы, и контрольной карты R . Применяется в тех же случаях, что и предыдущая карта. Однако она более проста, поэтому более пригодна для заполнения на рабочем месте.

2. Контрольные карты для регулирования по качественным признакам:

а) контрольная карта p (для доли дефектных изделий или процента брака), применяется для контроля и регулирования технологического процесса после проверки небольшой партии изделий и разделения их на доброкачественные и дефектные, т. е. определения их по качественным признакам. Доля дефектных изделий получена путем деления числа об-

наруженных дефектных изделий на число проверенных изделий. Может применяться также для определения интенсивности выпуска продукции, процента неявки на работу и т. д.;

б) контрольная карта pn (количество брака), применяется в случаях, когда контролируемым параметром является число дефектных изделий при постоянном объеме выборки n . Практически совпадает с картой p ;

в) контрольная карта c (число дефектов на одно изделие), используется, когда контролируется число дефектов, обнаруживаемых среди постоянных объемов продукции (автомобили – одна или 5 транспортных единиц, листовая сталь – один или 10 листов);

г) контрольная карта n (число дефектов на единицу площади), используется, когда площадь, длина, масса, объем, сорт непостоянны и обращаться с выборкой как с постоянным объемом невозможно.

При обнаружении дефектных изделий целесообразно прикреплять к ним разные ярлыки: для дефектных изделий, обнаруженных оператором (тип А), и для дефектных изделий, обнаруженных контролером (тип В). Например, в случае А – красные буквы по белому полю, в случае В – черные буквы по белому полю.

На ярлыке указывают номер детали, наименование изделия, технологический процесс, место работы, год, месяц и число, сущность дефекта, число отказов, причину возникновения дефектности, принятые меры воздействия.

В зависимости от целей и задач анализа качества продукции, а также возможностей получения необходимых для его осуществления данных аналитические методы его проведения существенно различаются. Влияет на это и этап жизненного цикла продукции, охватываемый деятельностью предприятия.

6.4. Последовательность применения методов

Последовательность применения инструментов качества может быть различной, в зависимости от специфики производства. Существует мнение, что стратификация и контрольный листок позволяют выявить проблему, а диаграмма рассеивания, гистограмма и контрольная карта способствуют анализу ситуации. Диаграмма Парето и схема Исикавы являются универсальными и способствуют и выявлению проблем, и их анализу.

Применение любого инструмента качества начинается со сбора данных. В первую очередь определяется цель сбора данных и упорядочиваются получаемые результаты, чтобы облегчить процедуру их обработки. Оптимальным инструментом для сбора данных является кон-

трольный листок. Форма контрольного листка разрабатывается в соответствии с конкретной производственной ситуацией. Для применения нескольких инструментов качества не существует определенного алгоритма, однако необходимым условием при составлении плана контроля продукции является системный подход. На первом этапе следует определить цель применения инструментов качества. Необязательно использовать все семь инструментов. Если причина несоответствия была обнаружена при применении части инструментов качества и ее можно устранить с помощью простого решения, следует незамедлительно приступить к действиям без применения остальных инструментов.

Накопленный опыт специалистов в области качества позволяет сформулировать общие пожелания по применению инструментов качества с целью эффективности данных методов:

- высококвалифицированные специалисты по контролю и технологии;
- разработка контрольного листка к каждому из инструментов качества;
- неоднократное использование некоторых инструментов качества во время анализа на одной или разных стадиях процесса;
- опробование различных путей классификации данных для более детального понимания сути проблемы.

На рис. 33 представлен примерный алгоритм применения семи инструментов качества.

Конечно, единого алгоритма на все случаи жизни не существует, вследствие различной специфики производства. Поэтому для каждого технологического процесса последовательность применения инструментов качества различна.

6.5. Цикл PDCA и семь инструментов качества

Рассмотрим связь семи инструментов качества и цикла PDCA (Шухарта–Деминга).

Цикл PDCA представляет собой логическую последовательность повторяющихся действий, направленную на непрерывное совершенствование процессов.

Цикл PDCA (рис. 34) широко используется при управлении процессами системы менеджмента качества. На этапе «Планирование» (Plan – P) устанавливают показатели качества (параметры, которым должна удовлетворять продукция, процесс или услуга) и разрабатывают план действий для их получения и способы достижения в заданные сроки, а также находят исполнителей.

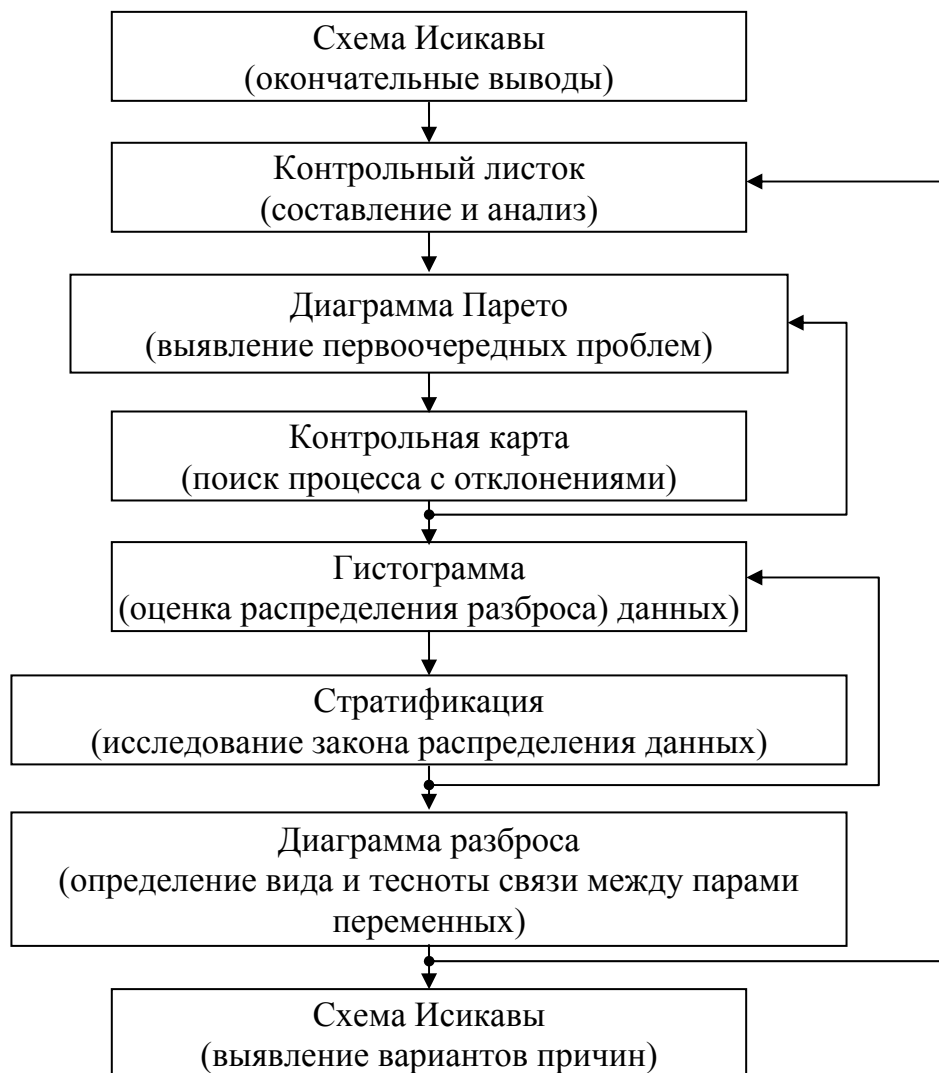


Рис. 33. Примерный алгоритм применения семи инструментов качества



Рис. 34. Цикл PDCA

На этапе «Выполнение» (Do – D) производится продукция с запланированными свойствами и проводится сбор данных, а также мониторинг процесса для определения его стабильности.

