Владимирский государственный университет

ОСНОВЫ
НАУЧНЫХ ИССЛЕДОВАНИЙ

Практикум для студентов направления подготовки «Техносферная безопасность»

Владимир 2026

Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Владимирский государственный университет
имени Александра Григорьевича и Николая Григорьевича Столетовых»

ОСНОВЫ НАУЧНЫХ ИССЛЕДОВАНИЙ

Практикум для студентов направления подготовки
«Техносферная безопасность»

Электронное издание

Владимир 2026

ISBN 978-5-9984-2237-9

УДК 631.3:636
ББК 40.715

Автор-составитель Е. А. Киндеев

Рецензенты:
Кандидат технических наук, доцент
доцент кафедры технологии машиностроения
Владимирского государственного университета
имени Александра Григорьевича и Николая Григорьевича Столетовых
А. В. Жданов

Ведущий специалист по охране труда
ООО «Фряновский керамический завод»
Е. К. Мегис

Издается по решению редакционно-издательского совета ВлГУ

Основы научных исследований [Электронное издание] : практикум для студентов направления подгот. «Техносферная безопасность» / авт.-сост. Е. А. Киндеев ; Владим. гос. ун-т им. А. Г. и Н. Г. Столетовых. – Владимир : Изд-во ВлГУ, 2026. – 104 с. – ISBN 978-5-9984-2237-9. – Электрон. дан. (2,30 Мб). – 1 электрон. опт. диск (CD-ROM). – Систем. требования: Intel от 1,3 ГГц ; Windows XP/7/8/10 ; Adobe Reader ; дисковод CD-ROM. – Загл. с титул. экрана.

Практикум содержит материалы для проведения практических занятий по дисциплине «Основы научных исследований».

Предназначено для студентов всех форм обучения направления подготовки 20.03.01 «Техносферная безопасность», изучающих дисциплину «Основы научных исследований».

Рекомендовано для формирования профессиональных компетенций в соответствии с ФГОС ВО.

Ил. 10. Табл. 21. Библиогр.: 8 назв.

ISBN 978-5-9984-2237-9

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ 5

Практическая работа № 1
ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ.. 7
Контрольные вопросы 14

Практическая работа № 2
МЕТОДИКА ОТСЕВА ГРУБЫХ ОШИБОК (ВЫЛЕТОВ) 15
Контрольные вопросы.. 22

Практическая работа № 3
МЕТОДИКА ПРОВЕРКИ ГИПОТЕЗЫ СООТВЕТСТВИЯ РАСПРЕДЕЛЕНИЯ ОПЫТНЫХ ДАННЫХ ЗАКОНУ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ. ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ 23
Контрольные вопросы 36

Практическая работа № 4
ПОЛИГОН И ГИСТОГРАММА РАСПРЕДЕЛЕНИЯ ЧАСТОТ НА ОСНОВАНИИ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ 37
Контрольные вопросы 42

Практическая работа № 5
СТАТИСТИЧЕСКИЕ МЕТОДЫ ПОСТРОЕНИЯ ПАРНЫХ ЗАВИСИМОСТЕЙ ПО ЭКСПЕРИМЕНТАЛЬНЫМ ДАННЫМ........ 43
Контрольные вопросы 48

Практическая работа № 6
МАТЕМАТИЧЕСКИЕ МЕТОДЫ ПЛАНИРОВАНИЯ ФАКТОРНОГО ЭКСПЕРИМЕНТА 49
Контрольные вопросы.. 60

Практическая работа № 7
МЕТОДЫ ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ 61
Контрольные вопросы 92

ЗАКЛЮЧЕНИЕ 93

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ 94

ПРИЛОЖЕНИЕ 96

ВВЕДЕНИЕ

Научное исследование представляет собой деятельность, нацеленную на получение новых знаний путем систематического изучения объективной реальности. Это процесс, включающий сбор фактов, анализ полученных данных, выдвижение гипотез и проверку их экспериментальным путём либо наблюдением, моделирование ситуаций и разработку рекомендаций по улучшению существующих процессов или созданию принципиально новых решений. Знания, полученные в результате научных исследований, имеют особую ценность благодаря своей обоснованности, доказанности и универсальности. Научное знание позволяет глубже понимать природу вещей, процессы, происходящие вокруг нас, даёт инструменты для управления этими процессами и решения практических задач.

Существует большое разнообразие направлений научных исследований, соответствующих различным областям человеческой деятельности и уровням познания мира. Например, фундаментальные исследования направлены на изучение общих законов мироздания, а прикладные исследования сосредоточены на применении полученных знаний для практических целей. Каждый вид исследований имеет свою специфику, методологию и требует особых подходов к организации процесса. Поэтому важно осознавать различия между ними и выбирать подходящие методы именно для конкретной области и поставленных задач.

Цель изучения дисциплины «Основы научных исследований» – это формирование у обучающихся методических принципов научного познания и творчества.

Задачи изучения дисциплины:

– научить студентов творчески решать задачи науки и производства;

– помочь освоить элементы технологии приобретения новых знаний, в том числе при самообразовании;

– способствовать формированию навыков обработки результатов НИР, полученных на основе ограниченной выборки.

Курс дисциплины «Основы научных исследований» предполагает проведение практических работ, при этом обучающиеся должны освоить некоторые применяемые методы научных исследований.

Основная цель настоящего практикума состоит в ознакомлении студентов с основными принципами и методами проведения научных исследований. Рассматриваются структура исследовательского процесса и необходимые шаги при проведении практической работы, проводится разбор важнейших методик сбора и анализа полученных данных.

Изучение материала позволит студентам научиться грамотно ставить проблему, чётко формулировать цели и задачи своего проекта и проводить качественные исследования.

Таким образом, освоив материал данного курса, обучающийся сможет приобрести практические навыки, необходимые каждому современному специалисту, который занимается наукой профессионально или интересуется ею как частью общего образования.

Практическая работа № 1
ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

Цель: изучить методику предварительной обработки данных экспериментальных исследований.

Содержание работы

1. Изучить цели предварительной обработки данных.
2. Выявить значение генеральной совокупности и выборки.
3. Вычислить характеристики эмпирических распределений.

Цели предварительной обработки

Методы обработки данных наблюдений базируются на положениях теории вероятностей и математической статистики.

Предварительная обработка результатов измерений или наблюдений необходима для того, чтобы в дальнейшем с наибольшей эффективностью, а главное – корректно, использовать для построения эмпирических зависимостей статистические методы.

Содержание предварительной обработки в основном состоит в определении характеристик распределения опытных данных, отсеивании грубых погрешностей измерения или погрешностей, неизбежно имеющих место при переписывании цифрового материала или при вводе информации в компьютер. Другим важным моментом предварительной обработки данных является проверка соответствия распределения результатов измерения закону нормального распределения. Если эта гипотеза неприемлема, то следует определить, какому закону распределения подчиняются опытные данные, и, если это возможно, преобразовать данное распределение к нормальному. Только после выполнения перечисленных выше операций можно перейти к построению эмпирических формул, применяя, например, метод наименьших квадратов.

Генеральная совокупность и выборка

Для изучения тех или иных явлений природы и общества проводят опыты или наблюдения. Результат опыта, как правило, является случайной величиной, т. е. такой величиной, значение которой нельзя предсказать заранее исходя из условий опыта. Случайная величина обладает целым набором допустимых значений, но в результате каждого отдельного опыта принимает лишь какое-то одно из них. Изменение случайной величины от опыта к опыту связано с неучитываемыми факторами (случайными факторами).

Различают дискретные и непрерывные случайные величины. Возможные значения дискретных случайных величин можно заранее перечислить (количество бракованных изделий в партии продукции). Значения непрерывной случайной величины не могут быть заранее перечислены, они непрерывно заполняют некоторый промежуток числовой оси (температура воздуха в рабочей зоне).

Набор допустимых значений сам по себе слабо характеризует случайную величину. Чтобы полностью охарактеризовать случайную величину, необходимо не только указать, какие значения она может принимать, но и как часто, т. е. указать вероятность появления тех или иных ее значений.

Соотношение, устанавливающее связь между возможными значениями случайной величины и вероятностями их появления, называется законом распределения случайной величины.

Распределение непрерывной случайной величины невозможно задать при помощи вероятностей отдельных значений, так как этих отдельных значений бесконечное количество. Для непрерывных случайных величин используется вероятность того, что в результате опыта значение случайной величины попадет в некоторую заранее намеченную совокупность чисел. Удобно пользоваться вероятностью события

$$X < x,$$

где X – случайная величина; x – произвольное действительное число из области значений случайной величины. Эта вероятность является функцией от x

$$P(X < x) = F(x)$$

и называется функцией распределения случайной величины.

Вместо полного определения случайной величины в виде законов распределения вероятностей ее часто определяют при помощи числовых характеристик, выражающих характерные особенности случайной величины и называемых моментами случайной величины.

На практике исследователь всегда располагает лишь ограниченным числом значений случайной величины, представляющим собой некоторую выборку из генеральной совокупности. Под генеральной совокупностью понимают все допустимые значения случайной величины.

При анализе какой-либо реальной случайной величины, непрерывно изменяющейся во времени (например, температуры, давления и т. п.) под наблюдаемыми значениями случайной величины понимают значения исследуемого параметра в дискретные моменты времени, разделенные таким интервалом, при котором соседние значения можно считать полученными из независимых опытов.

Выборка называется репрезентативной (представительной), если она дает достаточное представление об особенностях генеральной совокупности.

Функция распределения, получаемая по выборке, называется эмпирической или выборочной функцией распределения. Функция распределения, получаемая по генеральной совокупности, называется функцией теоретического распределения.

Если говорить о характеристиках распределений вероятностей, то характеристики теоретических распределений можно рассматривать как характеристики, существующие в генеральной совокупности, а характеристики эмпирических распределений – как выборочные характеристики.

Применяют и другую терминологию. Характеристики распределения вероятностей в генеральной совокупности называют параметрами, а выборочные (эмпирические) значения характеристик – оценками или статистиками.

Хорошая оценка параметра распределения должна обладать свойствами состоятельности, несмещенности и по возможности должна быть эффективной.

Оценка называется состоятельной, если при неограниченном увеличении объема выборки n ее значение сходится по вероятности к оцениваемому параметру.

Оценка называется несмещенной, если при любом объеме выборки n ее математическое ожидание равно оцениваемому параметру. Для несмещенной оценки отсутствует систематическая погрешность, зависящая от объема выборки n.

Оценка называется эффективной, если среди прочих оценок параметра она имеет наименьшую дисперсию.

Вычисление характеристик эмпирических распределений

Необходимо, прежде всего, отметить, что здесь и в дальнейшем речь идет только о непрерывно распределенных случайных величинах.

В случае одномерного эмпирического распределения произвольным моментом порядка k называется сумма k-тых степеней отклонений результатов наблюдений от произвольного числа c, деленная на объем выборки n:

$$m_k = \frac{1}{n}\sum_{i=1}^{n}(x_i - c)^k, \qquad (1.1)$$

где k может принимать любые значения натурального ряда чисел.

Если $c = 0$, то момент называют начальным. Начальный момент первого порядка называется математическим ожиданием (средним значением) случайной величины $\bar{x}$. Действительно, $\bar{x}$ можно определить и по формуле

$$m_1 = \bar{x} = \frac{1}{n}\sum_{i=1}^{n}(x_i - 0)^1. \qquad (1.2)$$

Чаще, чем начальные моменты, применяются центральные моменты.

При $c = \bar{x}$ момент называется центральным. Первый центральный момент равен нулю

$$m_1 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^1 = 0. \qquad (1.3)$$

Второй центральный момент называется смещенной дисперсией

$$m_2 = S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2. \qquad (1.4)$$

Несмещенную оценку для выборочной дисперсии теоретического распределения можно найти по формуле

$$\bar{S}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2. \qquad (1.5)$$

Стоит обратить внимание, что между формулами есть разница: когда мы вычисляем дисперсию, мы делим на n; когда вычисляем выборочную дисперсию, мы делим на $n - 1$ (размер выборки, уменьшенный на единицу), это так называемая **поправка Бесселя**, которая представляет собой действие деления на $n - 1$. Поправка Бесселя учитывает влияние того, что вместо истинного значения случайной величины используется математическое ожидание случайной величины $\bar{x}$, которое не равно истинному значению случайной величины, а лишь достаточно близко к нему. Поэтому несмещенная оценка получается несколько больше, чем смещенная дисперсия.

Выборочные среднеквадратические отклонения соответственно могут быть найдены по формулам:

$$S = \sqrt{S^2}, \tag{1.6}$$

$$\bar{S} = \sqrt{\bar{S}^2}. \tag{1.7}$$

Из других моментов чаще всего используют моменты третьего и четвертого порядков:

$$m_3 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3, \tag{1.8}$$

$$m_4 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4. \tag{1.9}$$

Выборочное значение коэффициента вариации v, являющееся мерой относительной изменчивости наблюдаемой случайной величины, вычисляют по формуле

$$v = \frac{\bar{S}}{\bar{x}}. \tag{1.10}$$

Коэффициент вариации может быть вычислен и в процентах

$$v = \frac{\bar{S}}{\bar{x}} \cdot 100\%. \tag{1.11}$$

Пример. Определить выборочные характеристики ряда наблюдений за пробегом до полного износа 26 автомобильных шин по следующим данным (табл. 1.1). X – пробег до полного износа, тыс. км.

Таблица 1.1

Вычисление характеристик случайной выборки

№ п/п	X, тыс. км	$X-\bar{\bar{X}}$ (1 знак после запятой)	$(X-\bar{X})^2$ (2 знака после запятой)	$(X-\bar{X})^3$ (3 знака после запятой)	$(X-\bar{X})^4$ (4 знака после запятой)
1	50,2	0,4	0,16	0,064	0,0256
2	46,4	–3,4	11,56	–39,304	133,6336
3	52,2	2,4	5,76	13,824	33,1776
4	49,3	–0,5	0,25	–0,125	0,0625
5	48,2	–1,6	2,56	–4,096	6,5536
6	50,5	0,7	0,49	0,343	0,2401
7	54,3	4,5	20,25	91,125	410,0625
8	47,5	–2,3	5,29	–12,167	27,9841
9	50,7	0,9	0,81	0,729	0,6561
10	48,4	–1,4	1,96	–2,744	3,8416
11	52,2	2,4	5,76	13,824	33,1776
12	55,0	5,2	27,04	140,608	731,1616
13	47,2	–2,6	6,76	–17,576	45,6976
14	50,4	0,6	0,36	0,216	0,1296
15	52,7	2,9	8,41	24,389	70,7281
16	45,3	–4,5	20,25	–91,125	410,0625
17	51,2	1,4	1,96	2,744	3,8416
18	49,7	–0,1	0,01	–0,001	0,0001
19	45,0	–4,8	23,04	–110,592	530,8416
20	49,6	–0,2	0,04	–0,008	0,0016
21	50,2	0,4	0,16	0,064	0,0256
22	47,2	–2,6	6,76	–17,576	45,6976
23	52,4	2,6	6,76	17,576	45,6976
24	48,5	–1,3	1,69	–2,197	2,8561
25	49,9	0,1	0,01	0,001	0,0001
26	50,6	0,8	0,64	0,512	0,4096
Σ	1294,8	–9,2E–14	158,74	8,508	2536,566

Таблица 1.2

Результаты расчета статистических данных
выборочных характеристик

$\bar{x}$	m_1	m_2	m_3	m_4	S	$\bar{S}$	v, %
49,8	–3,6E–15	6,105385	0,327231	97,56022	2,47	2,52	5,06

Вывод: для рассмотренной выборки из 26 шин пробег до износа составляет 49,8 ± 2,52 тыс. км.

Порядок выполнения работы

Записать основные цели предварительной обработки данных.

Записать основные характеристики эмпирических распределений (формулы 1.1 – 1.11).

Вычислить характеристики случайной выборки по своему варианту исходных данных (табл. П1) и заполнить таблицу результатов аналогично табл. 1.1.

Выполнить расчет статистических данных выборочных характеристик аналогично табл. 1.2 и сделать выводы.

Контрольные вопросы

1. Назовите цели предварительной обработки данных.
2. Что называют случайной величиной?
3. В чем различие дискретных и непрерывных случайных величин?
4. Что называют законом распределения случайной величины?
5. Что называют генеральной совокупностью?
6. Что значит состоятельная оценка параметра распределения?
7. Что значит несмещенная оценка параметра распределения?
8. Что значит эффективная оценка параметра распределения?
9. Как определить среднее значение выборки?
10. Как определить дисперсию выборки?
11. Как определяют среднеквадратическое отклонение?
12. Как определяют коэффициент вариации?

Практическая работа № 2
МЕТОДИКА ОТСЕВА ГРУБЫХ ОШИБОК (ВЫЛЕТОВ)

Цель: изучить виды ошибок и методику отсева грубых ошибок при обработке результатов экспериментов.

Содержание работы

Изучить основные виды ошибок в научных исследованиях.

Ознакомиться с методикой отсеивания грубых ошибок.

Проверить минимальное и максимальное значения экспериментальных данных на наличие вылетов.

Общие сведения

Каждый результат измерения является случайной величиной. Отклонение результата наблюдения от истинного значения измеряемой величины называется ошибкой наблюдения. Ошибка наблюдения также есть случайная величина, она является результатом действия только случайных факторов. Если обозначить ошибку через ΔX, результат измерения через X, а истинное значение измеряемой величины через A, то

$$\Delta X = X - A. \tag{2.1}$$

Различают ошибки трех видов:

1. *Систематические ошибки* постоянны во всей серии измерений или изменяются по определенному закону. Они могут быть выявлены и устранены введением соответствующих поправок в результаты измерений.

2. *Случайные ошибки* – ошибки измерения, остающиеся после устранения всех выявленных грубых и систематических ошибок. Они вызываются большим количеством таких факторов, эффекты действия которых невозможно выделить и учесть при обработке результатов измерений.

3. *Грубые ошибки (вылеты)* возникают вследствие нарушения основных условий измерения. Результат, содержащий грубую ошибку, резко отличается по величине от остальных результатов измерений.

Существует большое количество способов отсеивания грубых ошибок (аномальных значений). Если в распоряжении экспериментатора имеется выборка небольшого объема (n < 25), то можно воспользоваться методом вычисления максимального относительного отклонения

$$\frac{|x_i - \bar{x}|}{\bar{S}} \le \tau_{1-p}, \qquad (2.2)$$

где x_i – наибольший (или наименьший) элемент выборки, для которой подсчитывались $\bar{x}$ и $\bar{S}$; τ_{1-p} – табличное значение статистики τ (квантили распределения), вычисленной при доверительной вероятности $q = 1 - p$.

Таким образом, для выделения аномального значения вычисляют

$$\tau = |x_i - \bar{x}| \,/\, \bar{S}, \qquad (2.3)$$

которое затем сравнивают с табличным значением τ_{1-p} (табл. 2.1).

Таблица 2.1

Квантили распределения максимального относительного отклонения τ_{1-p}

n	Уровни значимости α				n	Уровни значимости α			
	0,10	0,05	0,025	0,01		0,10	0,05	0,025	0,01
3	1,41	1,41	1,41	1,41	15	2,33	2,49	2,64	2,80
4	1,65	1,69	1,71	1,72	16	2,35	2,52	2,67	2,84
5	1,79	1,87	1,92	1,96	17	2,38	2,55	2,70	2,87
6	1,89	2,00	2,07	2,13	18	2,40	2,58	2,73	2,90
7	1,97	2,09	2,18	2,27	19	2,43	2,60	2,75	2,93
8	2,04	2,17	2,27	2,37	20	2,45	2,62	2,78	2,96
9	2,10	2,24	2,35	2,46	21	2,47	2,64	2,80	2,98
10	2,15	2,29	2,41	2,54	22	2,49	2,66	2,82	3,01
11	2,19	2,34	2,47	2,61	23	2,50	2,68	2,84	3,03
12	2,23	2,39	2,52	2,66	24	2,52	2,70	2,86	3,05
13	2,26	2,43	2,56	2,71	25	2,54	2,72	2,88	3,07
14	2,30	2,46	2,60	2,76					

Если $\tau \le \tau_{1-p}$, то наблюдение не отсеивают, а если это условие не соблюдается, то наблюдение исключают. После исключения того или иного наблюдения или нескольких наблюдений характеристики эмпирического распределения должны быть пересчитаны по данным сокращенной выборки.

Уже высчитанные квантили распределения статистики τ при уровнях значимости α = 0,10; α = 0,05; α = 0,025; α = 0,01 или доверительной вероятности $1 - \alpha = p = 0{,}90$; 0,95; 0,975; 0,99 приводятся в справочниках по математической статистике. На практике обычно используют уровень значимости $\alpha = 0{,}05$ (результат получается с 95%-й доверительной вероятностью).

Процедуру отсева можно повторить и для следующего по абсолютной величине максимального относительного отклонения, но предварительно необходимо пересчитать $\bar{x}$ и $\bar{S}$ для выборки нового объема $n - 1$.

Рассмотрим **другой метод** отсева грубых погрешностей для малой выборки. В этом случае вычисляют

$$\tau' = \frac{|x_i - \bar{x}|}{\sqrt{\frac{n-1}{n}}\bar{S}} \qquad (2.4)$$

и полученный результат сравнивают с критическим значением, взятым из таблиц при соответствующих n и $1 - p$. В данную формулу в отличие от предыдущей введен уточняющий коэффициент

$$\frac{1}{\sqrt{\frac{n-1}{n}}}.$$

При объеме выборки $n > 25$ может быть использован метод, основанный на применении таблиц распределения Стьюдента. Эти таблицы (табл. 2.2) имеются практически в любой книге по математической статистике.

Последовательность действий при использовании данного метода следующая:

1. Из ряда наблюдений (измерений) выбирают наблюдение, имеющее наибольшее отклонение (положительное или отрицательное)

$$d_{\max} = \left|x_{\max(\min)} - \bar{x}.\right| \qquad (2.5)$$

2. Вычисляют максимальное относительное отклонение

$$\tau = d_{\max}/\bar{S}. \qquad (2.6)$$

По таблице распределения Стьюдента (см. табл. 2.2) находят процентные точки t-распределения Стьюдента $t_{(a,\, n-2)}$, где a – процентная точка нормированного выборочного отклонения. Принимают две точки $a = 5$ % и $a = 0{,}1$ %.