На этапе «Проверка» (Check – C) осуществляется диагностика процесса с точки зрения его управляемости, проверяется соответствие показателей качества запланированным значениям, выявляются отклонения.

На этапе «Корректировка» (Act – A) устраняются выявленные причины, осуществляются необходимые изменения.

После корректирующих мероприятий снова проверяется соответствие полученных показателей запланированным значениям. Цикл повторяется, пока результат не совпадет с планом. В связи с этим цикл PDCA способствует непрерывному улучшению качества и является эффективным инструментом для достижения наилучших результатов.

Теперь попытаемся разобраться, какие инструменты целесообразно применять на разных этапах цикла.

На этапе «Plan» целесообразно применять контрольные листки для систематизации большого числа различных данных. После предварительного сбора данных возможно применение других, более сложных, статистических методов и инструментов контроля качества. После построения диаграммы Парето по собранным данным определяются наиболее приоритетные факторы для улучшения. Если диаграмма Парето не позволяет сделать однозначные выводы, то применяют стратификацию данных. Также на данном этапе полезно использовать контрольные карты Шухарта.

На этапе «Do» применение контрольного листка также дает возможность систематизировать получаемые данные, проводить непрерывный мониторинг деятельности по всем направлениям и наглядно представить информацию для ее дальнейшего использования. Целесообразно применять контрольные карты на данном этапе, что позволит контролировать любое изменение показателей процесса, отслеживать его в управляемом состоянии.

На этапе «Check» для выявления наиболее значимых отклонений и определения факторов, оказывающих влияние на показатели процесса, рекомендуется использовать диаграмму Парето и диаграмму Исикавы. Для подтверждения взаимосвязи влияющих факторов и показателей процесса рекомендуется применять диаграмму разброса. Также возможно применять гистограммы для наглядного представления полученных данных и зрительной оценки закона распределения анализируемых данных. Более того, с помощью гистограмм возможно оценить стабильность процесса.

На этапе «Act» для определения того, какие направления изменений являются наиболее приоритетными и необходимыми, рационально использовать диаграмму Исикавы. Определив с ее помощью основные параметры, к улучшению которых следует стремиться, можно воспользоваться диаграммой разброса, чтобы установить, каким образом можно достичь данных параметров.

Как было сказано ранее, повысить эффективность использования статистических инструментов качества можно путем оптимизации порядка их применения, а также с помощью их совместного использования с другими инструментами и методами.

6.6. Мышление в формате А3

В настоящее время многие компании, проводя стратегическое планирование или решая производственные задачи, используют уникальный набор инструментов, которые стали известны как А3, потому что в Японии их печатают на одной стороне листа бумаги размера А3. А3 – это формат технических описаний, разработанный для передачи истории непрерывных улучшений в лаконичном, наглядном и стандартизированном виде.

В компании «Тойота» – инициаторе такой интерпретации формата А3 – на самом деле используют несколько различных типов содержательных документов, которые можно назвать «классическими» формами А3: план-проект работ команды, информационный отчет, отчет о решении задач и отчет А3 о текущем состоянии дел. Большинство используемых компанией «Тойота» форм включают 9 типовых элементов, необходимых для успешного управления проектом:

1. Основная тема.
2. Постановка задачи.
3. Целевая формулировка, определяющая задачи проекта.
4. Научная основа процесса (цикл PDCA), то есть метод, применяемый для изучения проблемы.
5. Методы системного анализа (5 «почему», диаграмма причин и следствий, формат экспериментов).
6. Предлагаемое решение.
7. График выполнения.
8. Графические иллюстрации, облегчающие восприятие информации.
9. Дата и отчетное подразделение или собственник формы указываются внизу формы.

Формат А3 стимулирует творческое мышление, заставляет менеджеров внимательнее прислушиваться к своей аудитории и излагать материал логично, чтобы в лаконичной и доступной форме зафиксировать все, что необходимо, на 1 странице.

Разработанные компанией «Тойота» документы-одностраничники по большей части основаны на графических изображениях процессов, таких как карты потока ценности, которые сжимают всю информацию в формат наглядного изображения, чтобы ее легче было донести до других людей.

А3 не только нагляден – формат одностраничника делает его еще и очень мобильным: он позволяет «управлять на ходу». В результате внедрения этого метода менеджеры начинают обсуждать ход выполнения поставленных задач чаще и в непосредственном контакте со своими подчиненными и руководителями. Многие менеджеры пробивают в А3 отверстия для скоросшивателя и помещают их в специальные блокноты или в папки со спиралью, другие просто складывают А3 и носят в заднем кармане.

А3 – это формат, содействующий написанию хорошей технической документации о научных исследованиях и управлении проектами.

Отчет А3 (отчет формата А3) представляет собой инструмент наглядного отображения информации. Отчет был впервые использован в компании «Тойота» как инструмент решения проблем. Задача ставилась – уместить всю необходимую информацию об анализе и решении проблемы на одной стороне листа формата А3 (11 × 17 дюймов). Такой размер листа был выбран, потому что это самый большой лист бумаги, который можно было отправить по факсу. Формат отчета оказался удачным, его лаконичность и в то же время информационная насыщенность и наглядность полностью удовлетворяли требования руководителей. В компании убеждены, что в меньшем можно выразить большее. Отчет получил распространение и стал использоваться в различных целях:

- для представления, анализа и решения проблемы;
- описания бизнес-процесса;
- представления бюджета, плана;
- презентации проекта, работы кайдзен-групп, кружков качества и т. д.

Отчет А3 воплощает собой процесс принятия решений в компании «Тойота», который опирается на цикл PDCA. То есть в отчете важен не сам формат, а последовательность, логика составления и чтения. Соотношение разделов отчета и этапов цикла приведено в табл. 24.