Вычисляют критическое значение относительного отклонения τ_p, которое выражается через критическое значение распределения Стьюдента $t_{(a,\, n-2)}$ по формуле

$$\tau_{(a,\,n)} = \frac{t_{(a,\,n-2)}\sqrt{n-1}}{\sqrt{n-2+\left[t_{(a,\,n-2)}\right]^2}}. \qquad (2.7)$$

Таблица 2.2

Процентные точки распределения Стьюдента

ν	40 %	25 %	10 %	5 %	2,5 %	1 %	0,5 %	0,25 %	0,1 %	0,05 %
1	0,3249	1,0000	3,0777	6,3138	12,7062	31,8205	63,6567	127,3213	318,3088	636,62
2	0,2887	0,8165	1,8856	2,9200	4,3027	6,9646	9,9248	14,0890	22,3271	31,599
3	0,2767	0,7649	1,6377	2,3534	3,1824	4,5407	5,8409	7,4533	10,2145	12,924
4	0,2707	0,7407	1,5332	2,1318	2,7764	3,7469	4,6041	5,5976	7,1732	8,6103
5	0,2672	0,7267	1,4759	2,0150	2,5706	3,3649	4,0321	4,7733	5,8934	6,8688
6	0,2648	0,7176	1,4398	1,9432	2,4469	3,1427	2,7074	4,3168	5,2076	5,9588
7	0,2632	0,7111	1,4149	1,8946	2,3646	2,9980	2,4995	4,0293	4,7853	5,4079
8	0,2619	0,7064	1,3968	1,8595	2,3060	2,8965	2,3554	3,8325	4,5008	5,0413
9	0,2610	0,7027	1,3830	1,8331	2,2622	2,8214	2,2498	3,6897	4,2968	4,7809
10	0,2602	0,6998	1,3722	1,8125	2,2281	2,7638	2,1693	3,5814	4,1437	4,5869
11	0,2596	0,6974	1,3634	1,7959	2,2010	2,7181	3,1058	3,4966	4,0247	4,4370
12	0,2590	0,6955	1,3562	1,7823	2,1788	2,6810	3,0545	3,4284	3,9296	4,3178
13	0,2586	0,6938	1,3502	1,7709	2,1604	2,6503	3,0123	3,3725	3,8520	4,2208
14	0,2582	0,6924	1,3450	1,7613	2,1448	2,6245	2,9768	3,3257	3,7874	4,1405
15	0,2679	0,6912	1,3406	1,7530	2,1314	2,6025	2,9467	3,2860	3,7328	4,0728
16	0,2576	0,6901	1,3368	1,7459	2,1199	2,5835	2,9208	3,2520	3,6862	4,0150
17	0,2573	0,6892	1,3334	1,7396	2,1098	2,5669	2,8982	3,2224	3,6458	3,9651
18	0,2571	0,6884	1,3304	1,7341	2,1009	2,5524	2,8784	3,1966	3,6105	3,9216
19	0,2569	0,6876	1,3277	1,7291	2,0930	2,5395	2,8609	3,1737	3,5794	3,8834
20	0,2567	0,6870	1,3253	1,7247	2,0860	2,5280	2,8453	3,1534	3,5518	3,8495
21	0,2566	0,6864	1,3232	1,7207	2,0796	2,5176	2,8314	3,1352	3,5272	3,8193
22	0,2564	0,6858	1,3212	1,7171	2,0739	2,5083	2,8188	3,1188	3,5050	3,7921
23	0,2563	0,6853	1,3195	1,7139	2,0687	2,4999	2,8703	3,1040	3,4850	3,7676
24	0,2562	0,6848	1,3178	1,7109	2,0639	2,4922	2,7969	3,0905	3,4668	3,7454
25	0,2561	0,6844	1,3163	1,7081	2,0595	2,4851	2,7874	3,0782	3,4502	3,7251

Окончание табл. 2.2

ν	40 %	25 %	10 %	5 %	2,5 %	1 %	0,5 %	0,25 %	0,1 %	0,05 %
26	0,2560	0,6840	1,3150	1,7056	2,0555	2,4786	2,7787	3,0669	3,4350	3,7066
27	0,2559	0,6837	1,3137	1,7033	2,0518	2,4727	2,7707	3,0565	3,4210	3,6896
28	0,2558	0,6834	1,3125	1,7011	2,0484	2,4671	2,7633	3,0469	3,4082	3,6739
29	0,2557	0,6830	1,3114	1,6991	2,0452	2,4620	2,7564	3,0380	3,3962	3,6594
30	0,2556	0,6828	1,3104	1,6973	2,0423	2,4573	2,7500	3,0298	3,3852	3,6460
32	0,2555	0,6822	1,3086	1,6939	2,0369	2,24487	2,7385	3,0149	3,3653	3,6218
34	0,2553	0,6818	1,3070	1,6909	2,0322	2,4411	2,7284	3,0020	3,3479	3,6007
36	0,2552	0,6814	1,3055	1,6883	2,0281	2,4345	2,7195	2,9905	3,3326	3,5821
38	0,2551	0,6810	1,3042	1,6860	2,0244	2,4286	2,7116	2,9803	3,3190	3,5657
40	0,2550	0,6807	1,3031	1,6839	2,0211	2,4233	2,7045	2,9712	3,3069	3,5510
42	0,2550	0,6804	1,3020	1,6820	2,0181	2,4185	2,6981	2,9630	3,2960	3,5377
44	0,2549	0,6801	1,3011	1,6802	2,0154	2,4141	2,6923	2,9555	3,2861	3,5258
46	0,2548	0,6799	1,3002	1,6787	2,0129	2,4102	2,6870	2,9488	3,2771	3,5150
48	0,2548	0,6796	1,2994	1,6772	2,0106	2,4066	2,6822	2,9426	3,2689	3,5051
50	0,2547	0,6794	1,2987	1,6759	2,0086	2,4033	2,6778	2,9370	3,2614	3,4960
55	0,2546	0,6790	1,2971	1,6730	2,0040	2,3961	2,6682	2,9247	3,2561	3,4764
60	0,2545	0,6786	1,2958	1,6706	2,0003	2,3901	2,6603	2,9146	3,2317	3,4602
65	0,2544	0,6783	1,2947	1,6686	1,9971	2,3851	2,6536	2,9060	3,2204	3,4466
70	0,2543	0,6780	1,2938	1,6669	1,9944	2,3808	2,6479	2,8987	3,2108	3,4350
80	0,2542	0,6776	1,2922	1,6641	1,9901	2,3739	2,6387	2,8870	3,1953	3,4163
90	0,2541	0,6772	1,2910	1,6620	1,9867	2,3685	2,6316	2,8779	3,1833	3,4019
100	0,2540	0,6770	1,2901	1,6602	1,9840	2,3642	2,6259	2,8707	3,1737	3,3905
120	0,2539	0,6765	1,2886	1,6577	1,9799	2,3578	2,6174	2,8599	3,1595	3,3735
150	0,2538	0,6761	1,2872	1,6551	1,9759	2,3515	2,6090	2,8492	3,1455	3,3566
200	0,2537	0,6757	1,2858	1,6525	1,9719	2,3451	2,6006	2,8385	3,1315	3,3398
250	0,2536	0,6755	1,2849	1,6510	1,9695	2,3414	2,5956	2,8322	3,1232	3,3299
300	0,2536	0,6753	1,2844	1,6499	1,9679	2,3388	2,5923	2,8279	3,1176	3,3233
400	0,2535	0,6751	1,2837	1,6487	1,9659	2,3357	2,5882	2,8227	3,1107	3,3150
500	0,2535	0,6750	1,2832	1,6479	1,9647	2,3338	2,5857	2,8195	3,1066	3,3101

4. Сравнивают значение относительного отклонения τ_p с вычисленными критическими значениями $\tau_{(5\,\%,\,n)}$ и $\tau_{(0{,}1\,\%,\;n)}$.

Максимальные относительные отклонения, полученные в процессе вычисления, могут быть разделены на три группы:

1) $\tau \leq \tau_{(5\,\%,\, n)}$;

2) $\tau_{(5\,\%,\, n)} \leq \tau \leq \tau_{(0{,}1\,\%,\, n)}$;

3) $\tau \geq \tau_{(0{,}1\,\%,\, n)}$.

Наблюдения, попавшие в первую группу, нельзя отсеивать ни в коем случае. Наблюдения второй группы можно отсеять, если только имеются какие-либо другие данные в пользу этой процедуры (например, результаты изучения физических, химических и других свойств данного объекта). Наблюдения третьей группы отсеивают всегда.

После исключения того или иного наблюдения характеристики эмпирического распределения должны быть пересчитаны по данным сокращенной выборки. После чего повторяют процедуру проверки для следующего по абсолютной величине наибольшего отклонения $d_{\max}$.

Пример 2.1. Проверить ряд наблюдений за наработкой деталей до отказа на наличие грубых ошибок (по лабораторной работе № 1).

Выбираем наблюдение, имеющее наибольшее отклонение (табл. П1).

$d_{\max} = |55 - 49{,}8| = 5{,}2.$

Вычисляем $\tau = \frac{d_{\max}}{\bar{S}} = \frac{5{,}2}{2{,}5} = 2{,}08.$

По таблице распределения Стьюдента (см. табл. 2.2) находим процентные точки t-распределения Стьюдента $t_{(a,\, n-2)}$

$t_{(5\,\%,\ 24)} = 1{,}7109,\ \ t_{(0{,}1\,\%,\, 24)} = 3{,}4668.$

Вычисляем критические значения относительного отклонения τ_p:

$$\tau_{(5\,\%,\, 26)} = \frac{1{,}7108 \cdot \sqrt{25}}{\sqrt{24 + 1{,}7109^2}} = \frac{8{,}5545}{5{,}1891} = 1{,}648;$$

$$\tau_{0{,}1\,\%,\, 26)} = \frac{3{,}4668 \cdot \sqrt{25}}{\sqrt{24 + 3{,}4668^2}} = \frac{17{,}334}{6{,}0015} = 2{,}888.$$

Сравниваем значение $\tau = 2{,}08$ с вычисленными критическими значениями $\tau_{(5\,\%,\ 26)}$ и $\tau_{(0{,}1\,\%,\ 26)}$: $1{,}648 < 2{,}08 < 2{,}888$.

Выделяющееся наблюдение не отсеивается, так как в пользу этой процедуры нет каких-либо дополнительных данных.

Пример 2.2. Произведено испытание образцов металлокерамики, пропитанной медью, предназначенных для изготовления антифрикционных накладок. Результаты измерений представлены в табл. 2.3.

Таблица 2.3

Результаты испытаний образцов

Номер образца	Твердость по Роквеллу при d = 2,5 мм; P = 600 Н	Модуль упругости $\sigma_в$, МПа	Относительное удлинение δ, %
1	107 … 109	435	16,5
2	106 … 109	460	17,0
3	102 … 107	440	15,0
4	108 … 109	417	14,8
5	108 … 109	436	16,0
6	105 … 110	460	18,4
7	107 … 110	435	16,5
8	102 … 105	445	15,5

Проверить, не являются ли результаты измерения 6-го образца по относительному удлинению грубой ошибкой, принять уровень значимости $\alpha = 0{,}05$:

Таблица 2.4

К вычислению $\bar{x}$ и $\bar{S}$

№ п/п	x	$x-\bar{x}$	$(x-\bar{x})^2$	τ
1	16,5	0,3	0,0827	0,3151
2	17,0	0,8	0,6202	0,8630
3	15,0	–1,2	1,4702	1,3288
4	14,8	–1,4	1,9952	1,5479
5	16,0	–0,2	0,0452	0,2329
6	18,4	2,2	4,7852	2,3973
7	16,5	0,3	0,0827	0,3151
8	15,5	–0,7	0,5077	0,7808
Σ/n	16,21			

Проверяем соответствие условию $\tau \le \tau_{1-p}$. При заданном уровне значимости $\tau_{1-p} = 2{,}17$ (см. табл. 2.1), для шестого образца $\tau = 2{,}3973$. Условие не выполняется, следовательно, наблюдение отсеивают.

Порядок выполнения работы

1. Записать основные виды ошибок при проведении экспериментальных исследований.
2. Изучить методики отсеивания грубых погрешностей.
3. Изучить порядок проверки результатов эксперимента на наличие грубых ошибок и примеры 2.1 и 2.2.
4. Провести проверку результатов измерений по своему варианту исходных данных (табл. П2) и заполнить таблицу результатов аналогично табл. 2.4, принять уровень значимости $\alpha = 0{,}05$. При обнаружении грубых ошибок исключить их из выборки и провести повторную проверку на наличие грубых ошибок.

Контрольные вопросы

1. Назовите и охарактеризуйте основные виды ошибок.
2. Как отсеивают ошибки методом вычисления максимального относительного отклонения?
3. Как отсеивают ошибки методом, основанным на применении таблиц распределения Стьюдента?

Практическая работа № 3
МЕТОДИКА ПРОВЕРКИ ГИПОТЕЗЫ СООТВЕТСТВИЯ РАСПРЕДЕЛЕНИЯ ОПЫТНЫХ ДАННЫХ ЗАКОНУ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ. ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ

Цель: изучить методику проверки гипотезы соответствия распределения опытных данных закону нормального распределения.

Содержание работы

Изучить методику проверки закона нормального распределения опытных данных.

Проверить нормальность распределения ряда экспериментальных наблюдений.

Вычислить характеристики для показателей асимметрии и эксцесса.

Общие сведения

Если распределение случайной величины подчиняется определенному закону и может быть приближенно описано кривой

$$y = ae^{-bx^2}, \tag{3.1}$$

то такое распределение называют нормальным. Так как к коэффициентам a и b предъявляется требование: $a > 0$ и $b > 0$, то можно говорить о семействе кривых нормального распределения. С увеличением коэффициента a кривая «вытягивается» в высоту, при увеличении коэффициента b кривая «сплющивается».

Нормальное распределение обладает и другими важными свойствами, которые позволяют считать это распределение основой математической статистики. Рассмотрим эти свойства:

1. Ордината y, которая определяет высоту кривой для каждой точки оси Ox (абсциссы), представляет собой плотность вероятности некоторого значения переменной x и определяется формулой

$$y = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{3.2}$$

где σ – среднеквадратическое отклонение теоретического распределения; μ – среднее значение (математическое ожидание) теоретического распределения.

Из формулы (3.2) следует, что нормальное распределение полностью определяется величинами μ и σ, поскольку $\pi = 3{,}141593...$ и $e = 2{,}718282...$ – математические постоянные. Математическое ожидание μ определяет положение кривой распределения по оси Ox. Среднеквадратическое отклонение σ определяет форму кривой. Чем больше σ (разброс данных), тем кривая становится более пологой (ее основание более широкое).

2. Кривая нормального распределения симметрична относительно среднего значения; математическое ожидание, мода и медиана совпадают.

3. Максимум ординаты кривой

$$y_{\max} = \frac{1}{\sigma\sqrt{2\pi}}, \tag{3.3}$$

что при $\sigma = 1$ составляет примерно 0,4.

Если $x \to \pm \infty$, то $y \to 0$, то есть очень большие и очень малые значения переменной x маловероятны.

Примерно 2/3 всех наблюдений лежит в площади, отсекаемой перпендикулярами к оси Ox ($\mu \pm \sigma$). При большом объеме выборки примерно 90 % всех наблюдений лежит между $-1{,}64\sigma$ и $+1{,}64\sigma$. Границы $-0{,}675\sigma$ и $+0{,}675\sigma$ называют вероятными отклонениями; в этом интервале находится около 50 % всех наблюдений.

Для статистических методов построения эмпирических зависимостей очень важно, чтобы результаты наблюдений оказались распределены по нормальному закону распределения, поэтому проверка нормальности распределения – основное содержание предварительной обработки результатов наблюдений.

Некоторое представление о близости эмпирического распределения к нормальному может дать анализ показателей асимметрии и эксцесса.

Показатель асимметрии можно определить, используя данные табл. 2.2, по формуле

$$g_1 = \frac{m_3}{m_2^{3/2}}. \tag{3.4}$$

Для симметричных распределений $m_3 = 0$ и $g_1 = 0$.

Для удобства сравнения эмпирического распределения и нормального в качестве показателя эксцесса принимают величину

$$g_2 = \frac{m_4}{m_2^2} - 3. \tag{3.5}$$

Для нормального распределения $\frac{m_4}{m_2^2} = 3$ и $g_2 = 0$.

Несмещенные оценки для показателей асимметрии и эксцесса определяют по формулам:

$$G_1 = \frac{\sqrt{n(n-1)}}{n-2} g_1, \tag{3.6}$$

$$G_2 = \frac{n-1}{(n-2)\cdot(n-3)}\left[(n+1)g_2 + 6\right]. \tag{3.7}$$

Для проверки гипотезы нормальности распределения следует также вычислить среднеквадратические отклонения для показателей асимметрии и эксцесса

$$S_{G_1} = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}}, \tag{3.8}$$

$$S_{G_2} = \sqrt{\frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}}. \tag{3.9}$$

Если выполняются условия $|G_1| \le 3S_{G_1}$, $|G_2| < 5S_{G_2}$, то гипотеза нормальности исследуемого распределения может быть принята.

Пример 3.1. Сравнить расход топлива автомобилей LADA GRANTA с рабочим объемом двигателя 1,6 л и LADA VESTA, который оснащен двигателем с рабочим объемом 1,8 л.

Результаты измерений занесены в табл. 3.1. Усредненное значение среднеквадратического отклонения

$$\overline{\sigma} = \sqrt{\frac{(n_1-1)\overline{D}_1 + (n_2-1)\overline{D}_2}{n_1+n_2-2}} = \sqrt{\frac{9\cdot 0{,}3 + 9\cdot 0{,}24}{10+10-2}} = 0{,}52.$$

Таблица 3.1

Расход топлива двигателей автомобилей на 100 км пробега

LADA GRANTA	x_1	8,3	8,2	9,2	9,3	9,1	7,9	8,2	8,5	8,0	9,2	Σ	$x_1 = 8{,}6$
	Δx_1	0,3	0,4	0,6	0,7	0,5	0,7	0,4	0,1	0,6	0,6		
	Δx_1^2	0,09	0,16	0,36	0,49	0,25	0,49	0,16	0,01	0,36	0,36	Σ = 2,73	$D_1 = 0{,}3$
LADA VESTA	x_2	8,9	9,4	8,5	9,3	8,3	9,5	8,5	8,1	8,9	8,4		$x_2 = 8{,}8$
	Δx_2	0,1	0,6	0,3	0,5	0,5	0,7	0,3	0,7	0,1	0,4		
	Δx_2^2	0,01	0,36	0,09	0,25	0,25	0,49	0,09	0,49	0,01	0,16	Σ = 2,2	$D_2 = 0{,}24$

Определяем значение критерия Стьюдента

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\tilde{\sigma}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{|8{,}6 - 8{,}8|}{0{,}52\sqrt{\frac{1}{10} + \frac{1}{10}}} = 0{,}86.$$

При числе степеней свободы $v = n - 1 = 10 - 1 = 9$ и уровне значимости $a = 0{,}95$ по табл. 2.2 определяем критическое значение критерия Стьюдента $t_{\text{крит}} = 1{,}8331$.

Поскольку $t < t_{\text{крит}}$ (0,86 < 1,8331), то отличие в расходах топлива сравниваемых автомобилей несущественно.

Существенность различий двух сравниваемых дисперсий проверяют с помощью критерия Фишера (F-критерия)

$$F = \frac{\sigma_x^2}{\sigma_y^2},$$

где σ_x^2 и σ_y^2 – соответственно большее и меньшее значения сравниваемых дисперсий.

Далее по справочным таблицам находят критическое значение критерия Фишера, соответствующее выбранному уровню риска α и числам степеней свободы $k_1 = n_1 - 1$ и $k_2 = n_2 - 1$. Если $F < F_{\text{кр}}$, то различие дисперсий считают несущественным.

Интервальные оценки измеряемых параметров

При малом числе наблюдений замена параметров «a» их точечными оценками $\tilde{a}$ может привести к серьезным ошибкам. В этом случае актуальной становится задача определения точности и надежности полученных оценок. Такая задача в математической статистике решается с помощью построения доверительных интервалов при заданных уровнях доверительной вероятности.

Допустим, что для параметра «a» получена из опыта оценка $\tilde{a}$, значения которой может быть отложено на числовой оси (рис. 3.1).

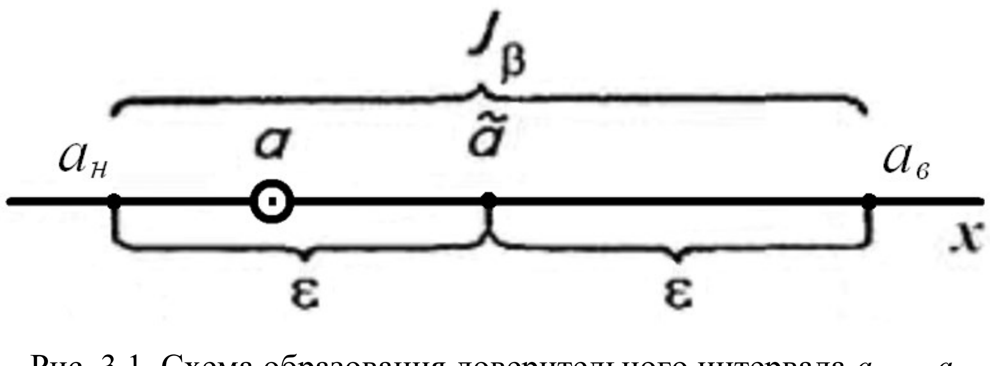


Рис. 3.1. Схема образования доверительного интервала $a_н \ldots a_в$

Достоверно о положении истинного значения параметра «a» можно сказать, что оно окажется в интервале от $-\infty$ до $+\infty$.

Такая информация является бесполезной. Но если назначить некоторый меньший уровень доверительной вероятности β (например, β = 0,9; 0,95; 0,99), но все-таки такой, чтобы событие можно было бы считать практически достоверным, то интервал возможного отклонения ε окажется меньше, и по его значению судят о точности оценки.

Иными словами, вероятность того, что истинное значение окажется внутри интервала ($\tilde{a}$ – ε) и ($\tilde{a}$ + ε) будет равна β, т. е.

$$P[(\tilde{a} - \varepsilon) < a < (\tilde{a} + \varepsilon)] = \beta.$$

Границами интервала будут точки $a_н$ (нижняя), $a_в$ (верхняя). Величина всего интервала

$$I_\beta = [(\tilde{a} - \varepsilon); (\tilde{a} + \varepsilon)] \qquad (3.10)$$

называется *доверительным интервалом*.

Ранее уже отмечалось, что для определения вероятности попадания случайной величины в определенный интервал необходимо знать закон ее распределения $F(x)$.

Вероятность того, что случайная величина x окажется в интервале $x_н \ldots x_в$, равна (рис. 3.2)

$$P(x_н < x < x_в) = F(x_в) - F(x_н). \qquad (3.11)$$

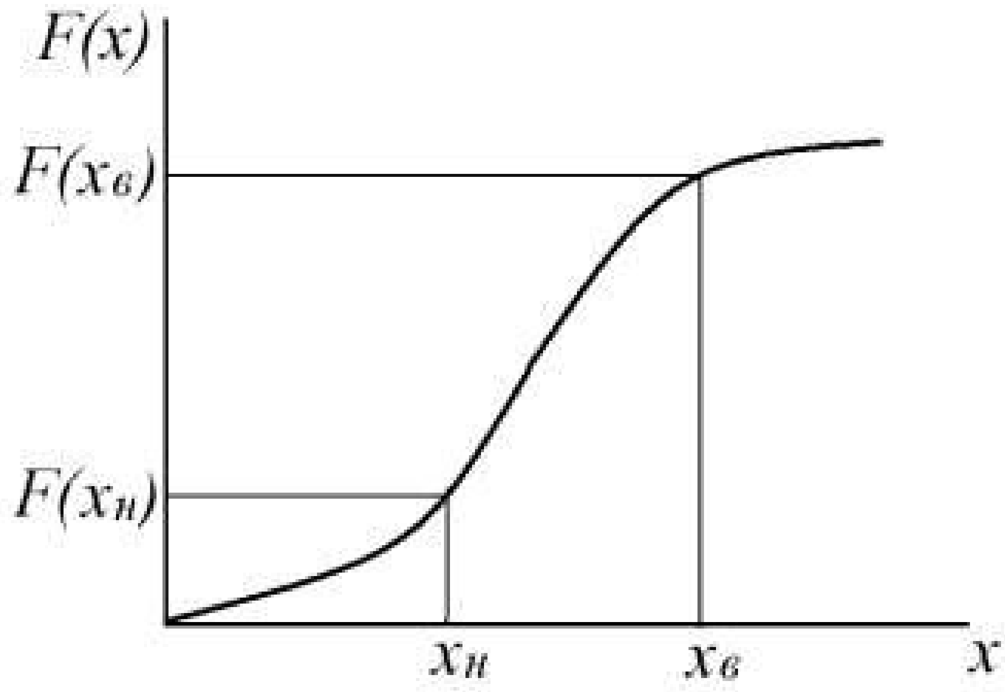


Рис. 3.2. Закон распределения случайной величины

Таким образом, наибольшая сложность в построении доверительных интервалов состоит в определении законов распределения выборочных значений оценок x и σ^2.

Еще в 1908 г. английский математик Уильям Сили Госсет, печатавшийся под псевдонимом Стьюдент, вывел закон распределения величины

$$t = \frac{x - m_x}{\sigma}\sqrt{n}. \tag{3.12}$$

С тех пор этот закон носит название **распределение Стьюдента**, или *t-распределение* с параметром $k = n - 1$ (число степеней свободы).

Плотность распределения Стьюдента равна

$$f_k(t) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi k}\cdot\Gamma\left(\frac{k}{2}\right)}\left(1 + \frac{t^2}{k}\right)^{-\frac{k+1}{2}}, \tag{3.13}$$

где $\Gamma(k) = (k-1)!$ – гамма-функция Эйлера.

Кривая распределения Стьюдента симметрична относительно начала координат, но медленнее приближается к нулю при $t \to \infty$, чем дифференциальная функция нормального распределения (рис. 3.3).

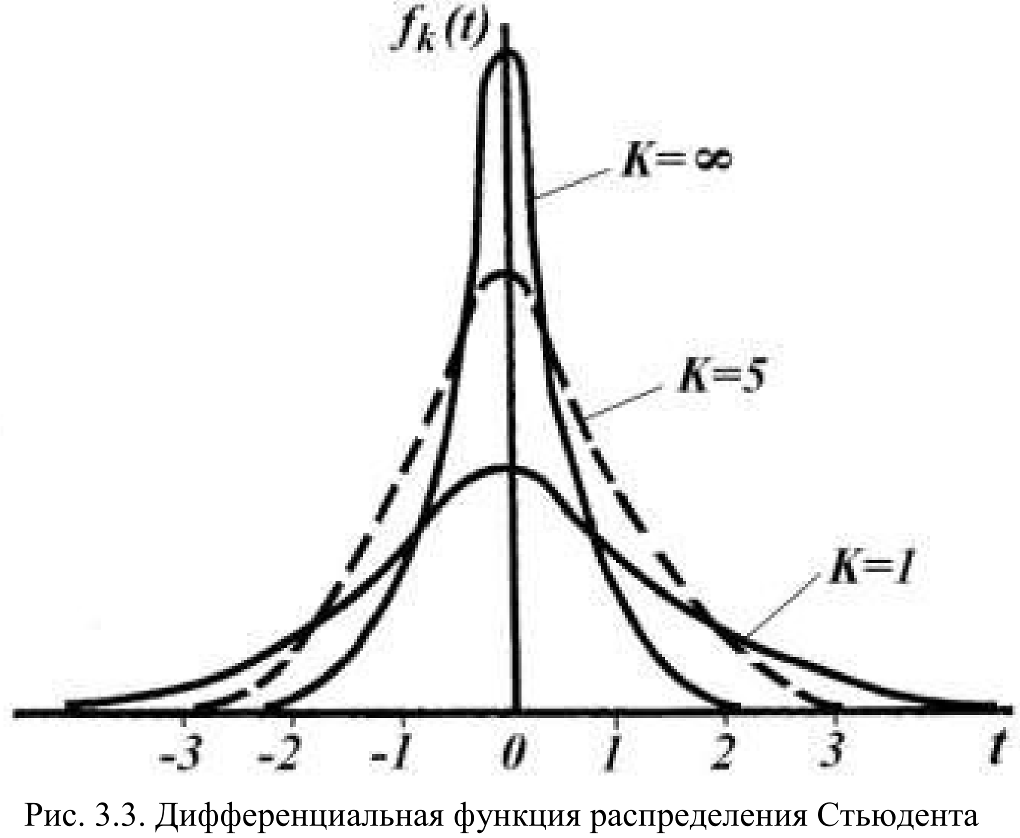


Рис. 3.3. Дифференциальная функция распределения Стьюдента

Считают, что при k = 30 и более распределение Стьюдента становится очень близким к нормальному распределению.

Разность $|(x - m_x)|$ представляет собой возможное отклонение математического ожидания от его оценки x, т. е. $|(\bar{x} - m_x)| \le \varepsilon$.

Поскольку из уравнения (3.12) следует, что

$(\bar{x} - m_x) = \dfrac{t\bar{\sigma}}{\sqrt{n}}$,то $\dfrac{t\bar{\sigma}}{\sqrt{n}} \le \varepsilon$.

Величина t зависит от числа степеней свободы $k = n - 1$ и от доверительной вероятности β, поэтому ее чаще обозначают t_β и находят по таблицам t-распределения Стьюдента.

Доверительный интервал для математического ожидания определится пределами

$$J_\beta = \left(\bar{x} - t_\beta\sqrt{\frac{\bar{D}}{n}};\ \bar{x} + t_\beta\sqrt{\frac{\bar{D}}{n}}\right), \tag{3.14}$$

где $\bar{D}$ – выборочное значение дисперсии, $\bar{D} = \bar{\sigma}^2$.

Рассмотренный выше метод построения доверительного интервала для математического ожидания является общепринятым в математической статистике. Наряду с этим методом ГОСТ Р 50779.22-2005 «Точечная оценка и доверительный интервал для среднего» предусматривает альтернативные методы его вычисления, позволяющие упростить работу и выполнить ее быстрее.

Прежде всего, это способ, основанный на учете «размаха» выборки. Этот метод рекомендуется применять в случаях, когда количество измерений не превышает двенадцати. Практическое удобство этого метода вычисления состоит в его простоте, а недостаток – в том, что вычисление на его основе приводит к более широкому доверительному интервалу.

Если результаты измерений ранжированы так, что $x_1 < x_2 < x_3 < \dots x_n$, то размах выборки можно вычислить как $W = x_n - x_1$.

Двусторонний доверительный интервал для математического ожидания m_x совокупности определяют по следующим формулам:

а) при доверительной вероятности $\beta = 1 - \alpha = 0{,}975$, где α – уровень значимости (или уровень риска)

$$(\bar{x} - q_{0,975}W) < m_x < (\bar{x} + q_{0,975}W); \tag{3.15}$$

б) при доверительной вероятности $\beta = 1 - \alpha = 0{,}995$

$$(\bar{x} - q_{0,995}W) < m_x < (\bar{x} + q_{0,995}W). \tag{3.16}$$

Значения коэффициентов $q_{1-\alpha/2}$ $(q_{0,975}; q_{0,995})$ находят с помощью табл. 3.2.

Помимо двусторонних доверительных интервалов могут быть односторонние. Односторонний доверительный интервал представляет собой оценку в виде

$$m_x < (\bar{x} + q_{0,95}W) \text{ или } m_x > (\bar{x} - q_{0,95}W), \tag{3.17}$$

при доверительной вероятности β = 0,95 и

$$m_x < (\bar{x} + q_{0,99}W) \text{ или } m_x > (\bar{x} - q_{0,99}W) \tag{3.18}$$

при доверительной вероятности β = 0,99.

ГОСТ Р 50779.22-2005 рекомендует следующую методику построения доверительных интервалов для математического ожидания m_x совокупности с использованием распределения Стьюдента.

Двусторонний доверительный интервал определяют по следующим формулам:

а) для доверительной вероятности 0,95:

$$\left(\bar{x} - \frac{t_{0,975}}{\sqrt{n}}S\right) < m_x < \left(\bar{x} + \frac{t_{0,975}}{\sqrt{n}}S\right), \tag{3.19}$$

б) для доверительной вероятности 0,99:

$$\left(\bar{x} - \frac{t_{0,995}}{\sqrt{n}}S\right) < m_x < \left(\bar{x} + \frac{t_{0,995}}{\sqrt{n}}S\right), \tag{3.20}$$

где S – принятое данным ГОСТом обозначение среднего квадратического отклонения выборки, т. е. $S = \sigma$.

Таблица 3.2

Значения $q_{1-\alpha}$ для одностороннего доверительного интервала и значения $q_{1-\alpha/2}$ для двустороннего доверительного интервала

N	$q_{1-\alpha/2}$ – для двустороннего доверительного интервала, доверительной вероятности		$q_{1-\alpha}$ – для одностороннего доверительного интервала	
	β = 0,95	β = 0,99	β = 0,95	β = 0,99
	q0,975	q0,995	q0,95	q0,99
2	6,353	31,828	3,157	15,910
3	1,304	3,008	0,885	2,111
4	0,717	1,316	0,529	1,023
5	0,507	0,843	0,388	0,685
6	0,399	0,628	0,312	0,523
7	0,333	0,507	0,263	0,429
8	0,288	0,429	0,230	0,366
9	0,255	0,374	0,205	0,322
10	0,230	0,333	0,186	0,288
11	0,210	0,302	0,170	0,262
12	0,194	0,277	0,158	0,241

Односторонние доверительные интервалы определяют в этой методике по одной из следующих формул:

а) для доверительной вероятности 0,95:

$$m_x < \left(\bar{x} + \frac{t_{0,95}}{\sqrt{n}} S\right) \text{ или } m_x > \left(\bar{x} - \frac{t_{0,95}}{\sqrt{n}} S\right); \tag{3.21}$$

б) для доверительной вероятности 0,99:

$$m_x < \left(\bar{x} + \frac{t_{0,99}}{\sqrt{n}} S\right) \text{ или } m_x < \left(\bar{x} - \frac{t_{0,99}}{\sqrt{n}} S\right). \tag{3.22}$$

Здесь $t_{0,975}$, $t_{0,995}$, $t_{0,95}$, $t_{0,99}$ – квантили распределения Стьюдента с количеством степеней свободы $v = n + 1$. Их значения даны в табл. 3.3.

В этой же таблице указаны также значения соотношений: $\frac{t_{0,975}}{\sqrt{n}}; \frac{t_{0,995}}{\sqrt{n}}; \frac{t_{0,95}}{\sqrt{n}}; \frac{t_{0,99}}{\sqrt{n}}$.

Если значения n больше 60, то значения t рекомендуется вычислять методом линейной интерполяции $120/n$, используя табл. 3.4.

Пример 3.2

$n = 250$; $120/n = 0,48$;

$t_{0,995} = 2,576 + 0,48\,(2,617 - 2,576) = 2,596$.

Таблица 3.3

Значения $t_{1-\alpha}$ и отношения $\frac{t_{1-\alpha}}{\sqrt{n}}$ для одностороннего интервала, значения $t_{1-\alpha/2}$ и отношения $\frac{t_{1-\alpha/2}}{\sqrt{n}}$ для двустороннего интервала

n	Доверительная вероятность для двустороннего доверительного интервала		Доверительная вероятность для одностороннего доверительного интервала		n	Доверительная вероятность для двустороннего доверительного интервала		Доверительная вероятность для одностороннего доверительного интервала	
	0,95	0,99	0,95	0,99		0,95	0,99	0,95	0,99
	$t_{0,975}$	$t_{0,995}$	$t_{0,95}$	$t_{0,99}$		$\frac{t_{0,975}}{\sqrt{n}}$	$\frac{t_{0,975}}{\sqrt{n}}$	$\frac{t_{0,975}}{\sqrt{n}}$	$\frac{t_{0,975}}{\sqrt{n}}$
1	2	3	4	5	6	7	8	9	10
2	12,71	63,66	6,314	31,82	2	8,985	45,013	4,465	22,501
3	4,303	9,925	2,920	6,965	3	2,484	5,730	1,686	4,021
4	3,182	5,841	2,353	4,541	4	1,591	2,920	1,177	2,270
5	2,776	4,604	2,132	3,747	5	1,242	2,059	0,953	1,676
6	2,571	4,032	2,015	3,365	6	1,049	1,646	0,823	1,374
7	2,447	3,707	1,943	3,143	7	0,925	1,401	0,734	1,188
8	2,365	3,499	1,895	2,998	8	0,836	1,237	0,670	1,060
9	2,306	3,355	1,860	2,896	9	0,769	1,118	0,620	0,966
10	2,262	3,250	1,833	2,821	10	0,715	1,028	0,580	0,892
11	2,228	3,169	1,812	2,764	11	0,672	0,956	0,546	0,833
12	2,201	3,106	1,796	2,718	12	0,635	0,897	0,518	0,785
13	2,179	3,055	1,782	2,681	13	0,604	0,847	0,494	0,744
14	2,160	3,012	1,771	2,650	14	0,577	0,805	0,473	0,708
15	2,145	2,977	1,761	2,624	15	0,554	0,769	0,455	0,668
16	2,131	2,947	1,753	2,602	16	0,533	0,737	0,438	0,651
17	2,120	2,921	1,746	2,583	17	0,514	0,708	0,423	0,627
18	2,110	2,898	1,740	2,567	18	0,497	0,683	0,410	0,605
19	2,101	2,878	1,734	2,552	19	0,482	0,660	0,398	0,586
20	2,093	2,861	1,729	2,539	20	0,468	0,640	0,387	0,568
21	2,086	2,845	1,725	2,528	21	0,455	0,621	0,376	0,552
22	2,080	2,831	1,721	2,518	22	0,443	0,604	0,367	0,537
23	2,074	2,819	1,717	2,508	23	0,432	0,588	0,358	0,523
24	2,069	2,807	1,714	2,500	24	0,422	0,573	0,350	0,510
25	2,064	2,797	1,711	2,492	25	0,413	0,559	0,342	0,498
26	2,060	2,787	1,708	2,485	26	0,404	0,547	0,335	0,487
27	2,056	2,779	1,706	2,479	27	0,396	0,535	0,328	0,477
28	2,052	2,771	1,703	2,473	28	0,388	0,524	0,322	0,467
29	2,048	2,763	1,701	2,467	29	0,380	0,513	0,316	0,658
30	2,045	2,756	1,699	2,462	30	0,373	0,503	0,310	0,449
40	2,024	2,707	1,682	2,430	40	0,320	0,428	0,266	0,384
50	2,008	2,680	1,676	2,404	50	0,284	0,379	0,237	0,340
60	2,000	2,664	1,673	2,393	60	0,258	0,344	0,216	0,309

Таблица 3.4

Значения $t_{0,975}$, $t_{0,995}$, $t_{0,95}$, $t_{0,99}$ при $n \geq 60$

n	$120/n$	$t_{0,975}$	$t_{0,995}$	$t_{0,95}$	$t_{0,99}$
0	2	2,00	2,664	1,673	2,393
120	1	1,980	2,617	1,658	2,358
∞	0	1,960	2,576	1,645	2,326

Представление результатов

1. Представляют выражение для оценки математического ожидания $\bar{x} = \frac{\sum x_i}{n}$ или $\bar{y} = \frac{y_1 n_1 + y_2 n_2 + \cdots + y_n n_n}{n}$.

2. Выражают доверительный интервал в форме неравенства (3.14) или (3.19).

3. Указывают количество результатов, исключенных из-за сомнительности и причины их исключения.

Пример 3.3

Изучается износ сопряжения «втулка шатуна – поршневой палец» у автомобильного двигателя. Ввиду большой трудоемкости опыта (измерения после ремонта, наблюдения в период всего цикла работы, разборка двигателя и измерения по окончании опытов) исследован износ только на 8 двигателях. Результаты измерений втулки шатуна первого цилиндра двигателя представлены в табл. 3.5.

Поскольку измерений мало и необходимости в группировке результатов измерений нет, то

$\bar{x} = \frac{\sum x_i}{n} = \frac{351{,}65}{8} = 43{,}96$ мкм.

Таблица 3.5

Результаты исследования износа втулки шатуна автомобильного двигателя

№ двигателя	Наработка общая, мото-часы	Диаметр начальный, мм	Диаметр после износа, мм	Износ за период наработки, мм	Износ, отнесенный к 1000 мото-часов, мкм	Отклонения от среднего, мкм	Квадрат отклонений, мкм^2
1	1800	42,031	42,125	0,094	52,22	8,26	68,22
2	2200	42,024	42,146	0,122	55,45	11,49	130,02
3	1540	42,028	42,084	0,056	36,36	–7,6	57,76
4	1500	42,034	42,087	0,053	35,33	–8,63	74,48
5	1700	42,027	42,092	0,065	38,24	–5,72	32,72
6	3240	42,032	42,191	0,159	49,07	5,11	26,11
7	1980	42,036	42,108	0,072	36,36	–7,6	57,76
8	3250	42,025	42,183	0,158	48,62	4,66	21,72
Σ	17210				351,65		468,79

Исправленное значение выборочной дисперсии равно

$$\overline{D} = \frac{\sum(x_i - \bar{x})^2}{n-1} = \frac{468{,}79}{8-1} = 66{,}97 \text{ мкм}^2.$$

Оценка среднеквадратического отклонения окажется равной

$$\overline{\sigma} = \sqrt{\overline{D}} = \sqrt{66{,}97} = 8{,}18 \text{ мкм}.$$

Доверительный интервал для среднего значения:

а) по общепринятой методике математической статистики

$$J_\beta = \left(\overline{m}_x - t_\beta \sqrt{\frac{\overline{D}_x}{n}};\ \overline{m}_x + t_\beta \sqrt{\frac{\overline{D}_x}{n}} \right),$$

если, например, $\beta = 0{,}95$ и $k = n - 1 = 8 - 1 = 7$, то $t_{0,95} = 2{,}36$; а в случае $\beta = 0{,}99$, при $k = 7$, $t_{0,99} = 3{,}5$; (значения t_β найдены по таблицам t-распределения Стьюдента):

$$J_{0,95} = \left(43{,}96 - 2{,}36 \sqrt{\frac{66{,}97}{8}}; 43{,}96 + 2{,}36 \sqrt{\frac{66{,}97}{8}} \right) = (37{,}13; 50{,}79).$$

$$J_{0,99} = \left(43{,}96 - 3{,}5 \sqrt{\frac{66{,}97}{8}}; 43{,}96 + 3{,}5 \sqrt{\frac{66{,}97}{8}} \right) = (33{,}83; 54{,}09);$$

б) по ГОСТ Р 50779.22-2005 с учетом «размаха» выборки

$W = x_{\max} - x_{\min} = 55{,}45 - 35{,}33 = 20{,}12$ мкм;

значения коэффициентов $q_{0,975} = 0{,}288$ и $q_{0,995} = 0{,}429$ находят по табл. 3.3;

$$J_{0,95} = \left[\left(\bar{x} - q_{0,975}W\right);\ \left(\bar{x} + q_{0,975}W\right)\right] =$$
$$= [(43{,}96 - 0{,}288 \cdot 20{,}12);\ (43{,}96 + 0{,}288 \cdot 20{,}12)]$$
$$= [(38{,}17);\ (49{,}75)];$$
$$J_{0,99} = \left[\left(\bar{x} - q_{0,995}W\right);\ \left(\bar{x} + q_{0,995}W\right)\right] =$$
$$= [(43{,}96 - 4{,}29 \cdot 20{,}12);\ (43{,}96 + 4{,}29 \cdot 20{,}12)]$$
$$= [(35{,}33);\ (52{,}59)];$$

в) по ГОСТ-Р50779.22-2005 с учетом вычисленных отношений $\frac{t_{1-\alpha/2}}{\sqrt{n}}$ и уравнению (3.22).

Отношение $\frac{t_{0,975}}{\sqrt{n}}$ по табл. 3.2 для β = 0,95 равно 0,836, а для β = 0,99 равно 1,237, учитывая при этом, что $n = 8$.

В этом случае если β = 0,95, то

$(43{,}96 - 0{,}836 \cdot 8{,}18) < m_x < (43{,}96 + 0{,}836 \cdot 8{,}18)$ или $37{,}12 < m_x << 50{,}80$;

а при β = 0,99

$(43{,}96 - 1{,}237 \cdot 8{,}18) < m_x < (43{,}96 + 1{,}237 \cdot 8{,}18)$ или $33{,}84 < m_x << 54{,}03$.

Отличия последнего варианта (в) от первого (а) можно объяснить только точностью используемых таблиц t-распределения Стьюдента.

Содержание отчета

В отчете изложить:

– проверку соответствия распределения ряда наблюдений путем анализа показателей асимметрии и эксцесса закону нормального распределения;

– определение несмещенных оценок для показателей асимметрии G_1 и эксцесса G_2;

– расчет среднеквадратических отклонений для показателей асимметрии и эксцесса (на основании данных табл. 3.1);

– расчет доверительного интервала для среднего значения износа сопряжения «втулка шатуна – поршневой палец» у автомобильного двигателя» по своему варианту исходных данных (табл. П3);

– сделать четыре вывода по работе.

Контрольные вопросы

1. Как изменяется кривая нормального распределения при изменении коэффициентов *a* и *b*?

2. Назовите основные свойства нормального распределения полученных экспериментальных данных.

3. Как определить асимметрию?

4. Как определить эксцесс?

5. Как определяют несмещенные оценки асимметрии и эксцесса?

6. Как определить среднеквадратические отклонения для показателей асимметрии и эксцесса?

7. Назовите условие для принятия гипотезы нормального распределения оценок исследуемого процесса, полученных опытным путем.

8. Как рассчитать доверительный интервал для математического ожидания измеряемой величины?

Практическая работа № 4
ПОЛИГОН И ГИСТОГРАММА РАСПРЕДЕЛЕНИЯ ЧАСТОТ НА ОСНОВАНИИ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

Цель: изучить методику построения полигона и гистограммы распределения частот.

Содержание работы

1. Определить интервалы распределения экспериментальных данных.
2. Рассчитать частоты распределения экспериментальных данных.
3. Начертить гистограмму, полигон частот и кумулятивную линию (интегральная функция случайной величины).

Общие сведения

Если данные наблюдения за непрерывной случайной величиной разделить на интервалы, то можно построить полигон и гистограмму частот.

Разбиение на интервалы можно выполнить по правилу Стерджесса. Число интервалов при этом определяется по формуле

$$k = 1 + 3{,}322 \cdot \lg n, \tag{4.1}$$

где k – количество интервалов (округляется до ближайшего целого числа); n – объем выборки.

Разбиение на интервалы выполняется в табличной форме (табл. 4.1). В этой же таблице ведется подсчет частот. Методика работы при этом следующая: таблицу наблюдений просматривают по порядку от первой до последней строчки и при чтении каждого результата соответствующую метку (точку или черточку) заносят в тот интервал, к которому относится данное наблюдение (эту работу удобно выполнять вдвоем).

По результатам подсчета абсолютных и относительных частот строят гистограмму *1* и полигон *2* распределений, а также кумулятивную линию *3* (рис. 4.1). Гистограмма и полигон распределений являются графическим отображением частот, которые, в свою очередь, представляют собой оценки плотностей вероятностей. Кумулятивная линия – график накопленных частот, в свою очередь оценивающих функцию распределения $F(x)$ в точке x.

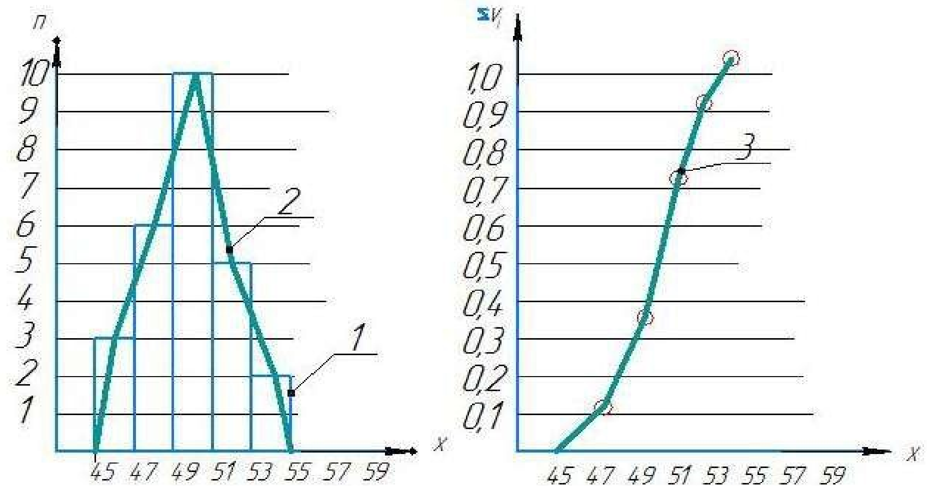


Рис. 4.1. Гистограмма (1), полигон (2) частот и кумулятивная линия (интегральная функция случайной величины) (3)

Пример 4.1. Построить гистограмму и полигон распределения частот по данным наблюдений за наработкой деталей до отказа.

В данном случае $k = 1 + 3{,}322 \cdot \lg 26 = 5{,}7$.

С другой стороны, разница между x_{max} и x_{min} (размах варьирования) составляет 55 – 45 = 10 тыс. ч. Исходя из этого, примем число интервалов равным 5 с шириной интервала, равной 2 тыс. ч.

Разбивку ряда наблюдений на интервалы и подсчет частот сводим в табл. 4.1.

Таблица 4.1

Разбивка массива исходных данных на интервалы и вычисление частот

№ п/п	Интервал	Середины интервалов	Подсчет частот	Частоты		
				абсолютные	относительные	относительные накопленные
1	От 45 до 47	46	///	3	0,115	0,115
2	От 47 до 49	48	//////	6	0,231	0,346
3	От 49 до 51	50	//////////	10	0,385	0,731
4	От 51 до 53	52	/////	5	0,192	0,923
5	От 53 до 55	54	//	2	0,077	1,000

По данным табл. 4.1 строим гистограмму и полигон распределения частот, а также кумулятивную линию (интегральную кривую) (см. рис. 4.1).

Пример 4.2. В качестве примера приведены результаты сопоставления фактического и теоретического распределения времени пребывания машин на станции технического обслуживания (табл. 4.2).

Таблица 4.2

Проверка гипотезы о законе распределения времени обслуживания машин

Время обслуживания t_i, ч	0…1	1…2	2…3	3…4	4…5	5…6	6…7	7…8	8…9	9…10	Σ
Наблюдаемая частота m_i	370	244	147	95	58	35	22	14	8	7	1000
Произведение $m_i \cdot t_i$	185	366	367,5	332,5	261	192,5	143	105	68	66,5	2087
Накопленная вероятность фактического ряда $\bar{F}(x)$	0,37	0,614	0,761	0,856	0,914	0,949	0,971	0,985	0,993	1,000	
Теоретическая частота nP_i	381	236	146	90	56	35	21	13	9	6	992
Разность $\Delta_i = m_i - nP_i$	–10,72	8,23	0,99	4,58	2,00	0,32	0,53	0,70	–0,24	1,90	
$(\Delta_i)^2 = (m_i - nP_i)^2$	114,95	67,69	0,98	20,97	4,02	0,10	0,28	0,49	0,06	3,61	
$(m_i - nP_i)^2/nP_i$	0,302	0,287	0,007	0,232	0,072	0,003	0,013	0,037	0,007	0,708	1,667
Накопленная частота теоретического ряда $F(x)$	0,381	0,616	0,763	0,853	0,909	0,944	0,965	0,978	0,987	0,992	
$\lvert D_\lambda \rvert = F(x) - \bar{F}(x)$	0,011	0,002	0,002	0,003	0,005	0,005	0,006	0,007	0,006	0,008	

В теории массового обслуживания построена модель, согласно которой время обслуживания подчиняется показательному (экспоненциальному) распределению

$$F(t) = 1 - e^{-\lambda t},$$

где λ – плотность потока событий;

$$\lambda = 1/ t_{\text{ср}},$$

где $t_{\text{ср}}$ – среднее время обслуживания машины.