Цикл PDCA в отчете формата А3

Оцени ситуацию	Название
	Сведения общего характера и история вопроса (ценность, ожидания, политика, цель или план в настоящий момент)
	Ситуация в текущий момент
Планируй	Рекомендации (затраты, результаты)
Делай	Внедрение (детали плана)
Проверяй и действуй	Доработка (ожидаемые результаты)

Отчет создается вместе с самой работой, которую он описывает. И потому работа над отчетом длительная и кропотливая. Обычно отчет претерпевает множество изменений, цель которых – оставить в отчете только самую важную и наглядную информацию. Весь отчет ориентирован на лиц, принимающих решение – руководителей компании, владельцев бизнеса, членов специального комитета.

Отчет А3 принципиально отличается от других обычных бумажных отчетов и служебных записок. Отчет А3 подразумевает обсуждения, размышления, показывает, что было сделано и как. Служебная записка, или отчет в классическом понимании, создается обычно с одной целью – получить статус «Согласовано» и подпись руководителя, а потом отправиться в папку и архив. К тому же ограничение по размеру отчета А3 позволяет быть уверенным в том, что автор оставил только самую важную информацию.

Структуру можно изменять для того, чтобы показать текущий статус или планировать решение проблем по мере их возникновения. На рис. 35 представлен формат А3 о текущем статусе, который можно использовать для отчета о выполнении хосин (стратегический инструмент управления изменениями в критичных бизнес-процессах).

Однако существуют и опасности. Формат А3 рассчитан на очень занятых людей, перегруженных бумажной работой и электронными отчетами. Всего один листок бумаги – это прекрасно, и менеджмент начинает настоятельно рекомендовать форму А3. Может возникнуть и соперничество, если сотрудники пытаются обойти друг друга, строя впечатляющие графики или втискивая на страницу как можно больше информации. Необходимо помнить, что задача А3 – понимание важности вопроса в целом; он создается для того, чтобы организация могла решать проблемы и получать результаты.

Тема			
-------------	--	--	--

I. История вопроса

II. Цели

III. Текущий статус

Цели	Деятель- ность и задачи	Оценка	Комментарии и следующий шаг

IV. Общая оценка

Или

V. Внеплановый анализ

VI. Будущие действия

Нерешен- ная проблема	Действие	Сроки	Ответст- венность

Рис. 35. Текущий статус АЗ

ПРИЛОЖЕНИЕ 1

ЗНАЧЕНИЯ ФУНКЦИИ ЛАПЛАСА

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,00000	0,50	0,19146	1,00	0,34134	1,50	0,43319	2,00	0,47725	3,00	0,49865
0,01	0,00399	0,51	0,19497	1,01	0,34375	1,51	0,43448	2,02	0,47831	3,05	0,49886
0,02	0,00798	0,52	0,19847	1,02	0,34614	1,52	0,43574	2,04	0,47932	3,10	0,49903
0,03	0,01197	0,53	0,20194	1,03	0,34849	1,53	0,43699	2,06	0,48030	3,15	0,49918
0,04	0,01595	0,54	0,20540	1,04	0,35083	1,54	0,43822	2,08	0,48124	3,20	0,49931
0,05	0,01994	0,55	0,20884	1,05	0,35314	1,55	0,43943	2,10	0,48214	3,25	0,49942
0,06	0,02392	0,56	0,21226	1,06	0,35543	1,56	0,44062	2,12	0,48300	3,30	0,49952
0,07	0,02790	0,57	0,21566	1,07	0,35769	1,57	0,44179	2,14	0,48382	3,35	0,49960
0,08	0,03188	0,58	0,21904	1,08	0,35993	1,58	0,44295	2,16	0,48461	3,40	0,49966
0,09	0,03586	0,59	0,22240	1,09	0,36214	1,59	0,44408	2,18	0,48537	3,45	0,49972
0,10	0,03983	0,60	0,22575	1,10	0,36433	1,60	0,44520	2,20	0,48610	3,50	0,49977
0,11	0,04380	0,61	0,22907	1,11	0,36650	1,61	0,44630	2,22	0,48679	3,55	0,49981
0,12	0,04776	0,62	0,23237	1,12	0,36864	1,62	0,44738	2,24	0,48745	3,60	0,49984
0,13	0,05172	0,63	0,23565	1,13	0,37076	1,63	0,44845	2,26	0,48809	3,65	0,49987
0,14	0,05567	0,64	0,23891	1,14	0,37286	1,64	0,44950	2,28	0,48870	3,70	0,49989
0,15	0,05962	0,65	0,24215	1,15	0,37493	1,65	0,45053	2,30	0,48928	3,75	0,49991
0,16	0,06356	0,66	0,24537	1,16	0,37698	1,66	0,45154	2,32	0,48983	3,80	0,49993
0,17	0,06749	0,67	0,24857	1,17	0,37900	1,67	0,45254	2,34	0,49036	3,85	0,49994
0,18	0,07142	0,68	0,25175	1,18	0,38100	1,68	0,45352	2,36	0,49086	3,90	0,49995
0,19	0,07535	0,69	0,25490	1,19	0,38298	1,69	0,45449	2,38	0,49134	3,95	0,49996
0,20	0,07926	0,70	0,25804	1,20	0,38493	1,70	0,45543	2,40	0,49180	4,00	0,49997
0,21	0,08317	0,71	0,26115	1,21	0,38686	1,71	0,45637	2,42	0,49224	4,05	0,49997
0,22	0,08706	0,72	0,26424	1,22	0,38877	1,72	0,45728	2,44	0,49266	4,10	0,49998
0,23	0,09095	0,73	0,26730	1,23	0,39065	1,73	0,45818	2,46	0,49305	4,15	0,49998

Окончание табл.