Среднее время обслуживания составляет

$$t_{\text{ср}} = \frac{\sum(n_i \cdot t_i)}{\sum n_i} = \frac{2087}{1000} = 2{,}1 \text{ ч},$$

где t_i – время середины каждого класса.

Плотность потока машин

$$\lambda = \frac{1}{t_{\text{ср}}} = \frac{1}{2{,}1} = 0{,}48.$$

Сумма квадратов отклонений составляет

$$\sum \chi^2 = \sum \frac{(m_i - n_i \cdot P_i)^2}{n_i \cdot P_i} = 1{,}667.$$

Число степеней свободы равно

$$r = K - 1 = 10 - 1 = 9,$$

где K – число классов в распределении времени обслуживания.

Критическое значение $\Sigma\chi^2$ (для $P = 0{,}99$ и $r = 9$, по таблице распределения Пирсона, χ^2) : $\sum \chi^2_{\text{крит}} = 2{,}09$.

Поскольку $\sum \chi^2 < \sum \chi^2_{\text{крит}}$, то можно полагать, что фактические данные не противоречат гипотезе о показательном (экспоненциальном) распределении времени обслуживания машин.

Вероятность того, что расхождение рядов можно объяснить случайными событиями, составляет $P = 0{,}99$.

В качестве меры расхождения между теоретическим и фактическим распределением по критерию А. Н. Колмогорова используют максимальное значение модуля разности между статистической и соответствующей теоретической функциями распределения (рис. 4.2).

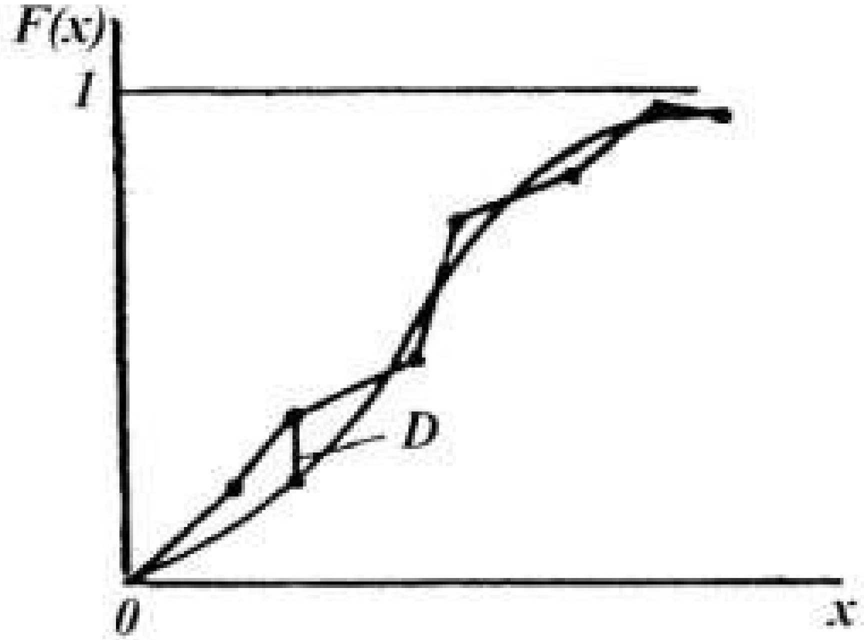


Рис. 4.2. Определение максимального модуля разности между фактическим и теоретическим распределением

А. Н. Колмогоров доказал, что при любой функции распределения $F(x)$ и достаточно большом числе наблюдений вероятность неравенства $D_{\max}\sqrt{n} \geq \lambda$ стремится к пределу

$$P_{\lambda} = 1 - \sum_{k \to \infty}^{\infty} (-1)^{K} \cdot e^{-2K^2\lambda^2}. \qquad (4.2)$$

При малых значениях вероятности P_{λ} гипотезу следует отвергать, а при достаточно большом – считать опытные данные не противоречащими выдвинутой гипотезе.

Наибольшая разность теоретического $F(x)$ и фактического распределения в предыдущем примере составляет $D_{\max} = 0{,}011$.

В этом случае $\lambda = D_{\max}\sqrt{n} = 0{,}348 \approx 0{,}35$, что соответствует $P_{\lambda} = 0{,}9997$, т. е. и по этому критерию опытные данные не противоречат гипотезе о показательном распределении времени обслуживания машин.

Содержание отчета

В отчете изложить:

– порядок построения гистограммы и полигона распределения частот по данным наблюдений за наработкой деталей до отказа с построением графиков (гистограмма, полигон и кумулятивная линия) с использованием примера 4.1.

– по своему варианту исходных данных (табл. П4) вычислить частоты абсолютные, относительные и относительные накопленные, разбив массив исходных данных на интервалы.

– методику сопоставления фактического распределения времени пребывания машин на станции технического обслуживания с теоретическим с использованием примера 4.2.

– сделать три вывода по работе.

Контрольные вопросы

1. Как определяется количество интервалов?
2. Что такое гистограмма и полигон частот?
3. Что такое кумулятивная линия?

Практическая работа № 5
СТАТИСТИЧЕСКИЕ МЕТОДЫ ПОСТРОЕНИЯ ПАРНЫХ ЗАВИСИМОСТЕЙ ПО ЭКСПЕРИМЕНТАЛЬНЫМ ДАННЫМ

Цель: изучить методику построения регрессионной зависимости.

Содержание работы

1. Описать методику построения регрессионной зависимости.
2. Выполнить лабораторные исследования.
3. Вычислить коэффициенты уравнения регрессии по экспериментальным данным.

Метод наименьших квадратов в простейшем случае двумерного пространства (на плоскости). Уравнение регрессии

Подавляющее большинство всех формул, используемых в естественно-научных и технических дисциплинах, относится к так называемым парным зависимостям типа $y = f(x)$. По результатам экспериментов такие формулы обычно строят при помощи метода наименьших квадратов.

Сама по себе процедура линейного парного регрессионного анализа (метода наименьших квадратов на плоскости) очень проста, для ее выполнения достаточно использовать калькулятор.

Регрессия – это односторонняя вероятностная зависимость, когда одна из переменных служит причиной для изменения другой.

Пусть имеется n пар наблюдений значений функции отклика $y_{i,}$, полученных при фиксированных значениях независимой переменной фактора x_i. Для графического изображения на плоскости этих пар наблюдений в виде экспериментальных точек с координатами ($x_i; y_i$) применяется система декартовых координат.

Координаты точек 1 – 8, изображенных на рис. 5.1, приведены в табл. 5.1. Подобные результаты наблюдений могут быть получены в любой экспериментальной работе.

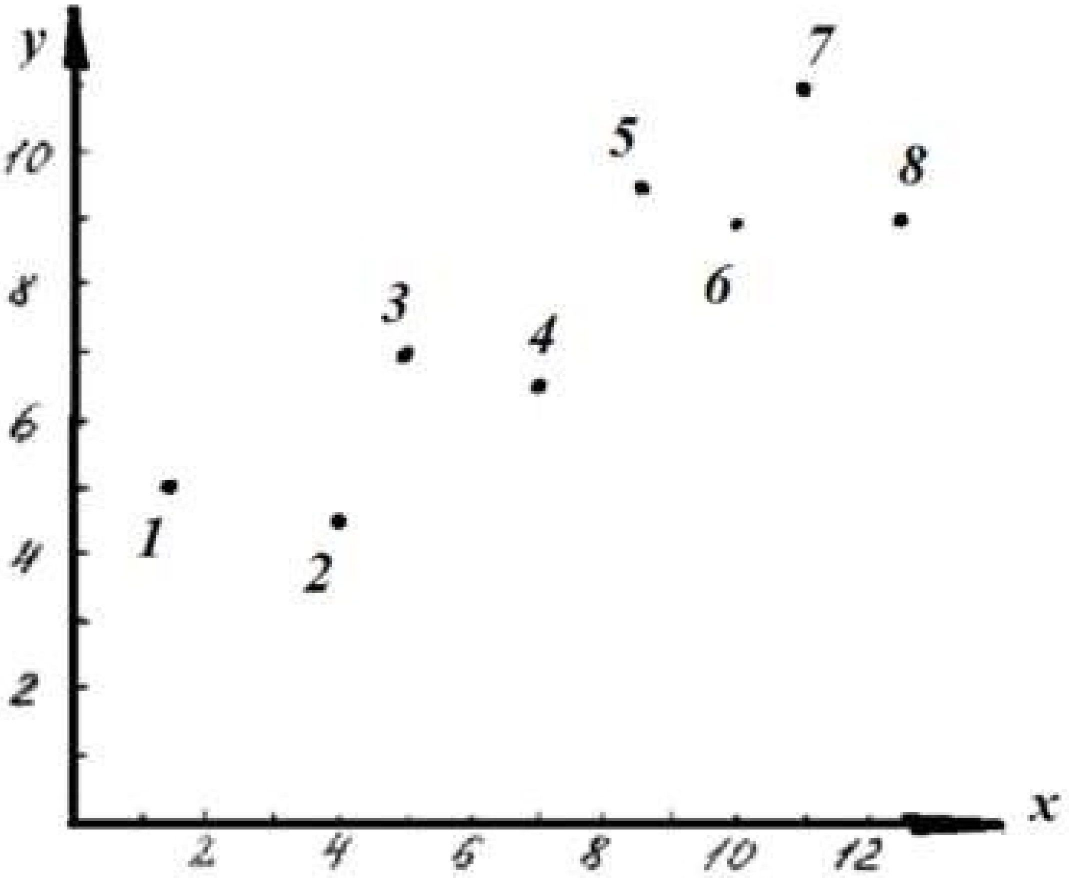


Рис. 5.1. Графическое изображение координат точек (x_i; y_i)

Таблица 5.1

Координаты точек для графика (рис. 5.1)

Точка / Координаты	1	2	3	4	5	6	7	8
x	1,5	4,0	5,0	7,0	8,5	10,0	11,0	12,5
y	5,0	4,5	7,0	6,5	9,5	9,0	11,0	9,0

Задача линейного регрессионного анализа (метода наименьших квадратов) состоит в том, чтобы, зная положение точек 1 – 8 на плоскости, так провести линию регрессии, чтобы сумма квадратов отклонений Δ_i^2 вдоль оси Oy (ординаты) этих точек от проведенной прямой была минимальной.

Для проведения вычислений по классическому методу наименьших квадратов (для проведения регрессионного анализа) к выдвигаемой гипотезе (к форме уравнения регрессии) предъявляется такое требование: это уравнение должно быть линейным по параметрам или допускать возможность линеаризации. Так, например, процедура проведения регрессионного анализа одинакова для уравнений $y = b_0 + b_1 \cdot x$ и $y = b_0 + b_1 \cdot z^2$, так как подстановка $x = z^2$ приводит второе уравнение к первому. Для упрощения усвоения начальных элементарных сведений, относящихся к методике построения формул по опытным данным, и более легкого освоения методики регрессионного анализа предположим, что при проведении парного линейного регрессионного анализа имеем дело только с уравнением прямой линии.

Уравнение прямой на плоскости в декартовых координатах

$$y = b_0 + b_1 x, \tag{5.1}$$

где b_0, b_1 – постоянные числа, геометрическая интерпретация которых дана ниже.

С учетом указанного выше задачу метода наименьших квадратов аналитически можно выразить следующим образом:

$$U = \sum_{i=1}^{n} [y_i - (b_0 + b_1 x_i)]^2 \to \min, \tag{5.2}$$

где $y_i - (b_0 + b_1 x_i) = \Delta_i$

$$\text{или } U = \sum_{i=1}^{n} \Delta_i^2 \to \min. \tag{5.3}$$

Формулы (5.2) и (5.3) кратко можно выразить так: сумма квадратов отклонений вдоль оси Oy должна быть минимальной (принцип Лежандра).

Построенная таким образом линия регрессии позволяет с некоторой вероятностью предсказать в интервале от $x = 1{,}5$ до $x = 12{,}5$ значения функции y при отсутствующих в табл. 5.1 значениях фактора x.

Для решения задачи, поставленной в формуле (5.2), необходимо в каждом конкретном случае вычислить значения коэффициентов b_0 и b_1, минимизирующие сумму отклонений U. Как известно из математического анализа, для этого необходимо вычислить частные производные функции U по коэффициентам b_0 и b_1 и приравнять их к нулю

$$\begin{cases} \dfrac{\partial U}{\partial b_0} = 0 \\ \dfrac{\partial U}{\partial b_1} = 0 \end{cases}. \tag{5.4}$$

Решая эту систему уравнений, находим искомые значения b_0 и b_i. Систему (5.4) называют системой нормальных уравнений. В нее подставляют значение U из формулы (5.2) и одновременно выполняют операцию дифференцирования

$$\begin{cases} \frac{\partial U}{\partial b_0} = \sum_{i=1}^{n}[y_i - (b_0 + b_1 x_i)] = 0 \\ \frac{\partial U}{\partial b_1} = \sum_{i=1}^{n}[y_i - (b_0 + b_1 x_i)]x_i = 0 \end{cases} \tag{5.5}$$

Преобразуем полученную систему нормальных уравнений

$$\begin{cases} b_0 n + b_1 \sum x_i = \sum y_i \\ b_0 \sum x_i + b_1 \sum x_i^2 = \sum (y_i \cdot x_i) \end{cases} \tag{5.6}$$

В формуле (5.6) и далее для краткости у знака суммы (Σ) опущены индексы. Систему (5.6) решаем с помощью определителей

$$b_0 = \frac{\theta_1}{\theta};\ b_1 = \frac{\theta_2}{\theta}, \tag{5.7}$$

где θ – главный определитель.

Имеем

$$\theta = \begin{vmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{vmatrix} = n\sum x_i^2 - \left(\sum x_i\right)^2 \tag{5.8}$$

$$\theta_1 = \begin{vmatrix} \sum y_i & \sum x_i \\ \sum x_i\, y_i & \sum x_i^2 \end{vmatrix} = \sum y_i \cdot \sum x_i^2 - \sum x_i y_i \cdot \sum x_i \tag{5.9}$$

$$\theta_2 = \begin{vmatrix} n & \sum y_i \\ \sum x_i & \sum x_i y_i \end{vmatrix} = n\sum x_i y_i - \sum x_i \cdot \sum y_i \tag{5.10}$$

Откуда

$$b_0 = \frac{\sum y_i \cdot \sum x_i^2 - \sum x_i y_i \cdot \sum x_i}{n \cdot \sum x_i^2 - (\sum x_i)^2} \qquad (5.11)$$

$$b_1 = \frac{n \cdot \sum x_i y_i - \sum x_i \cdot \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2} \qquad (5.12)$$

Как и другие статистические расчеты, вычисление коэффициентов регрессии удобно проводить в табличной форме.

Пример. По данным табл. 5.2 построить уравнение линейной регрессии.

На примере построения линии регрессии по данным табл. 5.1 можно рассмотреть практическую методику вычисления коэффициентов регрессии, которая приведена в табл. 5.2.

Таблица 5.2

Методика вычисления коэффициентов регрессии

№ п/п	x	y	x^2	y^2	xy	$x + y$	$(x + y)^2$
1	1,5	5,0	2,25	25,00	7,50	6,50	42,25
2	4,0	4,5	16,00	20,25	18,00	8,50	72,25
3	5,0	7,0	25,00	49,00	35,00	12,00	144,00
4	7,0	6,5	49,00	42,25	45,50	13,50	182,25
5	8,5	9,5	72,25	90,25	80,75	18,00	324,00
6	10,0	9,0	100,00	81,00	90,00	19,00	361,00
7	11,0	11,0	121,00	121,00	121,00	22,00	484,00
8	12,5	9,0	156,25	81,00	112,50	21,50	462,25
Σ	59,5	61,5	541,75	509,75	510,25	121,00	2072,00

Средние значения: $\bar{x} = 7{,}4$; $\bar{y} = 7{,}7$.

Для проверки правильности вычислений в табл. 5.2 можно использовать выражение

$$\sum (x_i + y_i)^2 = \sum x_i^2 + 2\sum x_i y_i + \sum y_i^2. \qquad (5.13)$$

Значения сумм подставляем в эту формулу.

Получаем: 2072 = 541,75 + 2 · 510,25 + 509,75;

2072 = 2072.

Следовательно, вычисления выполнены правильно. В формулы (5.11) и (5.12) подставляем найденные значения для сумм из табл. 5.2, в результате получаем: $b_0 = 3{,}7263$; $b_1 = 36{,}5313$.

Уравнение регрессии или формула, которая отображает с некоторой вероятностью зависимость y от x, построенная по экспериментальным точкам, изображенным на рис. 5.1, имеет вид

$$y = 3{,}73 + 0{,}53x. \tag{5.14}$$

Это пример парной линейной зависимости.

Содержание отчета

В отчете изложить:

– теоретические предпосылки получения уравнения регрессии;

– результаты расчетов (исходные данные в табл. П5) для заполнения табл. 5.2;

– вычисленные средние значения x и y;

– результаты проверки по формуле 5.13;

– расчетные значения коэффициентов уравнения регрессии;

– сделать 5 выводов.

Контрольные вопросы

1. Что такое регрессия?
2. В чем состоит задача линейного регрессионного анализа?
3. Какое требование предъявляется для проведения вычислений по классическому методу наименьших квадратов?
4. В чем суть принципа Лежандра?
5. Как определяются коэффициенты уравнения регрессии?

Практическая работа № 6
МАТЕМАТИЧЕСКИЕ МЕТОДЫ ПЛАНИРОВАНИЯ ФАКТОРНОГО ЭКСПЕРИМЕНТА

Цель: изучить математические методы планирования факторного эксперимента.

Содержание работы

1. Изучить применение метода планирования эксперимента при исследовании технологических процессов.
2. Составить матрицу планирования для двухфакторного эксперимента.
3. Провести исследования на основании матрицы планирования эксперимента.
4. Определить уравнение регрессии проведенного исследования (составить модель).
5. Проверить однородность выборочных дисперсий по критерию Кохрена.
6. Проверить значимость коэффициентов уравнения регрессии по критерию Стьюдента.
7. Проверить адекватность полученной модели по критерию Фишера.

Постановка задачи

Для решения практических вопросов выполняют два типа исследований. Задачей первого типа исследований является выявление зависимости выходного параметра, т. е. функции отклика, от влияющих на него факторов (интерполяционная задача). Задача второго типа исследований состоит в определении таких значений факторов, при которых функция отклика принимает экстремальное значение – минимальное или максимальное (оптимизационная задача). В этом случае функция отклика называется параметром оптимизации. Достоинством применения математических методов планирования экспериментов является значительное сокращение количества опытов и времени на их реализацию при проведении исследований обоих видов.

По результатам предварительных исследований задаются моделью процесса в виде линейного уравнения регрессии

$$y = b_0 + \Sigma b_i x_i + \Sigma b_{ij} x_i x_j. \qquad (6.1)$$

Для определения коэффициентов уравнения регрессии составляется матрица планирования полного факторного эксперимента (ПФЭ) 2^k или дробного факторного эксперимента (ДФЭ) 2^{k-p}, по которым проводятся опыты. Матрица планирования и результаты измерений заносятся в табл. 6.1.

Таблица 6.1

Матрица планирования и результаты измерений

Номер опыта	X_0	X_1	X_2	X_K	Y	Y	S^2
1	+ 1	+1	–1	+1	$y_{11}, y_{12}, \dots y_{1m}$	y_1	S_1^2
2	+ 1	–1	–1	+1	$y_{21}, y_{22}, \dots y_{2m}$	y_2	S_2^2
3	+1	+1	+1	+1	$y_{31}, y_{32}, \dots y_{3m}$	y_3	S_3^2
…	…	…	…	…	…	…	…
N	+1	–1	+1	–1	$y_{N1}, y_{N2}, \dots y_{Nm}$	y_n	S_N^2

В каждой строчке матрицы планирования определяют среднее значение измеряемой величины по m параллельным опытам

$$\bar{y}_i = \frac{\sum y_i}{m}, \qquad (6.2)$$

и дисперсию

$$S_i^2 = \frac{\sum (y_i - \bar{y}_n)^2}{(m-1)}. \qquad (6.3)$$

Однородность выборочных дисперсий проверяется по критерию Кохрена. Для этого составляется отношение максимальной дисперсии к сумме всех дисперсий

$$G = \frac{S_{\max}^2}{\sum S_i^2}. \qquad (6.4)$$

Полученное отношение сравнивается с табличным (табл. 6.2) значением $G_{(1-p,\, f_1, f_2)}$, где $p = 0{,}05$; $f_1 = m - 1$; $f_2 = N$. Если $G < G_{(1-p,\, f_1, f_2)}$, то дисперсии однородны.

Таблица 6.2

Квантили распределения Кохрена при p = 0,05

f_2	f_1													
	1	2	3	4	5	6	7	8	9	10	16	36	144	200
2	9985	9750	9392	9057	8772	8534	8332	8159	8010	7880	7341	6602	5813	5000
3	9669	8709	7977	7457	7071	6771	6530	6333	6167	6025	5466	4748	4031	3333
4	9065	7679	6841	6287	5895	5598	5365	5175	5017	4884	4366	3720	3093	2500
5	8412	6838	5981	5441	5065	4783	4564	4387	4241	4118	3645	3066	2513	2000
6	7808	6161	5321	4803	4447	4184	3980	3817	3682	3568	3135	2612	2119	1667
7	7271	5612	4800	4307	3974	3726	3535	3384	3259	3154	2756	2278	1833	1429
8	6798	5157	4377	3910	3595	3362	3185	3043	2926	2829	2462	2022	1616	1250
9	6385	4775	4027	3584	3286	3067	2901	2768	2659	2568	2226	1820	1446	1111
10	6020	4450	3733	3311	3029	2823	2666	2541	2439	2353	2032	1655	1308	1000
12	5410	3924	3264	2880	2624	2439	2299	2187	2098	2020	1737	1403	1100	0833
15	4709	3346	2758	2419	2195	2034	1911	1815	1736	1671	1429	1144	0889	0667
20	3894	2705	2205	1921	1735	1602	1501	1422	1357	1303	1108	0879	0675	0500
24	3434	2354	1907	1656	1493	1374	1286	1216	1160	1113	0942	0743	0567	0417
30	2929	1980	1593	1377	1237	1137	1061	1002	0958	0921	0771	0604	0457	0333
40	2370	1576	1259	1082	0968	0887	0827	0780	0745	0713	0595	0462	0347	0250
60	1737	1131	0895	0765	0682	0623	0583	0552	0520	0497	0411	0316	0234	0167
120	0998	0632	0495	0419	0371	0337	0312	0292	0279	0266	0218	0165	0120	0083

В качестве оценки для дисперсии воспроизводимости можно взять среднюю дисперсию

$$S^2_{\text{воспр}} = \frac{\sum S_i^2}{N} \tag{6.5}$$

с числом степеней свободы $f_{\text{воспр}} = N(m-1)$.

Коэффициенты уравнения регрессии определяются по формуле

$$b_j = \frac{\sum x_{ij} y_i}{N}. \tag{6.6}$$

Необходимо отметить, что дисперсия y, полученного по выборке объема m, в m раз меньше дисперсии единичного измерения

$$S_y^2 = \frac{S^2_{\text{воспр}}}{m}. \tag{6.7}$$

Поэтому дисперсию коэффициентов S_{b_j} можно определить следующим образом:

$$S^2_{b_j} = \frac{S^2_{\text{воспр}}}{N \cdot m}. \tag{6.8}$$

Значимость коэффициентов проверяется по критерию Стьюдента (табл. 6.3).