104

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,24	0,09483	0,74	0,27035	1,24	0,39251	1,74	0,45907	2,48	0,49343	4,20	0,49999
0,25	0,09871	0,75	0,27337	1,25	0,39435	1,75	0,45994	2,50	0,49379	4,25	0,49999
0,26	0,10257	0,76	0,27637	1,26	0,39617	1,76	0,46080	2,52	0,49413	4,30	0,49999
0,27	0,10642	0,77	0,27935	1,27	0,39796	1,77	0,46164	2,54	0,49446	4,35	0,49999
0,28	0,11026	0,78	0,28230	1,28	0,39973	1,78	0,46246	2,56	0,49477	4,40	0,49999
0,29	0,11409	0,79	0,28524	1,29	0,40147	1,79	0,46327	2,58	0,49506	4,45	0,50000
0,30	0,11791	0,80	0,28814	1,30	0,40320	1,80	0,46407	2,60	0,49534	4,50	0,50000
0,31	0,12172	0,81	0,29103	1,31	0,40490	1,81	0,46485	2,62	0,49560	4,55	0,50000
0,32	0,12552	0,82	0,29389	1,32	0,40658	1,82	0,46562	2,64	0,49585	4,60	0,50000
0,33	0,12930	0,83	0,29673	1,33	0,40824	1,83	0,46638	2,66	0,49609	4,65	0,50000
0,34	0,13307	0,84	0,29955	1,34	0,40988	1,84	0,46712	2,68	0,49632	4,70	0,50000
0,35	0,13683	0,85	0,30234	1,35	0,41149	1,85	0,46784	2,70	0,49653	4,75	0,50000
0,36	0,14058	0,86	0,30511	1,36	0,41309	1,86	0,46856	2,72	0,49674	4,80	0,50000
0,37	0,14431	0,87	0,30785	1,37	0,41466	1,87	0,46926	2,74	0,49693	4,85	0,50000
0,38	0,14803	0,88	0,31057	1,38	0,41621	1,88	0,46995	2,76	0,49711	4,90	0,50000
0,39	0,15173	0,89	0,31327	1,39	0,41774	1,89	0,47062	2,78	0,49728	4,95	0,50000
0,40	0,15542	0,90	0,31594	1,40	0,41924	1,90	0,47128	2,80	0,49744	5,00	0,50000
0,41	0,15910	0,91	0,31859	1,41	0,42073	1,91	0,47193	2,82	0,49760		
0,42	0,16276	0,92	0,32121	1,42	0,42220	1,92	0,47257	2,84	0,49774		
0,43	0,16640	0,93	0,32381	1,43	0,42364	1,93	0,47320	2,86	0,49788		
0,44	0,17003	0,94	0,32639	1,44	0,42507	1,94	0,47381	2,88	0,49801		
0,45	0,17364	0,95	0,32894	1,45	0,42647	1,95	0,47441	2,90	0,49813		
0,46	0,17724	0,96	0,33147	1,46	0,42785	1,96	0,47500	2,92	0,49825		
0,47	0,18082	0,97	0,33398	1,47	0,42922	1,97	0,47558	2,94	0,49836		
0,48	0,18439	0,98	0,33646	1,48	0,43056	1,98	0,47615	2,96	0,49846		
0,49	0,18793	0,99	0,33891	1,49	0,43189	1,99	0,47670	2,98	0,49856		

ПРИЛОЖЕНИЕ 2

**КРИТИЧЕСКИЕ ЗНАЧЕНИЯ КРИТЕРИЯ ПИРСОНА
(χ^2 -КРИТЕРИЙ)**

Число степеней свободы df	Уровень значимости α , %							
	20	10	5	2	1	0,5	0,2	0,1
1	1,642	2,706	3,841	5,412	6,635	7,879	9,550	10,83
2	3,219	4,605	5,991	7,824	9,210	10,60	12,43	13,82
3	4,642	6,251	7,815	9,837	11,34	12,84	14,80	16,27
4	5,989	7,779	9,488	11,67	13,28	14,86	16,92	18,47
5	7,289	9,236	11,07	13,39	15,09	16,75	18,91	20,52
6	8,558	10,64	12,59	15,03	16,81	18,55	20,79	22,46
7	9,803	12,02	14,07	16,62	18,48	20,28	22,60	24,32
8	11,03	13,36	15,51	18,17	20,09	21,95	24,35	26,12
9	12,24	14,68	16,92	19,68	21,67	23,59	26,06	27,88
10	13,44	15,99	18,31	21,16	23,21	25,19	27,72	29,59
11	14,63	17,28	19,68	22,62	24,72	26,76	29,35	31,26
12	15,81	18,55	21,03	24,05	26,22	28,30	30,81	32,91
13	16,98	19,81	22,36	25,47	27,69	29,82	32,54	34,53
14	18,15	21,06	23,68	26,87	29,14	31,32	34,09	36,12
15	19,31	22,31	25,00	28,26	30,58	32,80	35,63	37,70
16	20,47	23,54	26,30	29,63	32,00	34,27	37,15	39,25
17	21,61	24,77	27,59	31,00	33,41	35,72	38,65	40,79
18	22,76	25,99	28,87	32,35	34,81	37,16	40,14	42,31
19	23,90	27,20	30,14	33,69	36,19	38,58	41,61	43,82
20	25,04	28,41	31,41	35,02	37,57	40,00	43,07	45,31
21	26,17	29,62	32,67	36,34	38,93	41,40	44,52	46,80
22	27,30	30,81	33,92	37,66	40,29	42,80	45,99	48,27
23	28,43	32,01	35,17	38,97	41,64	44,18	47,39	49,73
24	29,55	33,20	36,42	40,27	42,98	45,56	48,81	51,18
25	30,68	34,38	37,65	41,57	44,31	46,93	50,22	52,62
26	31,79	35,56	38,89	42,86	45,64	48,29	51,63	54,05
27	32,91	36,74	40,11	44,14	46,96	49,64	53,02	55,48
28	34,03	37,92	41,34	45,42	48,28	50,99	54,41	56,89
29	35,14	39,09	42,56	46,69	49,59	52,34	55,79	58,30
30	36,25	40,26	43,77	47,96	50,89	53,67	57,17	59,70
31	37,36	41,42	44,99	49,23	52,19	55,00	58,54	61,10
32	38,47	42,58	46,19	50,49	53,49	56,33	59,90	62,49
33	39,57	43,75	47,40	51,74	54,78	57,65	61,26	63,87
34	40,68	44,90	48,60	53,00	56,06	58,96	62,61	65,25
35	41,78	46,06	49,80	54,24	57,34	60,27	63,95	66,62
36	42,88	47,21	51,00	55,49	58,62	61,58	65,30	67,99
37	43,98	48,36	52,19	56,73	59,89	62,88	66,63	69,35
38	45,08	49,51	53,38	57,97	61,16	64,18	67,97	70,70

Окончание табл.

Число степеней свободы df	Уровень значимости α , %							
	20	10	5	2	1	0,5	0,2	0,1
39	46,17	50,66	54,57	59,20	62,43	65,48	69,29	72,05
40	47,27	51,81	55,76	60,44	63,69	66,77	70,62	73,40
41	48,36	52,95	56,94	61,67	64,95	68,05	71,94	74,74
42	49,46	54,09	58,12	62,89	66,21	69,34	73,25	76,08
43	50,55	55,23	59,30	64,12	67,46	70,62	74,57	77,42
44	51,64	56,37	60,48	65,34	68,71	71,89	75,87	78,75
45	52,73	57,51	61,66	66,56	69,%	73,17	77,18	80,08
46	53,82	58,64	62,83	67,77	71,20	74,44	78,48	81,40
47	54,91	59,77	64,00	68,99	72,44	75,70	79,78	82,72
48	55,99	60,91	65,17	70,20	73,68	76,97	81,08	84,04
49	57,08	62,04	66,34	71,41	74,92	78,23	82,37	85,35
50	58,16	63,17	67,50	72,61	76,15	79,49	83,66	86,66
51	59,25	64,30	68,67	73,82	77,39	80,75	84,94	87,97
52	60,33	65,42	69,83	75,02	78,62	82,00	86,23	89,27
53	61,41	66,55	70,99	76,22	79,84	83,25	87,51	90,57
54	62,50	67,67	72,15	77,42	81,07	84,50	88,79	91,87
55	63,58	68,80	73,31	78,62	82,29	85,75	90,06	93,17
56	64,66	69,92	74,47	79,81	83,51	86,99	91,33	94,46
57	65,74	71,04	75,62	81,01	84,73	88,24	92,61	95,75
58	66,82	72,16	76,78	82,20	85,95	89,48	93,87	97,04
59	67,89	73,28	77,93	83,39	87,17	90,72	95,14	98,32
60	68,97	74,40	79,08	84,58	88,38	91,95	%,40	99,61
61	70,05	75,51	80,23	85,77	89,59	93,19	97,67	100,89
62	71,13	76,63	81,38	86,95	90,80	94,42	98,92	102,17
63	72,20	77,75	82,53	88,14	92,01	95,65	100,18	103,44
64	73,28	78,86	83,68	89,32	93,22	96,88	101,44	104,72
65	74,35	79,97	84,82	90,50	94,42	98,11	102,69	105,99
66	75,42	81,09	85,96	91,68	95,63	99,33	103,94	107,26
67	76,50	82,20	87,11	92,86	96,83	100,55	105,19	108,53
68	77,57	83,31	88,25	94,04	98,03	101,78	106,44	109,79
69	78,64	84,42	89,39	95,21	99,23	103,00	107,69	111,06
70	79,71	85,53	90,53	96,39	100,43	104,21	108,93	112,32
71	80,79	86,64	91,67	97,56	101,62	105,43	110,17	113,58
72	81,86	87,74	92,81	98,73	102,82	106,65	111,41	114,84
73	82,93	88,85	93,95	99,90	104,01	107,86	112,65	116,09
74	84,00	89,96	95,08	101,07	105,20	109,07	113,89	117,35
75	85,07	91,06	96,22	102,24	106,39	110,29	115,12	118,60
76	86,13	92,17	97,35	103,41	107,58	111,50	116,36	119,85
77	87,20	93,27	98,48	104,58	108,77	112,70	117,59	121,10
78	88,27	94,37	99,62	105,74	109,96	113,91	118,82	122,35
79	89,34	95,48	100,75	106,91	111,14	115,12	120,05	123,59
80	90,41	96,58	101,88	108,07	112,33	116,32	121,28	124,84

ПРИЛОЖЕНИЕ 3

КРИТИЧЕСКИЕ ЗНАЧЕНИЯ КОЭФФИЦИЕНТА КОХРЕНА (G-КРИТЕРИЯ)

$df_1 \backslash df_2$	1	2	3	4	5	6	7	8	9	10	16	36	144	∞
для уровня достоверности $\alpha = 0,05$														
2	9,985	9,750	9,392	9,057	8,772	8,534	8,332	8,159	8,010	7,880	7,341	6,602	5,813	5,000
3	9,669	8,709	7,977	7,457	7,071	6,771	6,530	6,333	6,167	6,025	5,466	4,748	4,031	3,333
4	9,065	7,679	6,841	6,287	5,892	5,598	5,365	5,175	5,017	4,884	4,366	3,720	3,093	2,500
5	8,412	6,838	5,981	5,440	5,063	4,783	4,564	4,387	4,241	4,118	3,645	3,066	2,513	2,000
6	7,808	6,161	5,321	4,803	4,447	4,184	3,980	3,817	3,682	3,568	3,135	2,612	2,119	1,667
7	7,271	5,612	4,800	4,307	3,974	3,726	3,535	3,384	3,259	3,154	2,756	2,278	1,833	1,429
8	6,798	5,157	4,377	3,910	3,595	3,362	3,185	3,043	2,926	2,829	2,462	2,022	1,616	1,250
9	6,385	4,775	4,027	3,584	3,286	3,067	2,901	2,768	2,659	2,568	2,226	1,820	1,446	1,111
10	6,020	4,450	3,733	3,311	3,029	2,823	2,666	2,541	2,439	2,353	2,032	1,655	1,308	1,000
12	5,410	3,924	3,264	2,880	2,624	2,439	2,299	2,187	2,098	2,020	1,737	1,403	1,100	0,833
15	4,709	3,346	2,758	2,419	2,159	2,034	1,911	1,815	1,736	1,671	1,429	1,144	0,889	0,667
20	3,894	2,705	2,205	1,921	1,735	1,602	1,501	1,422	1,357	1,303	1,108	0,879	0,675	0,500
24	3,434	2,354	1,907	1,656	1,493	1,374	1,286	1,216	1,160	1,113	0,942	0,743	0,567	0,417
30	2,929	1,980	1,593	1,377	1,237	1,137	1,061	1,002	0,958	0,921	0,771	0,604	0,457	0,333
40	2,370	1,576	1,259	1,082	0,968	0,887	0,827	0,780	0,745	0,713	0,595	0,462	0,347	0,250
60	1,737	1,131	0,895	0,766	0,682	0,623	0,583	0,552	0,520	0,497	0,411	0,316	0,234	0,167
120	0,998	0,632	0,495	0,419	0,371	0,337	0,312	0,292	0,279	0,266	0,218	0,165	0,120	0,083

Окончание табл.

$df_1 \backslash df_2$	1	2	3	4	5	6	7	8	9	10	16	36	144	∞
для уровня достоверности $\alpha = 0,01$														
2	9,999	9,950	9,794	9,586	9,373	9,172	8,988	8,823	8,674	7,539	7,949	7,067	6,062	5,000
3	9,933	9,423	8,831	8,355	7,933	7,606	7,335	7,107	6,912	6,743	6,059	5,153	4,230	3,333
4	9,676	8,643	7,814	7,212	6,761	6,410	6,129	5,897	5,702	5,536	4,884	4,057	3,251	2,500
5	9,279	7,885	6,957	6,329	5,875	5,531	5,259	5,037	4,854	4,697	4,094	3,351	2,644	2,000
6	8,828	7,218	6,258	5,635	5,195	4,866	4,608	4,401	4,229	4,048	3,529	2,858	2,229	1,667
7	8,276	7,664	5,685	5,080	4,659	4,347	4,105	3,911	3,751	3,616	3,105	2,494	1,929	1,429
8	7,945	6,162	5,209	4,627	4,226	3,932	3,704	3,522	3,373	3,248	2,779	2,214	1,700	1,250
9	7,544	5,727	4,810	4,251	3,870	3,592	3,378	3,207	3,067	2,950	2,514	1,992	1,521	1,111
10	7,175	5,358	4,469	3,934	3,572	3,308	3,106	2,945	2,813	2,704	2,297	1,811	1,376	1,000
12	6,528	4,751	3,919	3,428	3,099	2,861	2,680	2,535	2,419	2,320	1,961	1,535	1,157	0,833
15	5,747	4,069	3,317	2,882	2,593	2,386	2,228	2,104	2,002	1,918	1,612	1,251	0,934	0,667
20	4,799	3,297	2,654	2,288	2,048	1,877	1,748	1,646	1,567	1,501	1,248	0,960	0,709	0,500
24	4,247	2,871	2,295	1,970	1,759	1,608	1,495	1,406	1,338	1,283	1,060	0,810	0,595	0,417
30	3,632	2,412	1,913	1,635	1,454	1,327	1,232	1,157	1,100	1,054	0,867	0,658	0,480	0,333
40	2,940	1,951	1,508	1,281	1,135	1,033	0,957	0,898	0,853	0,816	0,668	0,503	0,363	0,250
60	2,151	1,371	1,069	0,902	0,796	0,722	0,668	0,625	0,594	0,567	0,461	0,344	0,245	0,167
120	1,252	0,759	0,585	0,489	0,429	0,387	0,357	0,334	0,316	0,302	0,242	0,178	0,125	0,083