Таблица 6.3

Процентные точки распределения Стьюдента

ν	40 %	25 %	10 %	5 %	2,5 %	1 %	0,5 %	0,25 %	0,1 %	0,05 %
1	0,3249	1,0000	3,0777	6,3138	12,7062	31,8205	63,6567	127,3213	318,3088	636,62
2	0,2887	0,8165	1,8856	2,9200	4,3027	6,9646	9,9248	14,0890	22,3271	31,599
3	0,2767	0,7649	1,6377	2,3534	3,1824	4,5407	5,8409	7,4533	10,2145	12,924
4	0,2707	0,7407	1,5332	2,1318	2,7764	3,7469	4,6041	5,5976	7,1732	8,6103
5	0,2672	0,7267	1,4759	2,0150	2,5706	3,3649	4,0321	4,7733	5,8934	6,8688
6	0,2648	0,7176	1,4398	1,9432	2,4469	3,1427	3,7074	4,3168	5,2076	5,9588
7	0,2632	0,7111	1,4149	1,8946	2,3646	2,9980	3,4995	4,0293	4,7853	5,4079
8	0,2619	0,7064	1,3968	1,8595	2,3060	2,8965	3,3554	3,8325	4,5008	5,0413
9	0,2610	0,7027	1,3830	1,8331	2,2622	2,8214	3,2498	3,6897	4,2968	4,7809
10	0,2602	0,6998	1,3722	1,8125	2,2281	2,7638	3,1693	3,5814	4,1437	4,5869
11	0,2596	0,6974	1,3634	1,7959	2,2010	2,7181	3,1058	3,4966	4,0247	4,4370
12	0,2590	0,6955	1,3562	1,7823	2,1788	2,6810	3,0545	3,4284	3,9296	4,3178
13	0,2586	0,6938	1,3502	1,7709	2,1604	2,6503	3,0123	3,3725	3,8520	4,2208
14	0,2582	0,6924	1,3450	1,7613	2,1448	2,6245	2,9768	3,3257	3,7874	4,1405
15	0,2679	0,6912	1,3406	1,7530	2,1314	2,6025	2,9467	3,2860	3,7328	4,0728
16	0,2576	0,6901	1,3368	1,7459	2,1199	2,5835	2,9208	3,2520	3,6862	4,0150
17	0,2573	0,6892	1,3334	1,7396	2,1098	2,5669	2,8982	3,2224	3,6458	3,9651
18	0,2571	0,6884	1,3304	1,7341	2,1009	2,5524	2,8784	3,1966	3,6105	3,9216
19	0,2569	0,6876	1,3277	1,7291	2,0930	2,5395	2,8609	3,1737	3,5794	3,8834
20	0,2567	0,6870	1,3253	1,7247	2,0860	2,5280	2,8453	3,1534	3,5518	3,8495

Окончание табл. 6.3

ν	40 %	25 %	10 %	5 %	2,5 %	1 %	0,5 %	0,25 %	0,1%	0,05 %
21	0,2566	0,6864	1,3232	1,7207	2,0796	2,5176	2,8314	3,1352	3,5272	3,8193
22	0,2564	0,6858	1,3212	1,7171	2,0739	2,5083	2,8188	3,1188	3,5050	3,7921
23	0,2563	0,6853	1,3195	1,7139	2,0687	2,4999	2,8703	3,1040	3,4850	3,7676
24	0,2562	0,6848	1,3178	1,7109	2,0639	2,4922	2,7969	3,0905	3,4668	3,7454
25	0,2561	0,6844	1,3163	1,7081	2,0595	2,4851	2,7874	3,0782	3,4502	3,7251
26	0,2560	0,6840	1,3150	1,7056	2,0555	2,4786	2,7787	3,0669	3,4350	3,7066
27	0,2559	0,6837	1,3137	1,7033	2,0518	2,4727	2,7707	3,0565	3,4210	3,6896
28	0,2558	0,6834	1,3125	1,7011	2,0484	2,4671	2,7633	3,0469	3,4082	3,6739
29	0,2557	0,6830	1,3114	1,6991	2,0452	2,4620	2,7564	3,0380	3,3962	3,6594
30	0,2556	0,6828	1,3104	1,6973	2,0423	2,4573	2,7500	3,0298	3,3852	3,6460
32	0,2555	0,6822	1,3086	1,6939	2,0369	2,4487	2,7385	3,0149	3,3653	3,6218
34	0,2553	0,6818	1,3070	1,6909	2,0322	2,4411	2,7284	3,0020	3,3479	3,6007
36	0,2552	0,6814	1,3055	1,6883	2,0281	2,4345	2,7195	2,9905	3,3326	3,5821
38	0,2551	0,6810	1,3042	1,6860	2,0244	2,4286	2,7116	2,9803	3,3190	3,5657
40	0,2550	0,6807	1,3031	1,6839	2,0211	2,4233	2,7045	2,9712	3,3069	3,5510
42	0,2550	0,6804	1,3020	1,6820	2,0181	2,4185	2,6981	2,9630	3,2960	3,5377
44	0,2549	0,6801	1,3011	1,6802	2,0154	2,4141	2,6923	2,9555	3,2861	3,5258
46	0,2548	0,6799	1,3002	1,6787	2,0129	2,4102	2,6870	2,9488	3,2771	3,5150
48	0,2548	0,6796	1,2994	1,677 2	2,0106	2,4066	2,6822	2,94 26	3,2689	3,5051
50	0,2547	0,6794	1,2987	1,6759	2,0086	2,4033	2,6778	2,9370	3,2614	3,4960
55	0,2546	0,6790	1,2971	1,6730	2,0040	2,3961	2,6682	2,9247	3,2561	3,4764
60	0,2545	0,67 86	1,2958	1,6706	2,0003	2,3901	2,6603	2,9146	3,2317	3,4602
65	0,2544	0,6783	1,2947	1,6686	1,9971	2,3851	2,6536	2,9060	3,2204	3,4466
70	0,2543	0,6780	1,2938	1,6669	1,9944	2,3808	2,6479	2,8987	3,2108	3,4350
80	0,2542	0,6776	1,2922	1,6641	1,9901	2,3739	2,6387	2,8870	3,1953	3,4163
90	0,2541	0,6772	1,2910	1,6620	1,9867	2,3685	2,6316	2,8779	3,1833	3,4019
100	0,2540	0,6770	1,2901	1,6602	1,9840	2,3642	2,6259	2,8707	3,1737	3,3905
120	0,2539	0,6765	1,2886	1,6577	1,9799	2,3578	2,6174	2,8599	3,1595	3,3735
150	0,2538	0,6761	1,2872	1,6551	1,9759	2,3515	2,6090	2,8492	3,1455	3,3566
200	0,2537	0,6757	1,2858	1,6525	1,9719	2,3451	2,6006	2,8385	3,1315	3,3398
250	0,2536	0,6755	1,2849	1,6510	1,9695	2,3414	2,5956	2,8322	3,1232	3,3299
300	0,2536	0,6753	1,2844	1,6499	1,9679	2,3388	2,5923	2,8279	3,1176	3,3233
400	0,2535	0,6751	1,2837	1,6487	1,9659	2,3357	2,5882	2,8227	3,1107	3,3150
500	0,2535	0,6750	1,2832	1,6479	1,9647	2,3338	2,5857	2,8195	3,1066	3,3101

Для всех коэффициентов уравнения регрессии составляют t-отношение

$$t_j = \frac{|b_j|}{S_{b_j}}, \tag{6.9}$$

которое сравнивают с табличным $t_{(1-p,\,f)}$ для уровня значимости $p = 0{,}05$ и числа степеней свободы $f = N(m-1)$. Если $t_j < t_{(1-p,\,f)}$, то принимается гипотеза равенства нулю генерального коэффициента регрессии, а соответствующий выборочный коэффициент b_j отсеивается как незначимый из уравнения регрессии. При этом остальные коэффициенты не пересчитывают.

Адекватность уравнения регрессии эксперименту проверяется по критерию Фишера. Для проверки адекватности составляют дисперсионное отношение

$$F = \frac{S^2_{\text{ад}}}{S^2_{\text{воспр}}}, \tag{6.10}$$

где $S^2_{\text{ад}}$ – дисперсия адекватности, которая определяется по формуле

$$S^2_{\text{ад}} = \frac{m\sum(\bar{y}_i - y_i)}{N-l}, \tag{6.11}$$

где l – число значимых коэффициентов в уравнении регрессии.

Если полученное дисперсионное отношение оказывается меньше табличного

$$F < F(1-p, v_1, v_2), \tag{6.12}$$

где p – уровень значимости; $v_1 = N - l$ – число степеней свободы дисперсии адекватности; $v_2 = N(m-1)$ – число степеней свободы дисперсии воспроизводимости, то уравнение считается адекватным эксперименту (табл. 6.4).

В противном случае для адекватного описания эксперимента необходимо увеличить порядок аппроксимирующего полинома.

Таблица 6.4

Процентные точки F-распределения Фишера, $p = 5$ %

ν_2	ν_1								
	1	2	3	4	5	6	7	8	9
1	161,45	199,50	215,71	224,58	230,16	233,99	236,77	238,88	140,54
2	18,513	19,000	19,164	19,247	19,296	19,330	19,353	19,371	19,385
3	10,128	9,5521	9,2766	9,1172	19,0135	8,9406	8,8868	8,8452	8,8123
4	7,7086	6,9443	6,5914	6,3883	6,2560	6,1631	6,0942	6,0410	5,9988
5	6,6079	5,7861	5,4095	5,1922	5,0503	4,9503	4,8759	4,8183	4,7725
6	5,9874	5,1433	4,7571	4,5337	4,3874	4,2839	4,2066	4,1468	4,0990
7	5,5914	4,7374	4,3468	4,1203	3,9715	3,8660	3,7870	3,7257	3,6767
8	5,3177	4,4590	4,0662	3,8378	3,6875	3,5806	3,5005	3,4381	3,3881
9	5,1174	4,2565	3,8626	3,6331	3,4817	3,3738	3,2927	3,2296	3,1789
10	4,9646	4,1028	3,7083	3,4780	3,3258	3,2172	3,1355	3,0717	3,0204
11	4,8443	3,9823	3,5874	3,3567	3,2039	2,0946	3,0123	2,9480	2,8962
12	4,7472	3,8853	3,4903	3,2592	3,1059	2,9961	2,9134	2,8486	2,7964
13	4,6672	3,8056	3,4105	3,1791	3,0254	2,9153	2,8321	2,7669	2,7144
14	4,6001	3,7389	3,3439	3,1122	2,9582	2,8477	2,7642	2,6987	2,6458
15	4,5431	3,6823	3,2974	3,0556	2,9013	2,7905	2,7066	2,6408	2,5876
16	4,4940	3,6337	3,2389	3,0069	2,8524	2,7413	2,6572	2,5911	2,5377
17	4,4513	3,5915	3,1968	2,9647	2,8100	2,6987	2,6143	2,5480	2,4943
18	4,4139	3,5546	3,1599	2,9277	2,7729	2,6613	2,5767	2,5102	2,4563
19	4,3808	3,5219	3,1274	2,8951	2,7401	2,6283	2,5435	2,4768	2,4227
20	4,3513	3,4928	3,0984	2,8661	2,7109	2,5990	2,5140	2,4471	2,3928
21	4,3248	3,4668	3,0725	2,8401	2,6848	2,5727	2,4876	2,4205	2,3661
22	4,3009	3,4434	3,0491	2,8167	2,6613	2,5491	2,4638	2,3965	2,3419
23	4,2793	3,4221	3,0280	2,7955	2,6400	2,5277	2,4422	2,3748	2,3201
24	4,2597	3,4028	3,0088	2,7763	2,6207	2,5082	2,4226	2,3551	2,3002
25	4,2417	3,3852	2,9912	2,7587	2,6030	2,4904	2,4047	2,3371	2,2821
26	4,2252	3,3690	3,9751	2,7426	2,5868	2,4741	2,3883	2,3205	2,2655
27	4,2100	3,3541	2,9604	2,7278	2,5719	2,4591	2,3732	2,3053	2,2501
28	4,1960	3,3404	2,9467	2,7141	2,5581	2,4453	2,3593	2,2913	2,2360
29	4,1830	3,3277	2,9340	2,7014	2,5454	2,4824	2,3463	2,2782	2,2229
30	4,1709	3,3158	2,9223	2,6896	2,5336	2,4205	2,3343	2,2662	2,2107
40	4,0848	3,2317	2,8387	2,6060	2,4495	2,3359	2,2490	2,1802	2,1240
60	4,0012	3,1504	2,7581	2,5252	2,3683	2,2540	2,1665	2,0970	2,0401
120	3,9201	3,0718	2,6802	2,4472	2,2900	2,1750	2,0867	2,0164	1,9588
160	3,8415	2,9957	2,6049	2,3719	2,2141	2,0986	2,0096	1,9384	1,8799

Окончание табл. 6.4

ν_2	ν_1									
	10	12	15	20	24	30	40	60	120	∞
1	241,88	243,91	245,95	248,01	249,05	250,09	251,14	252,20	253,25	254,32
2	19,396	19,413	19,429	19,446	19,454	19,462	19,471	19,479	19,487	19,496
3	8,7855	8,7446	8,7029	8,6602	8,6385	8,6166	8,5944	8,5720	8,5494	8,5265
4	5,9644	5,9117	5,8578	5,8025	5,7744	5,7459	5,7170	5,6870	5,6581	5,6281
5	4,7351	4,6777	4,6188	4,5581	4,5272	4,4957	4,4638	4,4314	4,3984	4,3650
6	4,0600	3,9999	3,9381	3,8742	3,8415	3,8082	3,7743	3,7398	3,7047	3,6688
7	3,6365	3,5747	3,5108	3,4445	3,4105	3,3758	3,3404	3,3043	3,2674	3,2298
8	3,3472	3,2840	3,2184	3,1503	3,1152	3,0794	3,0428	3,0053	2,9669	2,9276
9	3,1373	3,0729	3,0061	2,9365	2,9005	2,8637	2,8259	2,7872	2,7475	2,7067
10	2,9782	2,9130	2,8450	2,7740	2,7372	2,6996	2,6609	2,6211	2,5801	2,5379
11	2,8536	2,7876	2,7186	2,6464	2,6090	2,5705	2,5309	2,4901	2,4480	2,4045
12	2,7534	2,6866	2,6169	2,5436	2,5055	2,4663	2,4259	2,3842	2,3410	2,2962
13	2,6710	2,6037	2,5331	2,4589	2,4202	2,3803	2,3392	2,2966	2,2524	2,2064
14	2,6021	2,5342	2,4630	2,3879	2,3487	2,3082	2,2664	2,2230	2,1778	2,1307
15	2,5437	2,4753	2,4035	2,3275	2,2878	2,2468	2,2043	2,1601	2,1141	2,0658
16	2,4935	2,4247	2,3522	2,2756	2,2354	2,1938	2,1507	2,1058	2,0589	2,0096
17	2,4499	2,3807	2,3977	2,2304	2,1898	2,1477	2,1040	2,0584	2,0107	1,9604
18	2,4117	2,3421	2,2686	2,1906	2,1497	2,1071	2,0629	2,0166	1,9681	1,9168
19	2,3779	2,3080	2,2341	2,1555	2,1141	2,0712	2,0264	1,9796	1,9302	1,8780
20	2,3479	2,2776	2,2033	2,1242	2,0825	2,0391	1,9938	1,9464	1,8961	1,8432
21	2,3210	2,2504	2,1757	2,0960	2,0540	2,0102	1,9645	1,9165	1,8657	1,8117
22	2,2967	2,2258	2,1508	2,0707	2,0283	1,9842	1,9380	1,8895	1,8380	1,7831
23	2,2747	2,2036	2,1282	2,0476	2,0050	1,9605	1,9139	1,8649	1,8128	1,7570
24	2,2547	2,1834	2,1077	2,0267	1,9838	1,9390	1,8920	1,8424	1,7897	1,7331
25	2,2365	2,1649	2,0889	2,0075	1,9643	1,9192	1,8778	1,8217	1,7684	1,7110
26	2,2197	2,1479	2,0716	1,9898	1,9464	1,9010	1,8533	1,8027	1,7488	1,6906
27	2,2043	2,1323	2,0558	1,9736	1,9299	1,8842	1,8361	1,7851	1,7307	1,6717
28	2,1900	2,1179	2,0411	1,9586	1,9147	1,8687	1,8203	1,7689	1,7138	1,6541
29	2,1768	2,1045	2,0275	1,9446	1,9005	1,8543	1,8055	1,7537	1,6981	1,6377
30	2,1646	2,0921	2,0148	1,9317	1,8874	1,8409	1,7918	1,7396	1,6835	1,6223
40	2,0772	2,0035	1,9245	1,8389	1,7929	1,7444	1,6928	1,6373	1,5766	1,5089
60	1,9926	1,9174	1,8364	1,7480	1,7001	1,6491	1,5943	1,5343	1,4673	1,3893
120	1,9105	1,8337	1,7505	1,6587	1,6084	1,5543	1,4952	1,4290	1,3519	1,2539
160	1,8307	1,7522	1,6664	1,5705	1,5173	1,4591	1,3940	1,3180	1,2214	1,0000

Пример. В качестве примера можно привести исследование износа комплектов автомобильных тормозных колодок разных производителей.

Тормозные колодки четыре разных производителей (A, B, C, D) установлены на четырех автомобилях одной модели. Каждым автомобилем управлял закрепленный за ним водитель.

Поскольку износ зависит от состояния автомобиля и особенностей стиля вождения водителя, то при испытаниях каждый автомобиль прошел одинаковый пробег с комплектами тормозных колодок каждого из производителей.

Возникает вопрос – удалось ли за счет этого исключить влияние фактора «автомобиль» на характеристику износа тормозных колодок? Для ответа на этот вопрос необходимо провести дисперсионный анализ, который позволит определить значимость всех факторов.

После пробега 10 тыс. км износ тормозных колодок был замерен и усредненные по комплектам результаты занесены в табл. 6.5.

Общее число измерений $n = p \cdot q = 4 \cdot 4 = 16$,
где p – число уровней факторов *производитель тормозных колодок*; q – число уровней фактора *автомобиль*.

Таблица 6.5

Износ тормозных колодок, мм

Производитель тормозных колодок	Автомобиль				Сумма по строке	$\bar{x}_i$
	I	II	III	IV		
A	11,2	8,4	8,4	7,1	35,1	8,78
B	9,7	7,6	7,1	5,5	29,9	7,48
C	5,8	6,3	5,6	4,3	22,0	5,50
D	7,1	5,8	6,7	4,7	24,3	6,08
Σ	33,8	28,1	27,8	21,6	–	–
Среднее по столбцам	8,45	7,03	6,95	5,4	–	6,96

Общее среднее износа тормозных колодок

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{q} x_{ij} = \frac{111{,}3}{16} = 6{,}96 \text{ мм.}$$

Сумма квадратов отклонений всех наблюдаемых значений от их общего среднего значения равна

$$Q = \sum_{i=1}^{p}\sum_{j=1}^{q}\left(x_{ij} - \bar{x}\right)^2 = (11{,}2 - 6{,}96)^2 + (8{,}4 - 6{,}96)^2 +$$

$$+(8{,}4 - 6{,}96)^2 + (7{,}1 - 6{,}96)^2 + (9{,}7 - 6{,}96)^2 + (7{,}6 - 6{,}96)^2 +$$
$$+(7{,}1 - 6{,}96)^2 + (5{,}5 - 6{,}96)^2 + (5{,}8 - 6{,}96)^2 + (6{,}3 - 6{,}96)^2 +$$
$$+(5{,}6 - 6{,}96)^2 + (4{,}3 - 6{,}96)^2 + (7{,}1 - 6{,}96)^2 + (5{,}8 - 6{,}96)^2 +$$
$$+(6{,}7 - 6{,}96)^2 + (4{,}7 - 6{,}96)^2 = 49{,}46 \text{ мм}^2.$$

Сумма квадратов отклонений средних значений по группе «производители тормозных колодок»

$$Q_{\text{тк}} = \sum q \sum_{i=1}^{4}(\bar{x}_i - \bar{x})^2 = 4[(8{,}78 - 6{,}96)^2 + (7{,}48 - 6{,}96)^2 +$$
$$+(5{,}5 - 6{,}96)^2 + (6{,}08 - 6{,}96)^2] = 4 \cdot 6{,}47 = 25{,}9 \text{ мм}^2.$$

Сумма квадратов отклонений средних значений по группе «автомобили»

$$Q_{\text{авт}} = \sum p \sum_{j=1}^{4}\left(\bar{x}_j - \bar{x}\right)^2 = 4[(8{,}45 - 6{,}96)^2 + (7{,}03 - 6{,}96)^2 +$$
$$+(6{,}96 - 6{,}96)^2 + (5{,}4 - 6{,}96)^2] = 4 \cdot 4{,}66 = 18{,}63 \text{ мм}^2.$$

Остаточная сумма квадратов

$Q_R = Q - Q_{\text{тк}} - Q_{\text{авт}} = 49{,}46 - 25{,}9 - 18{,}63 = 4{,}93$ мм^2.

Числа степеней свободы:

$K = n - 1 = 16 - 1 = 15$ (для суммы Q);

$K_{\text{тк}} = q - 1 = 4 - 1 = 3$ (для суммы $Q_{\text{ш}}$);

$K_{\text{авт}} = P - 1 = 4 - 1 = 3$ (для суммы $Q_{\text{авт}}$).

Остаточная сумма степеней свободы (для остаточной суммы Q_R)

$K_R = (p - 1)(q - 1) = 3 \cdot 3 = 9.$

Несмещенные оценки дисперсий:

$$D = \frac{Q}{n - 1} = \frac{49{,}46}{15} = 3{,}3 \text{ мм}^2;$$

$$D_{\text{тк}} = \frac{Q_{\text{тк}}}{q - 1} = \frac{25{,}9}{3} = 8{,}63 \text{ мм}^2;$$

$$D_{\text{авт}} = \frac{Q_{\text{авт}}}{p - 1} = \frac{18{,}63}{3} = 6{,}21 \text{ мм}^2;$$

$$D_R = \frac{Q_R}{(p - 1)(q - 1)} = \frac{4{,}93}{9} = 0{,}55 \text{ мм}^2.$$

Значения F-критериев для соответствующих значений дисперсий окажутся равными

$$F_{\text{тк}} = \frac{D_{\text{тк}}}{D_R} = \frac{8{,}63}{0{,}55} = 15{,}76,$$

с табличными (для 5%-го уровня риска) значениями числа степеней свободы $K_{\text{тк}} = 3$ и $K_R = 9$, значение $F_{\text{кр}} = 3{,}86$.

Поскольку $F_{\text{тк}} > F_{\text{кр}}$, то влияние фактора «производители тормозных колодок» на износ следует признать существенным.

Значение F-критерия для определения значимости влияния «автомобили» будет равен

$$F_{\text{авт}} = \frac{D_{\text{авт}}}{D_R} = \frac{6{,}21}{0{,}55} = 11{,}34,$$

при критическом значении $F_{\text{кр}} = 3{,}86$ (для уровня риска 5 %).

Таким образом, $F_{\text{авт}} > F_{\text{кр}}$, значит, влияние фактора «автомобили» на износ шин существенно.

Необходимо заметить, что существенность влияния автомобилей на износ тормозных колодок указывает на несовершенство планирования опыта. В плане эксперимента можно предложить некоторые меры снижения такой зависимости.

Каждый комплект тормозных колодок испытан на своем автомобиле при одинаковом пробеге (10 тыс. км). Для снижения зависимости износа от фактора «автомобили» можно ввести требование перестановки комплектов тормозных колодок таким образом, чтобы они прошли на каждом из четырех автомобилей пробег в 2,5 тыс. км.

Содержание отчета

В отчете изложить:

– задачи, решаемые с использованием методики планирования эксперимента при исследовании технологических процессов;

– общую методику планирования эксперимента (факторы, влияющие на результаты исследуемого процесса, его оценки, уравнение регрессии, матрица планирования и результаты измерений;

– методику проверки однородности выборочных дисперсий по критерию Кохрена;

– методику проверки значимости коэффициентов уравнения регрессии по критерию Стьюдента;

– методику проверки адекватности уравнения регрессии эксперименту по критерию Фишера, произвести дисперсионный анализ (исходные данные в табл. П6) аналогично примеру 6.1;

– сделать 5 выводов.

Контрольные вопросы

1. Какие задачи решают с использованием методики планирования эксперимента?
2. В чём сущность методики планирования эксперимента?
3. Что проверяют при проведении экспериментальных исследований по критерию Кохрена?
4. Что проверяют при проведении экспериментальных исследований по критерию Стьюдента?
5. Что проверяют при проведении экспериментальных исследований по критерию Фишера?

Практическая работа № 7
МЕТОДЫ ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ

Цель: познакомиться с методикой планирования факторного эксперимента.

Содержание работы

1. Изучить теоретическую часть работы.
2. Составить конспект отчета об изученном материале.

1. Планирование эксперимента

Требования к планированию эксперимента:

1) число опытов должно быть минимальным, чтобы не усложнять процедуру эксперимента и не увеличивать его стоимость, но не в ущерб точности результата;
2) необходимо определить совокупность факторов, влияющих на результаты эксперимента, ранжировать их, выявить главные, несущественные переменные можно исключить;
3) условием корректности эксперимента следует считать одновременное варьирование всеми переменными (факторами), оказывающими взаимное влияние на исследуемый процесс;
4) ряд действий в эксперименте может быть заменен их моделями (прежде всего математическими), при этом адекватность моделей должна быть проверена и оценена;
5) необходимо разработать стратегию эксперимента и алгоритм ее реализации: серии эксперимента должны анализироваться после завершения каждой из них перед переходом к последующей серии.

План проведения эксперимента должен включать следующие разделы:

1. Наименование темы исследования.
2. Цель и задачи эксперимента.

3. Условия проведения эксперимента: параметр оптимизации и варьируемые факторы.