ПРИЛОЖЕНИЕ 4

КРИТИЧЕСКИЕ ЗНАЧЕНИЯ КОЭФФИЦИЕНТА СТЬЮДЕНТА (Т-КРИТЕРИЯ)

Число степеней свободы, <i>df</i>	Доверительная вероятность, <i>p</i> , %							
	80	90	95	98	99	99,5	99,8	99,9
1	3,0777	6,3138	12,7062	31,8205	63,6567	127,3213	318,3088	636,6192
2	1,8856	2,9200	4,3027	6,9646	9,9248	14,0890	22,3271	31,5991
3	1,6377	2,3534	3,1824	4,5407	5,8409	7,4533	10,2145	12,9240
4	1,5332	2,1318	2,7764	3,7469	4,6041	5,5976	7,1732	8,6103
5	1,4759	2,0150	2,5706	3,3649	4,0321	4,7733	5,8934	6,8688
6	1,4398	1,9432	2,4469	3,1427	3,7074	4,3168	5,2076	5,9588
7	1,4149	1,8946	2,3646	2,9980	3,4995	4,0293	4,7853	5,4079
8	1,3968	1,8595	2,3060	2,8965	3,3554	3,8325	4,5008	5,0413
9	1,3830	1,8331	2,2622	2,8214	3,2498	3,6897	4,2968	4,7809
10	1,3722	1,8125	2,2281	2,7638	3,1693	3,5814	4,1437	4,5869
11	1,3634	1,7959	2,2010	2,7181	3,1058	3,4966	4,0247	4,4370
12	1,3562	1,7823	2,1788	2,6810	3,0545	3,4284	3,9296	4,3178
13	1,3502	1,7709	2,1604	2,6503	3,0123	3,3725	3,8520	4,2208
14	1,3450	1,7613	2,1448	2,6245	2,9768	3,3257	3,7874	4,1405
15	1,3406	1,7531	2,1314	2,6025	2,9467	3,2860	3,7328	4,0728
16	1,3368	1,7459	2,1199	2,5835	2,9208	3,2520	3,6862	4,0150
17	1,3334	1,7396	2,1098	2,5669	2,8982	3,2224	3,6458	3,9651
18	1,3304	1,7341	2,1009	2,5524	2,8784	3,1966	3,6105	3,9216
19	1,3277	1,7291	2,0930	2,5395	2,8609	3,1737	3,5794	3,8834
20	1,3253	1,7247	2,0860	2,5280	2,8453	3,1534	3,5518	3,8495
21	1,3232	1,7207	2,0796	2,5176	2,8314	3,1352	3,5272	3,8193
22	1,3212	1,7171	2,0739	2,5083	2,8188	3,1188	3,5050	3,7921
23	1,3195	1,7139	2,0687	2,4999	2,8073	3,1040	3,4850	3,7676
24	1,3178	1,7109	2,0639	2,4922	2,7969	3,0905	3,4668	3,7454
25	1,3163	1,7081	2,0595	2,4851	2,7874	3,0782	3,4502	3,7251
26	1,3150	1,7056	2,0555	2,4786	2,7787	3,0669	3,4350	3,7066
27	1,3137	1,7033	2,0518	2,4727	2,7707	3,0565	3,4210	3,6896
28	1,3125	1,7011	2,0484	2,4671	2,7633	3,0469	3,4082	3,6739
29	1,3114	1,6991	2,0452	2,4620	2,7564	3,0380	3,3962	3,6594
30	1,3104	1,6973	2,0423	2,4573	2,7500	3,0298	3,3852	3,6460

Окончание табл.

Число степеней свободы, <i>df</i>	Доверительная вероятность, <i>p</i> , %							
	80	90	95	98	99	99,5	99,8	99,9
32	1,3086	1,6939	2,0369	2,4487	2,7385	3,0149	3,3653	3,6218
34	1,3070	1,6909	2,0322	2,4411	2,7284	3,0020	3,3479	3,6007
36	1,3055	1,6883	2,0281	2,4345	2,7195	2,9905	3,3326	3,5821
38	1,3042	1,6860	2,0244	2,4286	2,7116	2,9803	3,3190	3,5657
40	1,3031	1,6839	2,0211	2,4233	2,7045	2,9712	3,3069	3,5510
42	1,3020	1,6820	2,0181	2,4185	2,6981	2,9630	3,2960	3,5377
44	1,3011	1,6802	2,0154	2,4141	2,6923	2,9555	3,2861	3,5258
46	1,3002	1,6787	2,0129	2,4102	2,6870	2,9488	3,2771	3,5150
48	1,2994	1,6772	2,0106	2,4066	2,6822	2,9426	3,2689	3,5051
50	1,2987	1,6759	2,0086	2,4033	2,6778	2,9370	3,2614	3,4960
55	1,2971	1,6730	2,0040	2,3961	2,6682	2,9247	3,2451	3,4764
60	1,2958	1,6706	2,0003	2,3901	2,6603	2,9146	3,2317	3,4602
65	1,2947	1,6686	1,9971	2,3851	2,6536	2,9060	3,2204	3,4466
70	1,2938	1,6669	1,9944	2,3808	2,6479	2,8987	3,2108	3,4350
80	1,2922	1,6641	1,9901	2,3739	2,6387	2,8870	3,1953	3,4163
90	1,2910	1,6620	1,9867	2,3685	2,6316	2,8779	3,1833	3,4019
100	1,2901	1,6602	1,9840	2,3642	2,6259	2,8707	3,1737	3,3905
110	1,2893	1,6588	1,9818	2,3607	2,6213	2,8648	3,1660	3,3812
120	1,2886	1,6577	1,9799	2,3578	2,6174	2,8599	3,1595	3,3735
130	1,2881	1,6567	1,9784	2,3554	2,6142	2,8557	3,1541	3,3669
140	1,2876	1,6558	1,9771	2,3533	2,6114	2,8522	3,1495	3,3614
150	1,2872	1,6551	1,9759	2,3515	2,6090	2,8492	3,1455	3,3566
160	1,2869	1,6544	1,9749	2,3499	2,6069	2,8465	3,1419	3,3524
170	1,2866	1,6539	1,9740	2,3485	2,6051	2,8441	3,1389	3,3487
180	1,2863	1,6534	1,9732	2,3472	2,6034	2,8421	3,1361	3,3454
190	1,2860	1,6529	1,9725	2,3461	2,6020	2,8402	3,1337	3,3425
200	1,2858	1,6525	1,9719	2,3451	2,6006	2,8385	3,1315	3,3398
250	1,2849	1,6510	1,9695	2,3414	2,5956	2,8322	3,1232	3,3299
300	1,2844	1,6499	1,9679	2,3388	2,5923	2,8279	3,1176	3,3233
350	1,2840	1,6492	1,9668	2,3370	2,5899	2,8249	3,1137	3,3185
400	1,2837	1,6487	1,9659	2,3357	2,5882	2,8227	3,1107	3,3150
450	1,2834	1,6482	1,9652	2,3347	2,5868	2,8209	3,1084	3,3123
500	1,2832	1,6479	1,9647	2,3338	2,5857	2,8195	3,1066	3,3101
550	1,2831	1,6476	1,9643	2,3331	2,5848	2,8184	3,1051	3,3083
600	1,2830	1,6474	1,9639	2,3326	2,5840	2,8175	3,1039	3,3068

ПРИЛОЖЕНИЕ 5

ЗНАЧЕНИЯ КРИТЕРИЯ ФИШЕРА (F-КРИТЕРИЯ)

$df_1 \backslash df_2$	1	2	3	4	5	6	7	8	9	10	15	20
Уровень значимости $\alpha = 5\%$												
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	245,9	248,0
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,43	19,45
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,70	8,66
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,86	5,80
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,62	4,56
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,94	3,87
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,51	3,44
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,22	3,15
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,01	2,94
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,85	2,77
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,72	2,65
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,62	2,54
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,53	2,46
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,46	2,39
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,40	2,33
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,35	2,28
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,31	2,23

Продолжение табл.

$df_2 \backslash df_1$	1	2	3	4	5	6	7	8	9	10	15	20
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,27	2,19
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,23	2,16
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,20	2,12
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,18	2,10
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,15	2,07
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,13	2,05
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,11	2,03
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,09	2,01
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,07	1,99
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,06	1,97
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,04	1,96
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,03	1,94
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,01	1,93
Уровень значимости $\alpha = 2 \%$												
1	1012,5	1249,5	1350,5	1405,8	1440,6	1464,5	1481,8	1495,0	1505,3	1513,7	1539,1	1551,9
2	48,51	49,00	49,17	49,25	49,30	49,33	49,36	49,37	49,39	49,40	49,43	49,45
3	20,62	18,86	18,11	17,69	17,43	17,25	17,11	17,01	16,93	16,86	16,66	16,55
4	14,04	12,14	11,34	10,90	10,62	10,42	10,27	10,16	10,07	10,00	9,78	9,67
5	11,32	9,45	8,67	8,23	7,95	7,76	7,61	7,50	7,42	7,34	7,12	7,01
6	9,88	8,05	7,29	6,86	6,58	6,39	6,25	6,14	6,05	5,98	5,76	5,65
7	8,99	7,20	6,45	6,03	5,76	5,58	5,44	5,33	5,24	5,17	4,95	4,84

Продолжение табл.

113

8	8,39	6,64	5,90	5,49	5,22	5,04	4,90	4,79	4,70	4,63	4,42	4,30
9	7,96	6,23	5,51	5,10	4,84	4,65	4,52	4,41	4,33	4,26	4,04	3,92
10	7,64	5,93	5,22	4,82	4,55	4,37	4,23	4,13	4,04	3,97	3,76	3,64
11	7,39	5,70	4,99	4,59	4,34	4,15	4,02	3,91	3,83	3,76	3,54	3,43
12	7,19	5,52	4,81	4,42	4,16	3,98	3,85	3,74	3,66	3,59	3,37	3,25
13	7,02	5,37	4,67	4,28	4,02	3,84	3,71	3,60	3,52	3,45	3,23	3,11
14	6,89	5,24	4,55	4,16	3,90	3,72	3,59	3,48	3,40	3,33	3,11	3,00
15	6,77	5,14	4,45	4,06	3,81	3,63	3,49	3,39	3,30	3,23	3,02	2,90
16	6,67	5,05	4,36	3,97	3,72	3,54	3,41	3,30	3,22	3,15	2,93	2,82
17	6,59	4,97	4,29	3,90	3,65	3,47	3,34	3,23	3,15	3,08	2,86	2,74
18	6,51	4,90	4,22	3,84	3,59	3,41	3,27	3,17	3,09	3,02	2,80	2,68
19	6,45	4,84	4,16	3,78	3,53	3,35	3,22	3,12	3,03	2,96	2,74	2,63
20	6,39	4,79	4,11	3,73	3,48	3,30	3,17	3,07	2,98	2,91	2,70	2,58
21	6,34	4,74	4,07	3,69	3,44	3,26	3,13	3,02	2,94	2,87	2,65	2,53
22	6,29	4,70	4,03	3,65	3,40	3,22	3,09	2,99	2,90	2,83	2,61	2,49
23	6,25	4,66	3,99	3,61	3,36	3,19	3,05	2,95	2,87	2,80	2,58	2,46
24	6,21	4,63	3,96	3,58	3,33	3,15	3,02	2,92	2,83	2,77	2,55	2,43
25	6,18	4,59	3,93	3,55	3,30	3,13	2,99	2,89	2,81	2,74	2,52	2,40
26	6,14	4,56	3,90	3,52	3,28	3,10	2,97	2,86	2,78	2,71	2,49	2,37
27	6,11	4,54	3,87	3,50	3,25	3,07	2,94	2,84	2,76	2,69	2,46	2,34
28	6,09	4,51	3,85	3,47	3,23	3,05	2,92	2,82	2,73	2,66	2,44	2,32
29	6,06	4,49	3,83	3,45	3,21	3,03	2,90	2,80	2,71	2,64	2,42	2,30
30	6,04	4,47	3,81	3,43	3,19	3,01	2,88	2,78	2,69	2,62	2,40	2,28

Продолжение табл.

$df_2 \backslash df_1$	1	2	3	4	5	6	7	8	9	10	15	20
Уровень значимости $\alpha=1\%$												
1	4052,2	4999,5	5403,4	5624,6	5763,6	5859,0	5928,4	5981,1	6022,5	6055,8	6157,3	6208,7
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,43	99,45
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	26,87	26,69
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,20	14,02
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,72	9,55
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,56	7,40
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,31	6,16
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,52	5,36
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	4,96	4,81
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,56	4,41
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,25	4,10
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,01	3,86
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,82	3,66
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,66	3,51
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,52	3,37
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,41	3,26
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,31	3,16
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,23	3,08
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,15	3,00

Окончание табл.