4. Методика проведения исследования.

5. Обоснование количества опытов (объема эксперимента).

6. Средства и методика проведения измерений.

7. Материальное обеспечение эксперимента (перечень оборудования).

8. Методика обработки и анализа экспериментальных данных.

9. Календарный план проведения испытаний, в котором указываются сроки их выполнения, исполнители, представляемые данные эксперимента.

10. Смета расходов.

Цель и задачи эксперимента – исходный пункт плана. Они формулируются на основе анализа научной гипотезы, теоретических результатов собственного исследования либо исследований других авторов.

Цель определяет конечный результат эксперимента, то есть то, что исследователь должен получить в итоге. Например, подтвердить правильные научные гипотезы; проверить на практике адекватность, работоспособность и практическую пригодность моделей, методик; определить оптимальные условия технологического процесса и т. п.

В различных условиях цели требуют разных затрат, средств и методов измерения, времени эксперимента, отражаются на методике его проведения. Эти пункты плана будут различными, например, в условиях лабораторного, полевого и производственного экспериментов.

Задачи эксперимента определяют частные цели, с помощью которых может быть достигнута конечная цель либо пути ее достижения. Например, определение оптимальных показателей температуры и давления при изготовлении фулиреновых нанотрубок; установление оптимального соотношения исходных материалов; обоснование скорости протекания технологического процесса и др.

Частными задачами эксперимента при его планировании могут быть:

– проверка теоретических положений с целью подтверждения их истинности;

– проверка (уточнение) констант математических либо иных моделей;

– поиск оптимальных (допустимых) условий какого-либо процесса;

– построение интерполяционных аналитических зависимостей.

Частные задачи эксперимента могут иметь несколько уровней, то есть древовидную форму. Рекомендуется формулировать до четырех сложных задач и десять или более простых задач.

Величина, описывающая результат проведенного эксперимента, называется параметром оптимизации (откликом) системы на воздействие. Множество значений, которые принимает параметр оптимизации, называется областью его определения.

Параметр оптимизации должен быть количественным, задаваться числом и быть измеримым при любом фиксированном наборе уровней факторов. Он обязан характеризоваться однозначно – заданному набору уровней факторов должно соответствовать, с точностью ошибки эксперимента, одно значение параметра оптимизации. Параметр оптимизации должен всесторонне характеризовать объект исследования, удовлетворять требованию универсальности и полноты. Он должен иметь физический смысл, чтобы обеспечить последующую интерпретацию результатов эксперимента, быть простым и легко вычисляемым.

Параметр оптимизации (отклик) зависит от факторов, влияющих на эксперимент. Фактор (лат. *factor* – производящий) – причина какого-либо процесса, явления, определяющая его влияние на объект исследования, его характер или отдельные черты. Это измеряемая величина, и каждое значение, которое может принимать фактор, называется уровнем фактора.

Каждый фактор в эксперименте может принимать одно из нескольких значений. Фиксированный набор уровней нескольких факторов будет определять какие-то конкретные условия проведения эксперимента. Изменение хотя бы одного из факторов приводит к изменению условий и как следствие к изменению значения параметра оптимизации.

Варьируемые факторы в многофакторном эксперименте определяют цели и условия исследования. Например, факторами в эксперименте по поиску оптимальных условий при производстве наноматериалов могут быть: температура, время воздействия, количество окисла и т. п.

Большое количество факторов делает эксперимент очень сложным и требует довольно много времени. Поэтому весьма важны при планировании эксперимента сокращение числа факторов и выбор наиболее существенных. При этом можно руководствоваться принципом Парето, в соответствии с которым 20 % факторов определяют 80 % свойств системы.

Значимость факторов может быть определена опытным или аналитическим путем. В первом случае проводится ограниченный эксперимент. При этом один фактор изменяется, а остальные нет и т. д. Ранжирование значимых факторов осуществляется по величине их влияния на результат эксперимента. Те факторы, изменение которых больше отражается на конечном результате, считаются более важными. Несущественными факторами можно пренебречь.

Если факторов много, этот путь неэффективен, тогда используется аналитический путь, основанный на методах факторного анализа.

Если факторы зависимы, их можно рассчитать с помощью метода топологической декомпозиции и структуры по их влиянию на конечную цель. Задача определения рангов факторов заключается в выделении наиболее связной части графа. Она решается поэтапно.

Сначала определяются «достижимые множества» для каждой вершины графа (для каждого фактора). Затем определяются «контрдостижимые множества», каждое из которых включает все вершины, имеющие путь в вершину. В завершении определяют наиболее существенные вершины графа, составляющие сильно связанный граф. Самые существенные факторы оставляют, остальные отбрасывают.

Важнейшее требование эксперимента – управляемость факторов, а экспериментатор должен иметь возможность выбрать нужное значение фактора и поддерживать его постоянным на протяжении всего эксперимента. Фактор также должен быть операциональным, чтобы его можно было указать последовательностью операций, необходимых для задания того или иного значения.

Формализуя условия проведения эксперимента, важно также определиться с областью его проведения. Для этого необходимо оценить границы областей определения факторов. Здесь возможны ограничения нескольких типов: которые не могут быть нарушены ни при каких условиях (например, температура не может оказаться ниже абсолютного нуля); технико-экономические ограничения (например, стоимость оборудования или продолжительность исследуемого процесса); конкретные условия процесса.

Под моделью эксперимента обычно понимают модель черного ящика, в которой используется функция отклика, устанавливающая зависимость между параметром оптимизации и факторами

$$y = f(x_1, x_2, \dots, x_n).$$

Выбрать модель – значит выбрать вид этой функции и записать ее уравнение. Тогда останется только провести эксперимент по вычислению численных коэффициентов данной модели. Главное требование к модели эксперимента – способность предсказывать дальнейшее направление опытов с требуемой точностью. Среди всех возможных адекватных моделей необходимо выбирать ту, которая представляется наиболее простой.

Наиболее часто в планировании эксперимента выбирают полиномиальные модели первой (линейной) или второй степени

$$\begin{gathered} y = b_0 + b_1 x_1 + b_2 x_2, \\ y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_1 + b_{11} x_1^2 + b_{22} x_2^2. \end{gathered} \tag{7.1}$$

Методика проведения эксперимента – ключевая часть плана эксперимента. Она включает:

– последовательность действий исследователя;

– основные приемы и правила осуществления каждого этапа, использование приборов и оборудования;

– порядок измерения, фиксации результатов и методы их обработки;

– порядок анализа результатов эксперимента и формулирования выводов.

При разработке методики важно правильно обосновать количество опытов, которое гарантирует требуемую точность результата, а с другой стороны – не ведет к неоправданному перерасходу средств и времени на избыточные испытания.

При более чем десяти испытаниях обоснование количества опытов может быть осуществлено на основе неравенства Чебышева

$$P[|\bar{X}| - M(x) \leq \varepsilon] \leq 1 - \frac{D(x)}{N\varepsilon^2}, \qquad (7.2)$$

где $\bar{X}$ – среднее значение случайно измеряемой величины; $M(x)$ – математическое ожидание величины; ε – требуемая точность результата; $D(x)$ – дисперсия величины x, рассчитанная по результатам N проведенных опытов.

Неравенство можно сформулировать следующим образом: вероятность того, что разность между математическим ожиданием и среднестатистическим значением случайной величины X не превысит требуемую точность результата (ε), равна разности между единицей и отношением $\frac{D(x)}{N\varepsilon^2}$.

В неравенстве три неизвестных: N и статистические характеристики, зависящие от N. Поэтому процесс расчета N проводится методом проб и ошибок.

Если неравенство выполняется, то количество опытов достаточно. В противном случае количество опытов увеличивается.

Достаточное количество наблюдений (опытов) может быть определено при помощи таблицы достаточно больших чисел (табл. 7.1). Она показывает, что достаточное количество наблюдений зависит от степени уверенности в результатах эксперимента (доверительной вероятности) и величины допустимой ошибки (доверительного интервала). Иными словами, степень уверенности определяется величиной вероятности, с которой делается соответствующее заключение.

Таблица 7.1

Таблица достаточно больших чисел

P	ε									
	0,10	0,09	0,08	0,07	0,06	0,05	0,04	0,03	0,02	0,01
0,75	33	40	51	67	91	132	206	367	827	3308
0,80	41	50	64	83	114	164	256	456	1026	4105
0,85	51	63	80	105	143	207	323	575	1295	5180
0,90	67	83	105	138	187	270	422	751	1690	6763
0,91	71	88	112	146	199	287	449	798	1796	7185
0,92	76	94	119	156	212	306	478	851	1915	7662
0,93	82	101	128	167	227	328	512	911	2051	8207
0,94	88	109	138	180	245	353	552	982	2210	8843
0,95	96	118	150	195	266	384	600	1067	2400	9603
0,96	105	130	164	215	292	421	659	1171	2636	10544
0,965	111	137	173	226	308	444	694	1234	2778	11112
0,970	117	145	183	240	327	470	735	1308	2943	11773
0,975	125	155	196	256	348	502	784	1395	3139	12559
0,980	135	167	211	276	375	541	845	1503	3382	13529
0,985	147	182	231	301	410	591	924	1643	3697	14791
0,990	165	204	259	338	460	663	1036	1843	4146	16587
0,991	170	210	266	348	473	682	1066	1895	4264	17057
0,992	175	217	274	358	488	703	1098	1953	4395	17583
0,993	181	224	284	371	505	727	1136	2020	4545	18182
0,994	188	233	294	385	524	755	1179	2097	4718	18875
0,995	196	243	307	402	547	787	1231	2188	4924	19698
0,996	207	255	323	422	575	828	1294	2301	5177	20409
0,997	220	271	344	449	611	880	1376	2446	5504	22018
0,998	238	294	373	487	663	954	1492	2652	5968	23873
0,999	270	334	422	552	751	1082	1691	3007	6767	27069

Относительно выбора величины вероятности *P* нет какого-либо общего решения, одинакового при всех исследованиях. Чем ближе к единице будет величина рассматриваемой вероятности, тем надежнее будет заключение. В практике научных исследований доверительная вероятность обычно принимается $P = 0{,}9 - 0{,}99$. Требуемая точность при исследованиях устанавливается в зависимости от природы изучаемого явления. В большинстве случаев требуемая точность принимается равной $\varepsilon = 0{,}01 - 0{,}05$.

Например, если величина доверительной вероятности принята равной $P = 0{,}95$, а допустимая ошибка равна $\varepsilon = 0{,}05$, то достаточное число наблюдений в ходе эксперимента будет равно 384.

Другая важная составляющая плана эксперимента –это обоснование средств и методики измерений, что предполагает выбор измерительных приборов, аппаратуры и оборудования, позволяет фиксировать данные эксперимента; преобразовывать их к удобному виду; хранить, обеспечивать выдачу по запросам и т. п.

Система измерений должна формироваться с учетом требований метрологии – науки о методах и средствах измерений, выборе единиц, шкал и систем измерений; проблемах точности измерений. Методы измерений, которые могут быть применены в различных экспериментах, рассмотрены в предыдущей практической работе.

Эти методы измерения могут быть сведены в две группы: прямых (искомая величина измеряется непосредственно в ходе эксперимента) и косвенных измерений (искомая величина, полученная на основе результатов прямых измерений). Кроме того, по признаку единиц измерений различают абсолютные измерения, проводимые в единицах исследуемой величины, и относительные измерения, предполагающие фиксацию отношения измеряемой величины к ее некоторому предельному значению.

Рассмотренные основы организации и проведения эксперимента носят лишь обзорный характер, а сущность, содержание, условия применения вышеизложенных рекомендаций и последовательность использования того или иного метода проведения эксперимента требуют более детального изучения. Кроме того, следует четко понимать, что каждый метод проведения эксперимента будет иметь и свои особенности в зависимости от объекта исследования.

2. Основные понятия планирования эксперимента

Планирование эксперимента имеет свою определенную терминологию. Рассмотрим общие термины.

Эксперимент – система операций, воздействий и (или) наблюдений, направленных на получение информации об объекте при исследовательских испытаниях.

Опыт – воспроизведение исследуемого явления в определенных условиях проведения эксперимента при возможности регистрации его результатов. Опыт представляет собой отдельную элементарную часть эксперимента.

Планирование эксперимента – процедура выбора числа опытов и условий их проведения, необходимых для решения поставленной задачи с требуемой точностью. Все факторы, определяющие процесс, изменяются одновременно по специальным правилам, а результаты эксперимента представляются в виде математической модели.

Задачи, для решения которых может использоваться планирование эксперимента, чрезвычайно разнообразны. К ним относятся: поиск оптимальных условий, построение интерполяционных формул, выбор существенных факторов, оценка и уточнение констант теоретических моделей, выбор наиболее приемлемых из некоторого множества гипотез о механизме явлений, исследование диаграмм состав – свойство и т. д.

Поиск оптимальных условий является одной из наиболее распространенных научно-технических задач. Они возникают в тот момент, когда установлена возможность проведения процесса и необходимо найти наилучшие (оптимальные) условия его реализации. Такие задачи называются *задачами оптимизации.* Процесс их решения называется *процессом оптимизации* или просто оптимизацией. Примеры задачи оптимизации: выбор оптимального состава многокомпонентных смесей и сплавов, повышение производительности действующих установок, повышение качества продукции, снижение затрат на ее получение и т. п.

Выделяют следующие этапы построения математической модели:

1) сбор и анализ априорной информации;

2) выбор факторов и выходных переменных, области экспериментирования;

3) выбор математической модели, с помощью которой будут представляться экспериментальные данные;

4) выбор критерия оптимальности и плана эксперимента;

5) определение метода анализа данных;

6) проведение эксперимента;

7) проверка статистических предпосылок для полученных экспериментальных данных;

8) обработка результатов;

9) интерпретация и рекомендации.

Факторы определяют состояние объекта. Основное требование к факторам – управляемость. Под управляемостью понимается установление нужного значения фактора (уровня) и поддержание его в течение всего опыта. В этом состоит особенность активного эксперимента. Факторы могут быть количественными и качественными. Примеры количественных факторов: температура, давление, концентрация и т. п. Их уровням соответствует числовая шкала. Различные катализаторы, конструкции аппаратов, способы лечения, методики преподавания являются примерами качественных факторов. Уровням таких факторов не соответствует числовая шкала, и их порядок не играет роли.

Выходные переменные – это реакции (отклики) на воздействие факторов. Отклик зависит от специфики исследования и может быть экономическим (прибыль, рентабельность), технологическим (выход, надежность), психологическим, статистическим и т. д. Параметр оптимизации должен быть эффективным с точки зрения достижения цели, универсальным, количественным, выражаемым числом, имеющим физический смысл, быть простым и легко вычисляемым.

Затраты машинного времени можно значительно сократить, если на этапе оптимизации параметров использовать экспериментальную факторную математическую модель. Экспериментальные факторные модели, в отличие от теоретических, не используют физических законов, описывающих происходящие в объектах процессы, а представляют собой некоторые формальные зависимости выходных параметров от внутренних и внешних параметров объектов проектирования.

Экспериментальная факторная модель может быть построена на основе проведения экспериментов непосредственно на самом техническом объекте (физические эксперименты) либо на основе вычислительных экспериментов на ЭВМ с теоретической моделью.

При построении экспериментальной факторной модели (рис. 7.1) объект моделирования (проектируемая техническая система) представляется в виде «черного ящика», на вход которого подаются некоторые переменные X и Z, а на выходе можно наблюдать и регистрировать переменные Y.

Рис. 7.1. Экспериментальная факторная модель

В процессе проведения эксперимента изменение переменных X и Z приводит к изменениям выходных переменных Y. Для построения факторной модели необходимо регистрировать эти изменения и осуществить необходимую статистическую обработку для определения параметров модели.

При проведении физического эксперимента переменными X можно управлять, изменяя их величину по заданному закону. Переменными Z управлять нельзя, они принимают случайные значения. При этом значения переменных X и Z можно регистрировать с помощью соответствующих средств измерения. Кроме того, на объект воздействуют некоторые переменные E, которые нельзя наблюдать и контролировать. Переменные X называют контролируемыми управляемыми; переменные Z – контролируемыми, но неуправляемыми, а переменные E – неконтролируемыми и неуправляемыми.

Переменные X и Z называют факторами. Факторы X являются управляемыми и изменяются как детерминированные переменные, а факторы Z неуправляемые, изменяемые во времени случайным образом, то есть Z представляют собой случайные процессы. Пространство контролируемых переменных – факторов X и Z – образует факторное пространство.

Выходная переменная Y представляет собой вектор зависимых переменных моделируемого объекта. Ее называют откликом, а зависимость Y от факторов X и Z – функцией отклика. Геометрическое представление функции отклика называют поверхностью отклика.

Переменная Е действует в процессе эксперимента бесконтрольно. Если предположить, что факторы X и Z стабилизированы во времени и сохраняют постоянные значения, то под влиянием переменных E функция отклика Y может меняться как систематическим, так и случайным образом. В первом случае говорят о систематической помехе, а во втором – о случайной помехе. При этом полагают, что случайная помеха обладает вероятностными свойствами, не изменяемыми во времени.

Возникновение помех обусловлено ошибками методик проведения физических экспериментов, ошибками измерительных приборов, неконтролируемыми изменениями параметров и характеристик объекта и внешней среды.

В вычислительных экспериментах объектом исследования является теоретическая математическая модель, на основе которой необходимо получить экспериментальную факторную модель. Для ее получения необходимо определить структуру и численные значения параметров модели.

Под структурой модели понимается вид математических соотношений между факторами X, Z и откликом Y. Параметры представляют собой коэффициенты уравнений факторной модели. Структуру модели обычно выбирают на основе априорной информации об объекте с учетом назначения и последующего использования модели. Задача определения параметров модели полностью формализована. Она решается методами регрессионного анализа. Экспериментальные факторные модели называют также регрессионными моделями.

Регрессионную модель можно представить выражением

$$\vec{Y} = \left(\vec{X}, \vec{Z}, \vec{B}\right), \tag{7.3}$$

где $\vec{B}$ – вектор параметров факторной модели.

Вид вектор-функции φ определяется выбранной структурой модели и считается заданным, а параметры B подлежат определению на основе результатов эксперимента.

Различают эксперименты пассивные и активные.

Пассивным называется такой эксперимент, когда значениями факторов управлять нельзя и они принимают случайные значения. В таком эксперименте существуют только факторы Z. В процессе эксперимента в определенные моменты времени измеряются значения факторов Z и функций откликов Y. После проведения N опытов полученная информация обрабатывается статистическими методами, позволяющими определить параметры факторной модели. Такой подход к построению математической модели лежит в основе метода статистических испытаний (Монте-Карло).

Активным называется такой эксперимент, когда значениями факторов задаются и поддерживают их неизменными в заданных уровнях в каждом опыте в соответствии с планом эксперимента. Следовательно, в этом случае существуют только управляемые факторы *X*.

Основные особенности экспериментальных факторных моделей следующие: они статистические; представляют собой сравнительно простые функциональные зависимости оценок математических ожиданий выходных параметров объекта от ее внутренних и внешних параметров; дают адекватное описание установленных зависимостей лишь в области факторного пространства, в которой реализован эксперимент. Статистически регрессионная модель описывает поведение объекта в среднем, характеризуя его неслучайные свойства, которые в полной мере проявляются лишь при многократном повторении опытов в неизменных условиях.

3. Основные принципы планирования эксперимента

Для получения адекватной математической модели необходимо обеспечить выполнение определенных условий проведения эксперимента. Модель называют адекватной, если в оговоренной области варьирования факторов *X* полученные с помощью модели значения функций отклика *Y* отличаются от истинных не более чем на заданную величину. Методы построения экспериментальных факторных моделей рассматриваются в теории планирования эксперимента.

Цель планирования эксперимента: получение максимума информации о свойствах исследуемого объекта при минимуме опытов.

Такой подход обусловлен высокой стоимостью экспериментов, как физических, так и вычислительных, и вместе с тем необходимостью построения адекватной модели.

При планировании активных экспериментов используются следующие принципы:

– отказ от полного перебора всех возможных состояний объекта;

– постепенное усложнение структуры математической модели;

– сопоставление результатов эксперимента с величиной случайных помех;

– рандомизация опытов;

– оптимальное планирование эксперимента.

Детальное представление о свойствах поверхности отклика может быть получено лишь при условии использования густой дискретной сетки значений факторов, покрывающей все факторное пространство. В узлах этой многомерной сетки находятся точки плана, в которых проводятся опыты. Выбор структуры факторной модели основан на постулировании определенной степени гладкости поверхности отклика. Поэтому с целью уменьшения количества опытов принимают небольшое число точек плана, для которых осуществляется реализация эксперимента.

При большом уровне случайной помехи получается большой разброс значений функции отклика Y в опытах, проведенных в одной и той же точке плана. В этом случае оказывается, что чем выше уровень помехи, тем с большей вероятностью простая модель окажется работоспособной. Чем меньше уровень помехи, тем точнее должна быть факторная модель.

Кроме случайной помехи при проведении эксперимента может иметь место систематическая помеха. Наличие этой помехи практически никак не обнаруживается, и результат ее воздействия на функцию не поддается контролю. Однако если путем соответствующей организации проведения опытов искусственно создать случайную ситуацию, то систематическую помеху можно перевести в разряд случайных. Такой принцип организации эксперимента называют рандомизацией систематически действующих помех.

Наличие помех приводит к ошибкам эксперимента. Ошибки подразделяют на систематические и случайные соответственно наименованиям вызывающих их факторов – помех.

Рандомизацию опытов осуществляют только в физических экспериментах. Следует отметить, что в этих экспериментах систематическую ошибку может порождать наряду с отмеченными ранее факторами также неточное задание значений управляемых факторов, обусловленное некачественной калибровкой приборов для их измерения (инструментальная ошибка), конструктивными или технологическими факторами.

К факторам в активном эксперименте предъявляются определенные требования. Они должны быть:

– управляемыми (установка заданных значений и поддержание их постоянными в процессе опыта);

– совместными (их взаимное влияние не должно нарушать процесс функционирования объекта);

– независимыми (уровень любого фактора должен устанавливаться независимо от уровней остальных);

– однозначными (одни факторы не должны быть функцией других);

– непосредственно влияющими на выходные параметры.

Выбор параметров оптимизации (критериев оптимизации) является одним из главных этапов работы на стадии предварительного изучения объекта исследования, так как правильная постановка задачи зависит от правильности выбора параметра оптимизации, являющегося функцией цели.

Под параметром оптимизации понимают характеристику цели, заданную количественно. Параметр оптимизации является реакцией (откликом) на воздействие факторов, которые определяют поведение выбранной системы.

Реальные объекты или процессы, как правило, очень сложны. Они часто требуют одновременного учета нескольких, иногда очень многих параметров. Каждый объект может характеризоваться всей совокупностью параметров, или любым подмножеством этой совокупности, или одним-единственным параметром оптимизации. В последнем случае прочие характеристики процесса уже не выступают в качестве параметра оптимизации, а служат ограничениями. Другой путь – построение обобщенного параметра оптимизации как некоторой функции от множества исходных.

Параметр оптимизации (функции отклика) – это признак, по которому оптимизируется процесс. Он должен быть количественным, задаваться числом. Множество значений, которые может принимать параметр оптимизации, называется областью его определения. Области определения могут быть непрерывными и дискретными, ограниченными и неограниченными. Например, выход реакции – это параметр оптимизации с непрерывной ограниченной областью определения. Он может изменяться в интервале от 0 до 100 %. Число бракованных изделий, число зерен на шлифе сплава, число кровяных телец в пробе крови – вот примеры параметров с дискретной областью определения, ограниченной снизу.

Количественная оценка параметра оптимизации на практике не всегда возможна. В таких случаях пользуются приемом, называемым ранжированием. При этом параметрам оптимизации присваиваются оценки – ранги по заранее выбранной шкале: двухбалльной, пятибалльной и т. д. Ранговый параметр имеет дискретную ограниченную область определения. В простейшем случае область содержит два значения (да, нет; хорошо, плохо). Это может соответствовать, например, годной продукции и браку.

4. Виды параметров оптимизации

В зависимости от объекта и цели параметры оптимизации могут быть весьма разнообразными. Введем некоторую классификацию. Реальные ситуации, как правило, довольно сложны. Они часто требуют нескольких, иногда очень многих параметров. Каждый объект может характеризоваться сразу всей совокупностью параметров или любым подмножеством из этой совокупности. Движение к оптимуму возможно, если выбран один-единственный параметр оптимизации. Тогда прочие характеристики процесса уже не выступают в качестве параметров оптимизации, а служат ограничениями. Другой путь – построение обобщенного параметра оптимизации как некоторой функции от множества исходных.

Экономические параметры оптимизации, такие как прибыль, себестоимость и рентабельность, обычно используются при исследовании действующих промышленных объектов, тогда как затраты на эксперимент имеет смысл оценивать в любых исследованиях, в том числе и лабораторных. Если цена опытов одинакова, затраты на эксперимент пропорциональны числу опытов, которые необходимо поставить для решения данной задачи. Это в значительной мере определяет выбор плана эксперимента.

Среди технико-экономических параметров наибольшее распространение имеет производительность. Такие параметры, как долговечность, надежность и стабильность, связаны с длительными наблюдениями. Имеется некоторый опыт их использования при изучении дорогостоящих ответственных объектов, например радиоэлектронной аппаратуры.

Почти во всех исследованиях приходится учитывать количество и качество получаемого продукта. Как меру количества продукта используют выход, например, процент выхода химической реакции, выход годных изделий. Показатели качества чрезвычайно разнообразны. Характеристики количества и качества продукта образуют группу технико-технологических параметров.

Под рубрикой «прочие» сгруппированы различные параметры, которые реже встречаются, но не являются менее важными. Сюда попали статистические параметры, используемые для улучшения характеристик случайных величин или случайных функций. В качестве примеров можно назвать задачи на минимизацию дисперсии случайной величины, на уменьшение числа выбросов случайного процесса за фиксированный уровень и т. д. Последняя задача возникает, в частности, при выборе оптимальных настроек автоматических регуляторов или при улучшении свойств нитей (проволока, пряжа, искусственное волокно и др.).

Требования к параметрам оптимизации:

1. Параметр оптимизации должен быть количественным.

2. Параметр оптимизации должен выражаться одним числом. Иногда это получается естественно, как регистрация показания прибора. Например, скорость движения машины определяется числом на спидометре. Часто приходится проводить некоторые вычисления. Так бывает при расчете выхода реакции. В химии часто требуется получать продукт с заданным соотношением компонентов. Один из возможных вариантов решения подобных задач состоит в том, чтобы выразить соотношение одним числом (например, 2) и в качестве параметра оптимизации пользоваться значением отклонений (или квадратов отклонений) от этого числа.

3. Однозначность в статистическом смысле. Заданному набору значений факторов должно соответствовать одно значение параметра оптимизации, при этом обратное неверно: одному и тому же значению параметра могут соответствовать разные наборы значений факторов.

4. Наиболее важным требованием к параметрам оптимизации является его возможность действительно эффективной оценки функционирования системы. Представление об объекте не остается постоянным в ходе исследования. Оно меняется по мере накопления информации и в зависимости от достигнутых результатов. Это приводит к последовательному подходу при выборе параметра оптимизации. Так, например, на первых стадиях исследования технологических процессов в качестве параметра оптимизации часто используется выход продукта. Однако в дальнейшем, когда возможность повышения выхода исчерпана, начинают интересоваться такими параметрами, как себестоимость, чистота продукта и т. д. Оценка эффективности функционирования системы может осуществляться как для всей системы в целом, так и оценкой эффективности ряда подсистем, составляющих данную систему. Но при этом необходимо учитывать возможность того, что оптимальность каждой из подсистем по своему параметру оптимизации не исключает возможность гибели системы в целом. Это означает, что попытка добиться оптимума с учетом некоторого локального или промежуточного параметра оптимизации может оказаться неэффективной или даже привести к браку.

5. Требование универсальности или полноты. Под универсальностью параметра оптимизации понимают его способность всесторонне охарактеризовать объект. В частности, технологические параметры недостаточно универсальны: они не учитывают экономику. Универсальностью обладают, например, обобщенные параметры оптимизации, которые строятся как функции от нескольких частных параметров.

6. Параметр оптимизации желательно должен иметь физический смысл, быть простым и легко вычисляемым. Требование физического смысла связано с последующей интерпретацией результатов эксперимента. Не представляет труда объяснить, что значит максимум извлечения, максимум содержания ценного компонента. Эти и подобные им технологические параметры оптимизации имеют ясный физический смысл, но иногда для них может не выполняться, например, требование статистической эффективности. Тогда рекомендуется переходить к преобразованию параметра оптимизации. Второе требование, т. е. простота и легковычисляемость, также весьма существенно. Для процессов разделения термодинамические параметры оптимизации более универсальны. Однако на практике ими пользуются мало: их расчет довольно труден. Из приведенных двух требований первое является более существенным, потому что часто удается найти идеальную характеристику системы и сравнить ее с реальной характеристикой.

5. Факторы

После выбора объекта исследования и параметра оптимизации нужно рассмотреть все факторы, которые могут влиять на процесс. Если какой-либо существенный фактор окажется неучтенным и примет произвольные значения, не контролируемые экспериментатором, то это значительно увеличит ошибку опыта. При поддержании этого фактора на определенном уровне может быть получено ложное представление об оптимуме, так как нет гарантии, что полученный уровень является оптимальным.

С другой стороны, большое число факторов увеличивает число опытов и размерность факторного пространства.

Выбор факторов эксперимента является весьма существенным, от этого зависит успех оптимизации.

Фактор – измеряемая переменная величина, принимающая в некоторый момент времени определенное значение и влияющая на объект исследования.

Факторы должны иметь область определения, внутри которой задаются его конкретные значения. Область определения может быть непрерывной или дискретной. При планировании эксперимента значения факторов принимаются дискретными, что связано с уровнями факторов. В практических задачах области определения факторов имеют ограничения, которые носят либо принципиальный, либо технический характер.

Факторы разделяются на количественные и качественные.

К *количественным* относятся те факторы, которые можно измерять, взвешивать и т. д.

Качественные факторы – это различные вещества, технологические способы, приборы, исполнители и т. п.

Хотя качественным факторам не соответствует числовая шкала, при планировании эксперимента к ним применяют условную порядковую шкалу в соответствии с уровнями, т. е. производится кодирование. Порядок уровней здесь произволен, но после кодирования он фиксируется.

Требования к факторам эксперимента:

1. Факторы должны быть управляемыми, это значит, что выбранное нужное значение фактора можно поддерживать постоянным в течение всего опыта. Планировать эксперимент можно только в том случае, если уровни факторов подчиняются воле экспериментатора. Например, экспериментальная установка смонтирована на открытой площадке. Здесь температурой воздуха мы не можем управлять, ее можно только контролировать, и потому при выполнении опытов температуру как фактор мы не можем учитывать.

2. Чтобы точно определить фактор, нужно указать последовательность действий (операций), с помощью которых устанавливаются его конкретные значения. Такое определение называется операциональным. Так, если фактором является давление в некотором аппарате, то совершенно необходимо указать, в какой точке и с помощью какого прибора оно измеряется и как оно устанавливается. Введение операционального определения обеспечивает однозначное понимание фактора.

3. Точность замеров факторов должна быть возможно более высокой. Степень точности определяется диапазоном изменения факторов. В длительных процессах, измеряемых многими часами, минуты можно не учитывать, а в быстрых процессах приходится учитывать доли секунды.

Исследование существенно усложняется, если фактор измеряется с большой ошибкой или значения факторов трудно поддерживать на выбранном уровне (уровень фактора «плывет»), приходится применять специальные методы исследования, например, конфлюэнтный анализ.

4. Факторы должны быть однозначны. Трудно управлять фактором, который является функцией других факторов. Но в планировании могут участвовать другие факторы, такие как соотношения между компонентами, их логарифмы и т. п. Необходимость введения сложных факторов возникает при желании представить динамические особенности объекта в статической форме. Например, требуется найти оптимальный режим подъема температуры в реакторе. Если относительно температуры известно, что она должна нарастать линейно, то в качестве фактора вместо функции (в данном случае линейной) можно использовать тангенс угла наклона, то есть градиент.

5. При планировании эксперимента одновременно изменяют несколько факторов, поэтому необходимо знать требования к совокупности факторов. Прежде всего выдвигается требование совместимости. Совместимость факторов означает, что все их комбинации осуществимы и безопасны. Несовместимость факторов наблюдается на границах областей их определения. Избавиться от нее можно сокраще-

нием областей. Положение усложняется, если несовместимость проявляется внутри областей определения. Одно из возможных решений – разбиение на подобласти и решение двух отдельных задач.

6. При планировании эксперимента важна независимость факторов, то есть возможность установления фактора на любом уровне вне зависимости от уровней других факторов. Если это условие невыполнимо, то невозможно планировать эксперимент.

6. Требования к совокупности факторов

При планировании эксперимента обычно одновременно изменяется несколько факторов. Поэтому очень важно сформулировать требования, которые предъявляются к совокупности факторов. Прежде всего выдвигается требование совместимости. Совместимость факторов означает, что все их комбинации осуществимы и безопасны. Это очень важное требование. Представьте себе, что вы поступили легкомысленно, не обратили внимания на требование совместимости факторов и запланировали такие условия опыта, которые могут привести к взрыву установки или осмолению продукта. Согласитесь, что такой результат очень далек от целей оптимизации.

Несовместимость факторов может наблюдаться на границах областей их определения. Избавиться от нее можно сокращением областей. Положение усложняется, если несовместимость проявляется внутри областей определения. Одно из возможных решений – разбиение на подобласти и решение двух отдельных задач.

При планировании эксперимента важна независимость факторов, то есть возможность установления фактора на любом уровне вне зависимости от уровней других факторов. Если это условие невыполнимо, то невозможно планировать эксперимент. Итак, мы подошли ко второму требованию – отсутствию корреляции между факторами. Требование некоррелированности не означает, что между значениями факторов нет никакой связи. Достаточно, чтобы связь не была линейной.

7. Планирование эксперимента. План эксперимента

При проведении активного эксперимента задается определенный план варьирования факторов, то есть эксперимент заранее планируется.

План эксперимента – совокупность данных, определяющих число, условия и порядок реализации опытов.

Планирование эксперимента – выбор плана эксперимента, удовлетворяющего заданным требованиям.

Точка плана – упорядоченная совокупность численных значений факторов, соответствующая условиям проведения опыта, точка факторного пространства, в которой проводится эксперимент. Точке плана с номером i соответствует вектор-строка

$$\vec{X}_{i1} = (X_{i1}, X_{i2}, \dots, X_{in}) \tag{7.4}$$

Общая совокупность таких векторов X_i, где i принимает значения от 1 до L, образует план эксперимента, а совокупность различных векторов, число которых обозначим N, – *спектр плана*.

В активном эксперименте факторы могут принимать только фиксированные значения. Фиксированное значение фактора называют *уровнем фактора*. Количество принимаемых уровней факторов зависит от выбранной структуры факторной модели и принятого плана эксперимента. Минимальный $X_{j\min}$ и максимальный $X_{i\max}$, $j = 1, \dots, n$ (n – число факторов) уровни всех факторов выделяют в факторном пространстве некоторый гиперпараллелепипед, представляющий собой область планирования. В области планирования находятся все возможные значения факторов, используемые в эксперименте.

Вектор $\vec{X}^0 = (X_1^0, X_2^0, \dots, X_n^0)$ задает точку центра области планирования. Координаты этой точки X_{j0} обычно выбирают из соотношения

$$X_j^0 = \frac{(X_{j\max} + X_{j\min})}{2}. \tag{7.5}$$

Точку X^0 называют центром эксперимента. Она определяет основной уровень факторов X_j^0, $j = 1, \dots, n$. Центр эксперимента стремятся выбрать как можно ближе к точке, которая соответствует искомым оптимальным значениям факторов, как это представляется на основе априорной информации об объекте исследования.

Интервалом (или шагом) варьирования фактора X_j называют величину, вычисляемую по формуле

$$\Delta X_j = \frac{(X_{j\max} - X_{j\min})}{2}, \tag{7.6}$$

где $j = 1, \ldots, n$.

Факторы нормируют, а их уровни кодируют. В кодированном виде верхний уровень обозначают +1, нижний –1, а основной – 0 (рис. 7.2). Нормирование факторов осуществляют на основе соотношения

$$x_j = \frac{(X_j - X_{0j})}{\Delta X_j}. \tag{7.7}$$

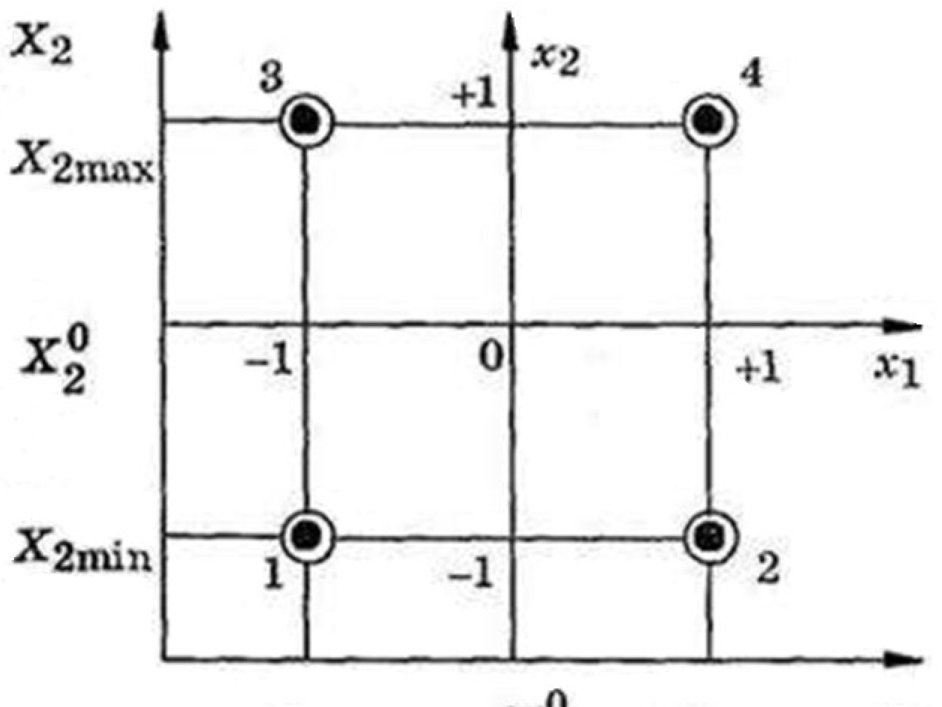


Рис. 7.2. Геометрическое представление области планирования при двух факторах: X_1 и X_2

Точки 1, 2, 3, 4 являются точками плана эксперимента. Например, значения факторов X_1 и X_2 в точке 1 равны соответственно $X_{1\min}$ и $X_{2\min}$, а нормированные их значения $X_{1\min} = -1$, $X_{2\min} = -1$.

После установления нулевой точки выбирают интервалы варьирования факторов. Это связано с определением таких значений факторов, которые в кодированных величинах соответствуют +1 и –1. Интервалы варьирования выбирают с учетом того, что значения факторов, соответствующие уровням +1 и –1, должны быть достаточно отличимы от значения, соответствующего нулевому уровню. Поэтому во всех случаях величина интервала варьирования должна быть больше удвоенной квадратичной ошибки фиксирования данного фактора. С другой стороны, чрезмерное увеличение величины интервалов варьирования нежелательно, так как это может привести к снижению эффективности поиска оптимума. При этом очень малый интервал варьирования уменьшает область эксперимента, что замедляет поиск оптимума.

При выборе интервала варьирования целесообразно учитывать, если это возможно, число уровней варьирования факторов в области эксперимента. От числа уровней зависят объем эксперимента и эффективность оптимизации.

План эксперимента удобно представлять в матричной форме. Матрица плана представляет собой прямоугольную таблицу, содержащую информацию о количестве и условиях проведения опытов. Строки матрицы плана соответствуют опытам, а столбцы – факторам. Размерность матрицы плана $L \times n$, где L – число опытов, n – число факторов. При проведении повторных (дублирующих) опытов в одних и тех же точках плана матрица плана содержит ряд совпадающих строк.

Матрица спектра плана – матрица, в которую входят только различающиеся между собой строки матрицы плана. Размерность матрицы спектра плана $N \times n$, где N – число точек плана, различающихся между собой хотя бы одной координатой U.

Матрица спектра плана имеет вид

$$X = \begin{vmatrix} x_1 \\ \bar{x}_2 \\ \bar{x}_i \\ \bar{x}_n \end{vmatrix} = \begin{vmatrix} x_{11} \\ x_{21} x_{22} \\ \bar{x}_{i1}\ \bar{x}_{i2}\ \dots\ \bar{x}_{ij}\ \dots\ \bar{x}_{in} \\ \bar{x}_{n1}\ \bar{x}_{n2}\ \dots\ \bar{x}_{nj}\ \dots\ \bar{x}_{nn} \end{vmatrix}. \quad (7.8)$$

Структура экспериментальной факторной модели

Под структурой экспериментальной факторной математической модели понимается вид математических соотношений между факторами *X*, *Z* и откликом *Y*. В качестве факторов принимают внутренние и внешние параметры технической системы, подлежащие оптимизации в процессе ее проектирования.

Внутренние параметры системы – это параметры ее элементов, *внешние* являются параметрами внешней среды, воздействующие на систему во время ее работы. Функциями отклика *Y* являются выходные параметры технической системы, которые характеризуют ее эффективность и качество процессов функционирования. Выходные параметры системы принимаются в качестве критериев оптимальности.

Структура факторной модели выбирается на основе априорной информации, используя принцип постепенного ее усложнения. Параметры факторной математической модели определяются методами регрессионного анализа. При определении параметров этими методами нет необходимости различать виды факторов, то есть подразделять факторы на управляемые *X* и неуправляемые *Z*. Поэтому в дальнейшем все они будут обозначаться буквой *X*. Тогда факторную модель можно представить векторным уравнением регрессии вида

$$\vec{Y} = \vec{\varphi}\left(\vec{X}, \vec{B}\right). \quad (7.9)$$

Для определения параметров используются результаты эксперимента. Результаты эксперимента можно представить функцией вида

$$Y = \varphi\left(\vec{X}\right) + \varepsilon, \quad (7.10)$$

где ε – аддитивная помеха случайного характера с нормальным распределением.

В качестве базисных функций используют переменные простейших полиномов, системы ортогональных полиномов, тригонометрические функции. Наиболее часто пользуются простейшими полиномами первой и второй степеней. Например, полином первой степени, описывающий функцию отклика Y при двух факторах x_1 и x_2, может иметь вид

$$y = b_0 + b_1x_1 + b_2x_2 \tag{7.11}$$

или

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_1x_2, \tag{7.12}$$

а полином второй степени будет иметь вид

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_1x_2 + b_4x_{12} + b_5x_{22}. \tag{7.13}$$

Базисные функции в случае использования последнего выражения имеют вид

$f_0(X) = 1;$

$f_1(X) = x_1;$

$f_2(X) = x_2;$

$f_3(X) = x_1x_2;$

$f_4(X) = x_{12};$

$f_5(X) = x_{22}.$

8. Планы экспериментов и их свойства. Виды экспериментов

Для проведения активных экспериментов разработано множество различных планов. Планы учитывают как особенности структуры регрессионных моделей, так и требования их эффективности с позиций повышения точности получаемых моделей и снижения затрат на проведение эксперимента.

При построении линейных моделей или нелинейных, содержащих только взаимодействия факторов, но без квадратов этих факторов, каждый фактор можно варьировать только на двух уровнях. Для получения таких моделей используют планы первого порядка.

Известно несколько разновидностей планов первого порядка. Они предназначены для планирования следующих видов экспериментов:

– однофакторного (классического) эксперимента;

– полного факторного эксперимента;

– дробного факторного эксперимента.

Если в регрессионную модель входят факторы в квадрате или с более высокими степенями, то необходимо не менее трех уровней варьирования факторов. При построении квадратичных моделей применяют планы второго порядка.

Планы различают по степени насыщенности и композиционности. План называют насыщенным, если общее число точек плана равно числу неизвестных параметров регрессионной модели. Такой план позволяет получить экспериментальную факторную модель при минимальных затратах, так как обеспечивает минимум числа опытов.

План называется композиционным, если в спектр его в качестве составной части входят точки спектра плана, который был реализован при построении более простой модели. Композиционность плана позволяет реализовать принцип постепенного усложнения модели при минимальных затратах, так как при этом используются результаты опытов, выполненных для получения простой модели. Многие планы второго порядка являются композиционными.

8.1. План однофакторного эксперимента

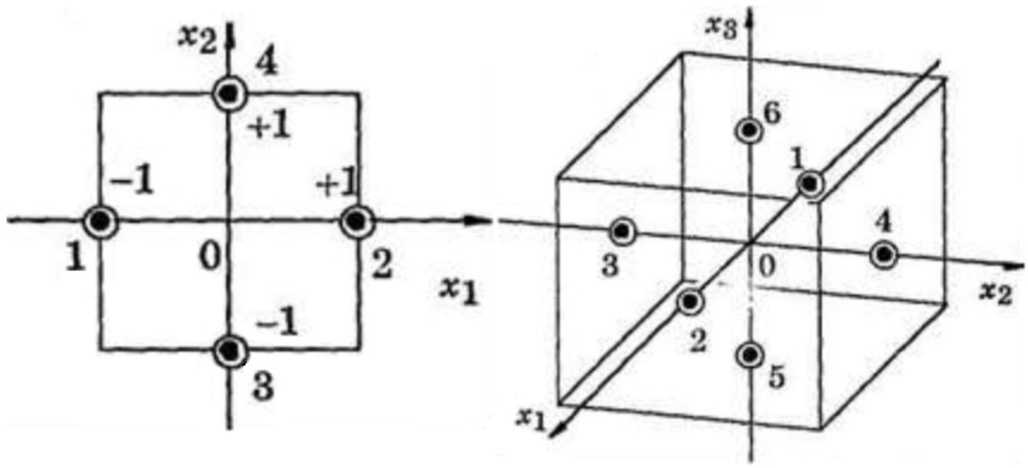
Однофакторный (классический) эксперимент предназначен для получения линейной экспериментальной факторной модели вида

$$y = b_0 + b_1x_1 + b_2x_2 + \ ... + b_nx_n. \tag{7.14}$$

Однофакторный эксперимент предусматривает поочередное варьирование каждого из факторов при фиксированных на некотором уровне значениях остальных факторов. Фактор X_i варьируют на двух уровнях – X_{iB} и X_i. Все остальные факторы при этом должны находиться в точке центра эксперимента X_{j0}, j, i. Для нормированных факторов $x_{jB} = +1$, $x_{iH} = -1$, $x_j = 0$. С учетом этого составим матрицу спектра плана однофакторного эксперимента

$$X = \begin{vmatrix} -1 & 0 & ... & 0 \\ +1 & 0 & ... & 0 \\ 0 & -1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & -1 \\ 0 & 0 & ... & +1 \end{vmatrix}. \tag{7.15}$$

Число точек плана в этом случае $N = 2n$, где n – количество факторов (рис. 7.3).



а) б)

Рис. 7.3. Схема однофакторного эксперимента: а – при $n = 2$; б – при $n = 3$

Вектор базисных функций имеет вид

$$\vec{f}(\vec{x}) = (1, x_1, x_1, \dots, x_n). \tag{7.16}$$

8.2. План полного факторного эксперимента

Спектр плана полного факторного эксперимента (ПФЭ) содержит все возможные комбинации значений факторов на всех уровнях их изменения. Число точек N спектра плана определяется по формуле

$$N = U^n, \tag{7.17}$$

где U – число уровней варьирования факторов; n – количество факторов.

Рассмотрим особенности и свойства ПФЭ, применяемые при построении линейных регрессий. Для получения линейной регрессии достаточно варьировать факторы на двух уровнях, $U = 2$. Тогда число точек спектра плана будет

$$N = 2n. \tag{7.18}$$

Такой план принято обозначать ПФЭ2n.

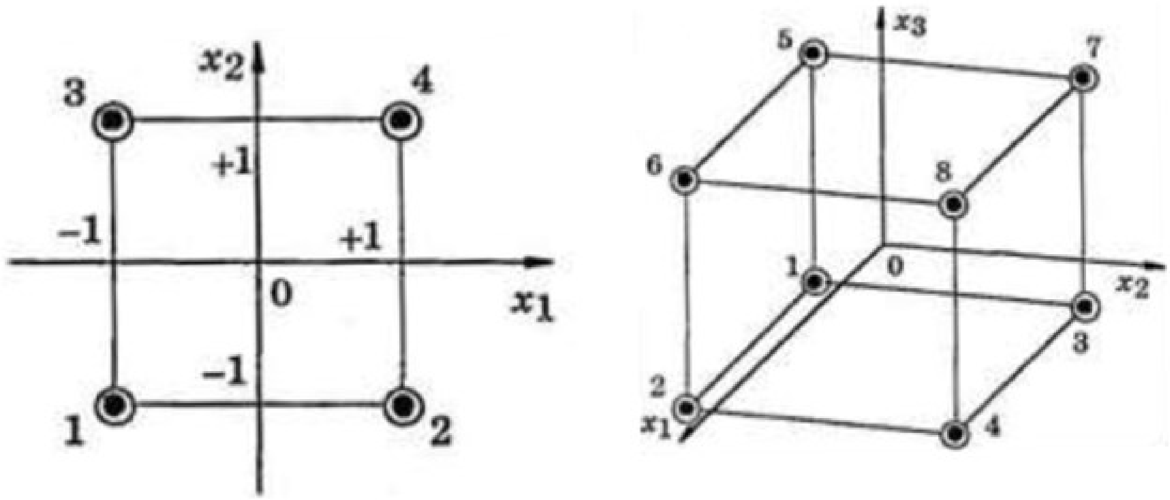
Рассмотрим порядок составления матрицы спектра плана, полагая, что факторы нормированы и, следовательно, могут принимать значения только либо +1, либо –1.