20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,09	2,94
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,03	2,88
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	2,98	2,83
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	2,93	2,78
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	2,89	2,74
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	2,85	2,70
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	2,81	2,66
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06	2,78	2,63
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,75	2,60
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,73	2,57
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,70	2,55

СПИСОК ЛИТЕРАТУРЫ

1. Большая советская энциклопедия. В 30 томах. – Москва : «Советская энциклопедия», 1969–1978.
2. Словарь Лопатникова [Электронный ресурс] // Официальный сайт Л.И. Лопатникова, 2012–2013. – Режим доступа: <http://slovarlopatnikov.ru/slovar>
3. Виды познавательной деятельности [Электронный ресурс] // Информ.-справ. портал, 2013. – Режим доступа: studfilosed.ru/voprosy-i-otvety-po-filosofii/27-vidy-poznavatelnoj-deyatelnosti-obydenno-prakticheskaya.html
4. Познание [Электронный ресурс] // Информ.-справ. портал, 2013. – Режим доступа: ru.wikipedia.org/wiki/%CF%E7%E0%E8%E5#.D0.9D.D0.B0.D1.83.D1.87.D0.BD.D0.BE.D0.B5
5. Основы научных исследований : учеб.-метод. пособие / В.А. Власов, А.А. Степанов, Л.М. Зольникова, Б.Б. Мойзес. – Томск : Изд-во ТПУ, 2007. – 202 с.
6. Казарян В.П. Философия науки : учебное пособие. В 2 частях. Часть 1 / В.П. Казарян. – Москва : Макс Пресс, 2003. – 154 с.
7. Голубинцев В.О. Философия для технических вузов / В.О. Голубинцев, А.А. Данцев, В.С. Любченко. – Ростов-на-Дону : Феникс, 2001. – 512 с.
8. Философский энциклопедический словарь / Гл. ред. Л.Ф. Ильичев. – Москва : Советская энциклопедия, 1983. – 836 с.
9. Валеев Г.Х. Методология и методы психолого-педагогических исследований: учеб. пособие для студентов 3–5-х курсов педагогических вузов по специальности «031000 – Педагогика и психология» [Электронный ресурс] // Информ.-справ. портал, 2015. – Режим доступа: www.pedlib.ru/Books/1/0171/1_0171-30.shtml#book_page_top
10. РМГ 29-99 ГСИ. Метрология. Основные термины и определения [Электронный ресурс] // Информ.-справ. портал, 2015. – Режим доступа: www.gosthelp.ru/text/RMG2999GSIMetro-logiyaOsno.html
11. Физика. Влажность воздуха. Точка росы. Психрометр. Гигрометр [Электронный ресурс] // Информ.-справ. портал, 2015. – Режим доступа: www.tepka.ru/fizika/7.31.html
12. Метод и методика измерений [Электронный ресурс] // Информ.-справ. портал, 2005–2015. – Режим доступа: www.metrob.ru/HTML/metrology/izmerenia/metod-izmerenia.html
13. Яворский В.А. Планирование научного эксперимента и обработка экспериментальных данных : учеб.-метод. пособие / В.А. Яворский. – Москва : МФТИ, 2006. – 24 с.

14. Кравченко Н.С. Методы обработки результатов измерений и оценки погрешностей в учебном лабораторном практикуме : учебное пособие / Н.С. Кравченко, О.Г. Ревинская. – Томск : Изд-во Томского политехнического университета, 2011. – 86 с.
15. Гмурман В.Е. Теория вероятностей и математическая статистика : учеб. пособие для вузов / В.Е. Гмурман. – 9-е изд., стер. – Москва : Высш. шк., 2003. – 479 с.
16. Айвазян С.А. Прикладная статистика: основы моделирования и первичная обработка данных : справочное изд. / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин. – Москва : Финансы и статистика, 1983. – 471 с.
17. Корчагин А.Б. Надежность технических систем и техногенный риск. В 2 частях. Часть 1. Основы теории / А.Б. Корчагин, В.С. Сердюк, А.И. Бокарев. – Омск : Изд-во ОмГТУ, 2011. – 229 с.
18. Князев Б.А. Начала обработки экспериментальных данных : учебное пособие / Б.А. Князев, В.С. Черкасский. – Новосибирск : Новосибир. госуд. ун-т, 1996. – 93 с.
19. Пискунов Н.С. Дифференциальное и интегральное исчисления для вузов. В 2 томах. Том 2 / Н.С. Пискунов. – 13-е изд. – Москва : Наука, Главная редакция физико-математической литературы, 1985. – 560 с.
20. Орлов А.И. Математика случая: вероятность и статистика – основные факты : учебное пособие / А.И. Орлов. – Москва : МЗ-Пресс, 2004. – 110 с.
21. Останин А.Н. Применение математических методов и ЭВМ : учеб. пособие для вузов / А.Н. Останин, В.А. Гугля, Н.Н. Гурский. – Минск : Вышэйшая школа, 1989. – 279 с.
22. Шрамков Е.Г. Электрические измерения. Средства и методы измерений (общий курс) / Е.Г. Шрамков. – Москва : Высшая школа, 1972. – 520 с.
23. Вентцель Е.С. Теория вероятностей : учеб. для вузов / Е.С. Вентцель. – 6-е изд., стер. – Москва : Высш. шк., 1999. – 576 с.
24. Лемешко Б.Ю. О выборе числа интервалов в критериях согласия типа χ^2 / Б.Ю. Лемешко, Е.В. Чимитова // Заводская лаборатория. Диагностика материалов. – 2003. – Т. 69. – С. 61–67.
25. Бурдун Г.Д. Основы метрологии / Г.Д. Бурдун, Б.Н. Марков. – Москва : Изд-во стандартов, 1985. – 120 с.
26. Соболев И.М. Численные методы Монте-Карло / И.М. Соболев. – Москва : Наука, 1973. – 312 с.

27. Голодная Н.Ю. Статистическая проверка статистических гипотез : уч.-практ. пособие / Н.Ю. Голодная, Г.И. Шуман. – Владивосток : Владивостокский госуд. ун-т экономики и сервиса, 1994. – 48 с.
28. Мойзес О.Е. Информатика : учебное пособие / О.Е. Мойзес, Е.А. Кузьменко, А.В. Кравцов. – 2-е изд., перераб. и доп. – Томск : Изд-во Томского политехнического университета, 2009. – 152 с.

Учебное издание

МОЙЗЕС Борис Борисович
ПЛОТНИКОВА Инна Васильевна
РЕДЬКО Людмила Анатольевна

СТАТИСТИЧЕСКИЕ МЕТОДЫ КОНТРОЛЯ КАЧЕСТВА И ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

Учебное пособие

Научный редактор *доктор технических наук,
профессор Б.И. Капранов*

Редактор *В.Ю. Пустовалова*
Компьютерная верстка *В.Д. Пяткова*
Дизайн обложки *А.И. Сидоренко*

Подписано к печати 12.09.2016. Формат 60×84/16. Бумага «Снегурочка».
Печать XEROX. Усл. печ. л. 6,92. Уч.-изд. л. 6,26.
Заказ 391-16. Тираж 100 экз.



Издательство

ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