Для составления матрицы спектра плана используется следующее простое правило: в первой строке матрицы все факторы равны –1, в первом столбце знаки единиц меняются поочередно; во втором столбце они чередуются через два; в третьем – через 4; в четвертом – через 8 и так далее по степеням двойки.

При $n = 2$ число точек плана $N = 2 \cdot 2 = 4$, а матрица спектра плана имеет вид

$$X = \begin{vmatrix} -1 & -1 \\ +1 & -1 \\ -1 & +1 \\ +1 & +1 \end{vmatrix}. \tag{7.19}$$

Точки плана ПФЭ2n располагаются в вершинах n- мерного гиперкуба (рис. 7.4).



а) *б)*

Рис. 7.4. Точки плана ПФЭ2n: а – при n = 2; б – при n = 3

Посредством ПФЭ можно построить как простейшую линейную модель технической системы вида

$$y = b_0 + b_1x_1 + b_2x_2 + \ \dots + b_nx_n, \quad (7.21)$$

так и нелинейную.

Порядок выполнения работы

1. Написать краткий отчет об изучении разделов работы.
2. Ответить на контрольные вопросы.

Контрольные вопросы

1. Перечислите основные требования к планированию эксперимента.
2. Содержание плана эксперимента.
3. Перечислите цели и задачи эксперимента.
4. Что такое фактор эксперимента?
5. Назовите разновидности планов эксперимента.

ЗАКЛЮЧЕНИЕ

Содержание практикума соответствует образовательному стандарту по дисциплине «Основы научных исследований» для направления подготовки «Техносферная безопасность». С целью ограничения объема издания материалы ряда глав даны в сокращенном виде.

Опыт научных исследований в области технических наук определил сложившуюся последовательность выполнения этапов исследовательских работ. Практикум охватывает вопросы, которые помогут студенту участвовать в исследовательских работах, правильно обрабатывать полученные в ходе исследований результаты измерений.

Учебный материал изложен в последовательности, способствующей углубленному изучению вопросов, касающихся математического аппарата, используемого при научных исследованиях.

Освоение материала практикума даст желаемый эффект только при обязательном закреплении теоретических знаний практическими работами.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Голубева, Н. В. Математическое моделирование систем и процессов [Электронный ресурс] : учеб. пособие для вузов / Н. В. Голубева. – 4-е изд., испр. и доп. – СПб. : Лань, 2024. – 244 с. – ISBN 978-5-8114-1424-6. – URL: https://e.lanbook.com/book/393023 (дата обращения: 28.08.2025).

2. Асхаков, С. И. Основы научных исследований [Электронный ресурс] : учеб. пособие / С. И. Асхаков. – Карачаевск : Изд-во КЧГУ, 2020. – 348 с. – URL: https://e.lanbook.com/book/161998 (дата обращения: 28.08.2025).

3. Трутнев, Н. В. Основы научных исследований в технических системах : практикум / Н. В. Трутнев, Е. А. Лялин. – Пермь : Прокростъ, 2024. – 103 с. – ISBN 978-5-94279-628-0.

4. Курасов, В. С. Испытания автомобилей и тракторов [Электронный ресурс] : учеб. пособие для вузов / В. С. Курасов, В. М. Погосян, В. В. Драгуленко. – 2-е изд., стер. – СПб. : Лань, 2024. – 84 с. – URL: https://e.lanbook.com/book/402002 (дата обращения: 28.08.2025). – ISBN 978-5-507-47655-8.

5. Поливаев, О. И. Испытание сельскохозяйственной техники и энергосиловых установок [Электронный ресурс] : учеб. пособие / О. И. Поливаев, О. М. Костиков. – СПб. : Лань, 2016. – 276 с. – URL: https://e.lanbook.com/book/209738 (дата обращения: 28.08.2025). – ISBN 978-5-8114-2108-4.

6. Трубицын, В. А. Основы научных исследований [Электронный ресурс] : учеб. пособие / В. А. Трубицын, А. А. Порохня, В. В. Мелешин. – Ставрополь : Изд-во СКФУ, 2016. – 149 с. – URL: https://e.lanbook.com/book/155174 (дата обращения: 28.08.2025).

7. Кошурников, А. Ф. Основы научных исследований : учеб. пособие / А. Ф. Кошурников ; Пермская ГСХА. – Пермь : Прокрость, 2014. – 317 с.

8. Никитин, А. М. Основы научных исследований [Электронный ресурс] : учеб.-метод. пособие по выполнению практ. работ для студентов направлений подгот. 13.03.02 – Электроэнергетика и электротехника, 35.03.06 – Агроинженерия / А. М. Никитин. – Брянск : Брянск. гос. аграр. ун-т, 2024. – 110 с. // Цифровой образовательный ресурс IPR SMART : [сайт]. – URL: https://www.iprbookshop. ru/147598.html (дата обращения: 01.11.2025).

ПРИЛОЖЕНИЕ

Таблица П1

Исходные данные для выполнения практической работы № 1

	Варианты исходных данных														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0,38	4,6	452	0,8896	6,688	0,34	3,6	350	0,8087	4,674	1,14	7,2	741	1,13	24,144
2	0,38	5,4	446	0,7504	7,38	0,35	4,2	350	0,7112	4,947	1,14	9,8	741	0,828	23,532
3	0,38	6,2	449	0,6555	6,815	0,34	4,8	353	0,6082	3,981	1,14	8	741	1,012	24,249
4	0,38	7	452	0,5787	6,615	0,34	3,8	353	0,7591	5,117	1,14	6,6	741	1,2351	23,159
5	0,39	4	452	1,0485	4,901	0,34	3,2	353	0,9316	4,183	1,14	6	741	1,3596	21,573
6	0,26	2	367	1,3805	4,495	0,53	4,4	456	1,027	9,631	1,04	6	697	1,2324	22,031
7	0,26	2,3	367	1,2117	4,328	0,53	5,4	456	0,8376	9,333	1,04	8	697	0,9247	21,66
8	0,26	2,6	367	1,0759	4,126	0,53	6	456	0,7534	9,063	1,04	9,6	697	0,7738	20,601
9	0,26	1,71	364	1,654	4,201	0,53	4	456	1,137	8,782	1,04	5,6	697	1,3314	21,227
10	0,26	2,3	364	1,2317	3,514	0,53	3,8	456	1,2007	7,168	1,04	5,8	697	1,2909	20,503
11	0,5	5,6	540	0,9379	9,766	0,65	5,4	547	1,0174	13,192	0,95	6,6	647	1,0272	20,186
12	0,51	7,4	540	0,7216	9,554	0,65	6,2	547	0,8867	13,032	0,95	7,8	647	0,8721	19,623
13	0,5	8	540	0,6539	9,337	0,65	7	547	0,7853	12,618	0,97	9	647	0,7746	18,592
14	0,5	5	540	1,0557	8,976	0,66	5	547	1,1197	12,887	0,96	6	647	1,1492	20,229
15	0,51	4,6	540	1,1771	7,321	0,66	4,6	547	1,2236	11,494	0,96	5,8	647	1,1916	19,715
16	0,63	6,4	640	1,0249	13,118	0,81	5,8	650	1,1872	16,586	0,8	5,8	550	0,9763	16,039
17	0,63	8	640	0,815	12,84	0,8	8	650	0,8486	16,305	0,81	7	550	0,8166	15,102
18	0,63	10	643	0,6529	12,584	0,81	9	650	0,765	15,6	0,81	8	550	0,7144	14,45
19	0,63	6,2	643	1,0562	12,267	0,81	5,2	650	1,3358	14,336	0,81	5	550	1,1465	15,097
20	0,63	5,6	643	1,1795	10,873	0,81	5,6	650	1,2375	15,689	0,81	5,4	550	1,0535	15,849
21	0,71	7,4	699	0,9922	14,727	0,88	6,6	700	1,1381	18,363	0,62	5,4	447	0,8118	11,407
22	0,71	9,2	699	0,7955	14,529	0,88	9	700	0,8382	18,245	0,62	5,9	447	0,7427	10,921
23	0,69	11	699	0,6667	13,65	0,89	9,8	700	0,7785	17,709	0,62	6,6	447	0,6643	9,944

Окончание табл. П1

	Варианты исходных данных														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
24	0,7	6,2	699	1,2057	12,904	0,9	5,8	700	1,3309	17,236	0,62	5	447	0,8779	11,56
25	0,7	6	699	1,254	11,872	0,9	5,2	700	1,4716	18,741	0,62	4,6	447	0,9569	11,376
26	0,73	7,6	749	1,0227	14,945	0,93	6,8	744	1,1545	19,907	0,44	3,8	350	0,8282	6,375
27	0,73	9,4	749	0,8267	14,496	0,93	9	744	0,8706	19,369	0,45	4,4	350	0,7318	6,203
28	0,73	10,2	749	0,7638	14,027	0,93	8	744	0,9817	19,875	0,46	4,8	350	0,6847	5,571
29	0,74	7	749	1,1264	14,304	0,93	6	744	1,3209	17,607	0,46	3,4	350	0,9684	6,698
30	0,74	6,4	749	1,2464	12,253	0,93	6,4	744	1,2402	19,169	0,47	3	350	1,1334	5,375

Таблица П2

Исходные данные для выполнения практической работы № 2

	Варианты исходных данных														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0,8087	1,2133	0,8526	0,8896	1,13	1,2007	6,6	6,2	6	1,4473	4,6	5	6	5,6	0,8777
2	0,7112	1,1077	0,9615	0,7504	0,828	1,0174	5,8	7,8	7,8	1,1058	5,2	5,4	6,8	8	0,6639
3	0,6082	0,8465	0,7541	0,6555	1,012	0,8867	7,4	9	9	0,96	7	6	6,4	7	0,7706
4	0,7591	1,31	1,0537	0,5787	1,7351	0,7853	5,4	7	5,2	1,6509	6,4	7	5	4,4	0,9657
5	0,9316	1,2512	1,1887	1,0485	1,3596	1,1197	3,8	6,8	6,4	1,3689	5,8	6,5	5,6	5	0,9186
6	1,7301	1,1616	1,0933	1,3805	1,2324	1,3805	6,4	7,2	7	1,1283	7,4	5,8	4,6	5	0,8652
7	0,8376	0,9939	0,8991	1,2117	0,9247	1,2117	7,8	8,8	7	1,1618	9	5	3	6,7	0,6931
8	0,7534	0,8701	0,7922	1,0759	0,7738	1,0759	8,8	9,6	8	1,0125	12	5,4	4,2	6,4	0,8074
9	1,137	1,2942	1,2474	1,654	1,3314	1,654	5,6	6	7,5	1,0715	5,8	7	5,4	4,6	1,2956
10	1,2007	1,2437	1,1646	1,2317	1,2909	1,2317	6	6,4	5,8	1,3903	5,2	6,4	5	4,8	1,1554
11	1,0174	1,1301	1,0287	0,9379	1,0272	0,9379	7,4	7,2	4,8	0,9709	7,8	7	4,4	5,2	1,1355
12	0,8867	1,0179	0,8727	0,7216	0,8721	0,7216	8,8	9	3,3	1,1889	9,4	6,4	4	6	0,9068
13	0,7853	0,8676	1,2383	0,6539	0,7746	0,6539	6,2	8	5,5	0,8459	10,6	6	5	5,6	0,7749
14	1,1197	1,2922	0,9505	1,0557	1,1492	1,13	8	6,2	6	0,7735	7	5,6	6	4,8	1,6981
15	1,2236	1,4513	1,426	1,1771	1,1916	0,828	5,4	6,8	6,5	0,7232	5,6	5	5,5	5	1,383
16	1,1872	1,1536	1,0617	1,0249	0,9763	1,012	7,6	5,4	5,5	1,0723	7,8	4,2	5	5,4	1,1676
17	0,8486	0,9263	0,9165	0,815	0,8166	1,2351	8,8	6	8	0,7456	9,6	6	4,2	6,3	0,9338
18	0,765	0,5875	0,841	0,6529	0,7144	1,3596	9,6	5,8	8,1	0,7312	10,8	5	4,6	5,8	0,8296
19	1,3358	1,3405	1,1862	1,0562	1,1465	1,2324	6,8	5	5	1,1876	6,2	4	6	5,6	1,4408
20	1,2375	1,4411	1,1225	1,1795	1,0535	1,1353	7,2	3,9	5,6	1,053	7,6	3,3	5,5	6	1,5005

Таблица П3

Исходные данные для выполнения практической работы № 3
(Износ втулки шатуна автомобильного двигателя, отнесенный к 1000 мото-часов, мкм)

	Варианты исходных данных														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	58,98	57,61	34,63	57,55	59,62	56,83	49,83	58,65	42,39	51,45	47,14	46,6	35,52	36,33	50,75
2	56,32	50,44	59,11	34,81	33,61	53,8	59,19	45,4	33,93	34,52	57,77	32,43	41,12	37,02	54,86
3	38,82	40,7	58,46	38,02	32,98	51,6	32,71	44,88	39,86	45,07	50,26	41,29	34,94	42,39	50,65
4	55,81	30,06	31,86	55,57	53,21	32,16	30,27	41,6	52,47	52,68	32,67	32,23	48,66	49,74	42,73
5	37,84	52,92	30,25	39,46	58,83	59,45	46,5	51,32	46,29	32,72	42,85	32,07	56,45	43,21	33,86
6	55,88	44,04	47,99	30,14	56,35	48,59	31,82	44,32	38,65	33,67	56,45	51,3	41,73	45,89	37,96
7	45,44	53,57	44,36	39,89	31,83	45,78	31,06	48,19	38,62	35,77	34,35	47,98	34,63	55,65	56,35
8	41,51	54,19	38,12	33,94	50,68	41,5	34,53	58,94	34,44	47,67	58,72	46,87	42,77	44,85	34,89
9	52,85	54,39	55,79	51,71	34,41	49,88	43,29	37,96	43,17	54,5	43,43	34,19	58,02	58,47	59,58
10	46,69	58,59	30,27	37,01	59,19	50,12	46,13	43,14	41,02	50,38	44,65	57,72	38,11	32,97	35,16

Таблица П4

Исходные данные для выполнения практической работы № 4

	Варианты исходных данных														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	2,5	1,5	2	2	2,5	3	1,5	2	5	2,5	3,5	3,5	2	5	3,875
2	3,5	3,5	3,5	2	5	3,5	3,5	3,125	3,5	2,5	5	2	3,5	3,5	3,875
3	2	2	3	2	3	3	2	2,25	2,5	2,5	2,5	1	1	2,5	2,125
4	6	5	4,5	5	5,5	5,5	5	5,125	4	5	3,5	5	4,5	4	4
5	4,5	4	3,5	3,5	4,5	4	4	3,875	4,5	4	2,5	4,5	2,5	4,5	3,5
6	4	4	4,5	5	3,5	4,5	4	4,375	5	3	3	3	1	5	3,5
7	5	3,5	2	3,5	3,5	3,5	3,5	3,5	3	3	2,5	5	4,5	3	3,25
8	3,5	2	3,5	4,5	2,5	5	2	3,375	3	4	4	3	4	3	3,5
9	3,5	3	3	4,5	2	3,5	3	3,5	4,5	3,5	2,5	5	3	4,5	3,625
10	2,5	2	2	3,5	1	2,5	2	2,5	3,5	1,5	1	4	1,5	3,5	2,375
11	3,5	2,5	3	5	1,5	4	2,5	3,5	4	6	4,5	4	4,5	4	4,25
12	3,5	3,5	3,5	6	1	3,5	3,5	4,125	1,5	3,5	1,5	4	4	1,5	2,125
13	3	4	3,5	3	3,5	2,5	4	3,375	3	3,5	2	3,5	2,5	3	2,625
14	2,5	2,5	4	2	4,5	4	2,5	2,75	4,5	4	5,5	2,5	3,5	4,5	4,5
15	2,5	3	3	3,5	2	2,5	3	3	3	1,5	4	2,5	3,5	3	3,375
16	1,5	2	2	1,5	2	1,5	2	1,75	5	5	5	6	6	5	5,25
17	3,5	3,5	1,5	2	3	1,5	3,5	2,625	4,5	3,5	3,5	4	3	4,5	3,875
18	4	5	3,5	4	3,5	2,5	5	4,125	2	2	3,5	2	3,5	2	2,75
19	2	4,5	5	3,5	3,5	2,5	4,5	3,75	2,5	4,5	2,5	3	3	2,5	2,625
20	5	5	4	4	5	4	5	4,5	4,5	3	4,5	2,5	2,5	4,5	4
21	2,5	4,5	5,5	4	4	3,5	4,5	4,125	5,5	4	4,5	3,5	2,5	5,5	4,5
22	2,5	3	3,5	4,5	1,5	3	3	3,375	4	5	5	3	4	4	4,25
23	5	3,5	3	3,5	4,5	4,5	3,5	3,75	2	3,5	3	4,5	5,5	2	3,125
24	4,5	4	5	4,5	5	5,5	4	4,5	4	4,5	4,5	5,5	6	4	4,625
25	4,5	4	5	4,5	5	5,5	4	4,5	3,5	5	5	2,5	4	3,5	4
26	3	3	4	5	2	4	3	3,75	3	4	3,5	4	4,5	3	3,5
27	4,5	4	3	2,5	5	3,5	4	3,5	5,5	4	4,5	3,5	2,5	5,5	4,5
28	4	6	5	4	5	3	6	4,75	3,5	4	6	1,5	4	3,5	4,25
29	5,5	3,5	3,5	4	5	5,5	3,5	4,125	3	4	5	2	4	3	3,75
30	2	3	3,5	4	1,5	2,5	3	3,125	1,5	4	4	3,5	6	1,5	3,25

Таблица П5

Исходные данные для выполнения практической работы № 5

Вариант	Координата	1	2	3	4	5	6	7	8
1	*x*	-0,3	1,4	0,4	1,5	1	0,9	2,3	2,2
	y	-1	-3,1	-5,9	-7,8	-9,1	-10,1	-13,6	-13,6
2	*x*	2,1	2,8	4,7	6,4	6,1	8,2	10	11,7
	y	-0,9	-4	-5,3	-7,3	-9,8	-10,1	-13,1	-13,6
3	*x*	1,9	2	3,4	6,3	6,9	9	8,9	11,1
	y	-2,2	-2,9	-5,1	-8,1	-8,3	-9,8	-13,4	-15,3
4	*x*	1,1	0,2	1,8	2,6	2,1	3,1	2,8	2,9
	y	-0,6	-1,9	-2	-5	-4,4	-7	-6,5	-8,6
5	*x*	0,1	1,4	-0,3	1,4	0,5	1,6	1,6	1,5
	y	0,8	4,4	4,7	8	8,4	11,4	13	13,8
6	*x*	0,6	1	0,2	1,1	2,1	2,2	2,9	3,5
	y	0,8	2,4	3,1	3,1	4,4	5,8	6	6,9
7	*x*	1,1	1,4	3,5	5	4	6,2	7,3	8,6
	y	1,3	4,6	4,7	7,8	9,7	11,6	12,9	14,9
8	*x*	-0,8	-3,3	-3,8	-7	-8,9	-8,6	-10,7	-12,9
	y	1,3	0,5	1,5	1,5	1,4	2,7	2,3	3,6
9	*x*	1,2	0,1	1,8	0,6	2,4	3,3	2,4	3,6
	y	2,3	4,2	6,5	8,2	9,3	11,5	14,2	15,9
10	*x*	-1,4	-3,8	-5,9	-7,2	-11	-11,6	-13,9	-16,9
	y	-2,1	-3,1	-4,1	-6,7	-7,7	-9,3	-12,2	-13,6
11	*x*	2	2,5	4,5	4,9	6,6	7,6	9	12,2
	y	-1,2	-3,7	-5,4	-7,2	-10,3	-11,3	-13	-15,9
12	*x*	1,5	3,6	5,1	6,7	7,8	8,9	11,1	13,7
	y	-1,8	-3,7	-5,6	-6,2	-9,9	-11,1	-12,8	-14,9
13	*x*	-0,6	-1,6	-0,9	-1	-1,8	-2	-2,2	-2,8
	y	2,1	2,7	3,8	4,6	6,1	7,6	7,4	10,5
14	*x*	0,7	-0,1	0,5	2,1	1,5	3,1	3	4
	y	0,8	0,9	-1	-0,5	-0,4	-0,4	0,2	-0,5
15	*x*	0,7	3,4	4,3	5,5	7,8	7,6	10,2	11,6
	y	-0,1	-0,9	-2,1	-2,3	-4,6	-4,6	-5,5	-6,6

Таблица П6

Исходные данные для выполнения практической работы № 6

Вариант	Марка шин	Номера автомобилей			
		I	II	III	IV
1	А	14,5	15,1	14,15	14,7
	Б	14,7	15,3	14,35	14,9
	В	14,85	15,45	14,5	15,05
	Г	14,75	15,35	14,4	14,95
2	А	14,15	14,75	14,8	14,25
	Б	14,3	14,9	14,95	14,4
	В	14,35	14,95	15	14,45
	Г	14,45	15,05	15,1	14,55
3	А	14,95	14,7	14,5	14,3
	Б	15,7	15,45	15,25	15,05
	В	15,15	14,9	14,7	14,5
	Г	15,9	15,65	15,45	15,25
4	А	15,2	15,3	15,75	15,25
	Б	14,9	15	15,45	14,95
	В	14,8	14,9	15,35	14,85
	Г	15,25	15,35	15,8	15,3
5	А	14,85	15	14,85	15,35
	Б	14,75	14,9	14,75	15,25
	В	14,4	14,55	14,4	14,9
	Г	14,1	14,25	14,1	14,6
6	А	15,35	15,1	14,85	14,65
	Б	15,65	15,4	15,15	14,95
	В	14,7	14,45	14,2	14
	Г	14,95	14,7	14,45	14,25
7	А	14,45	14,6	14,4	14,25
	Б	14,7	14,85	14,65	14,5
	В	14,95	15,1	14,9	14,75
	Г	14,95	15,1	14,9	14,75
8	А	15,1	14,8	15,05	15,35
	Б	15,05	14,75	15	15,3
	В	14,75	14,45	14,7	15
	Г	15,05	14,75	15	15,3
9	А	14,95	14,15	14,6	14,9
	Б	15,1	14,3	14,75	15,05
	В	15	14,2	14,65	14,95
	Г	15,2	14,4	14,85	15,15

Окончание табл. П6

Вариант	Марка шин	Номера автомобилей			
		I	II	III	IV
10	А	15,35	15,25	15,2	14,95
	Б	14,7	14,6	14,55	14,3
	В	15,05	14,95	14,9	14,65
	Г	15,3	15,2	15,15	14,9
11	А	14,8	14,75	14,95	14,9
	Б	15,25	15,2	15,4	15,35
	В	14,65	14,6	14,8	14,75
	Г	15,35	15,3	15,5	15,45
12	А	15,8	15,35	15,6	15,3
	Б	15,15	14,7	14,95	14,65
	В	15,05	14,6	14,85	14,55
	Г	15,4	14,95	15,2	14,9
13	А	14,2	14,9	14,1	14,4
	Б	14,7	15,4	14,6	14,9
	В	14,45	15,15	14,35	14,65
	Г	14,95	15,65	14,85	15,15
14	А	15,55	15,4	15,75	15,7
	Б	14,95	14,8	15,15	15,1
	В	15,6	15,45	15,8	15,75
	Г	15,7	15,55	15,9	15,85
15	А	15,1	14,5	14,75	14,3
	Б	15,4	14,8	15,05	14,6
	В	15,75	15,15	15,4	14,95
	Г	15,6	15	15,25	14,8

Учебное электронное издание

ОСНОВЫ НАУЧНЫХ ИССЛЕДОВАНИЙ

Практикум для студентов направления подготовки
«Техносферная безопасность»

Автор-составитель
КИНДЕЕВ Евгений Александрович

Редактор О. В. Балашова
Технический редактор Ш. Ш. Амирсейидов
Компьютерная верстка Л. В. Макаровой
Корректор Н. В. Пустовойтова
Выпускающий редактор А. А. Амирсейидова

Системные требования: Intel от 1,3 ГГц; Windows XP/7/8/10; Adobe Reader; дисковод CD-ROM.

Тираж 9 экз.

Издательство Владимирского государственного университета
имени Александра Григорьевича и Николая Григорьевича Столетовых.
600000, Владимир, ул. Горького, 87.