СБОРНИК ЗАДАЧ ПО МАТЕМАТИКЕ

ДЛЯ ВТУЗОВ

4

Под редакцией А. В. Ефимова и А. С. Поспелова

Москва
Издательство
Физико-математической литературы
2003

ББК 22.171
С 23
УДК 51(075.8)

Коллектив авторов:
Э. А. ВУКОЛОВ, А. В. ЕФИМОВ, В. Н. ЗЕМСКОВ,
А. С. ПОСПЕЛОВ

Сборник задач по математике для втузов. В 4 частях. Ч. 4: Учебное пособие для втузов / Под общ. ред. А. В. Ефимова и А. С. Поспелова. — 3-е изд. перераб. и доп. —М.: Издательство Физико-математической литературы, 2003.—432 с.—ISBN 5-94052-037-5 (Ч. 4).

Содержит задачи по специальным курсам математики: теории вероятностей и математической статистике. Во всех разделах приводятся необходимые теоретические сведения. Все задачи снабжены ответами, а наиболее сложные — решениями. Решение части задач предполагает использование ЭВМ. Для студентов высших технических учебных заведений.

Учебное издание

ВУКОЛОВ Эдуард Александрович, ЕФИМОВ Александр Васильевич, ЗЕМСКОВ Владимир Николаевич, ПОСПЕЛОВ Алексей Сергеевич

СБОРНИК ЗАДАЧ ПО МАТЕМАТИКЕ ДЛЯ ВТУЗОВ

Часть 4

Редактор *Е. А. Привалов*
Компьютерная графика *М. Н. Грицук*
Компьютерный верстка *Е. С. Нехаева*

ИД № 01389 от 30.03.2000
Гигиеническое заключение № 77.99.02.953.Д.003724.07.01
от 05.07.2001

Подписано в печать 20.01.2003. Формат 60×84/16.
Печать офсетная с готовых диапозитивов.
Усл.-печ. л. 25,2. Уч.-изд. л. 30,5. Тираж 4 000 экз.
Заказ № 1315.

Издательство Физико-математической литературы
119071 Москва В-71, Ленинский проспект, 15

Отпечатано в полном соответствии с качеством
предоставленных диапозитивов на ФГУИПП «Вятка».
610033, г. Киров, ул. Московская, 122.

ISBN 5-94052-037-5 (Ч. 4)
ISBN 5-94052-033-2

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ ТИТУЛЬНЫХ РЕДАКТОРОВ

Настоящее издание «Сборника задач по математике для втузов» подверглось значительной перестановке глав и их распределению по томам. В результате первый том содержит алгебраические разделы курса высшей математики, в том числе векторную алгебру и аналитическую геометрию, определители и матрицы, системы линейных уравнений, линейную алгебру и новый раздел — общую алгебру.

Второй том полностью посвящен изложению основ математического анализа, дифференциальному и интегральному исчислениям функций одной и нескольких переменных, а также дифференциальным уравнениям.

В третьем томе собраны специальные разделы математического анализа, которые в различных наборах и объемах изучаются в технических вузах и университетах. Сюда относятся такие разделы, как векторный анализ, элементы теории функций комплексной переменной, ряды и их применение, операционное исчисление, методы оптимизации, уравнения в частных производных, а также интегральные уравнения.

Наконец, четвертый том содержит теоретические введения, типовые примеры и циклы задач по теории вероятностей и математической статистике.

Указанные выше изменения составляют лишь структурную переработку Сборника, никоим образом не затрагивая ни расположения материала внутри соответствующей главы, ни последовательности нумерации примеров и задач.

В смысловом отношении авторы внесли только следующие изменения. Во всех разделах Сборника исключены теоретические введения и циклы задач, связанные с численными методами. Дело в том, что в настоящее время существует целый ряд программных оболочек, каждая из которых реализует достаточно полный набор стандартных методов приближенного решения задач, а основные навыки работы с компьютером можно приобрести уже в школе. Авторы посчитали также необходимым добавить один новый раздел «Основы общей алгебры» и предложить цикл задач по тензорной алгебре в разделе «Линейная алгебра» в первый, «алгебраический» том Сборника. Это связано с тем, что круг идей и методов общей алгебры все глубже проникает в наукоемкие отрасли промышленности и, следовательно, становится необходимой частью образования и подготовки специалистов по инженерным специальностям.

Кроме отмеченного выше, авторами выполнена стандартная техническая работа по исправлению ошибок, описок и других неточностей, учтены также все замечания, возникавшие в процессе работы с предыдущими изданиями Сборника.

А. В. Ефимов, А. С. Поспелов

Глава 18

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

§ 1. Случайные события

1. Понятие случайного события. Предметом теории вероятностей являются модели *экспериментов со случайными исходами* (*случайных экспериментов*). При этом рассматриваются только такие эксперименты, которые можно повторять (воспроизводить) при неизменном комплексе условий произвольное число раз (по крайней мере теоретически). Для реально воспроизводимого эксперимента понятие «наблюдаемый результат» означает, что существует принципиальная возможность зарегистрировать данный результат опыта с помощью того или иного прибора (в простейшем случае, например, визуально). Любой наблюдаемый результат интерпретируется как случайный исход опыта (*случайное событие*).

Математическая формализация модели случайного эксперимента включает в себя:

1) построение множества элементарных исходов Ω,

2) описание поля событий для данного эксперимента,

3) задание вероятностного распределения на поле событий.

Под *множеством элементарных исходов* понимают множество взаимоисключающих исходов такое, что результатом эксперимента всегда является один и только один исход. Любое подмножество данного множества Ω интерпретируется как событие (возможно, и ненаблюдаемое). Совокупность всех наблюдаемых событий составляет *поле событий* для данного эксперимента.

Понятия, связанные с 2) и 3), определяются строго в аксиоматической теории вероятностей, которая составляет основу всех современных курсов теории вероятностей и их приложений.

Множество Ω для данного эксперимента может быть *дискретным*, *непрерывным* или иметь более сложную структуру. К дискретным относятся конечные или счетные множества элементарных исходов, к непрерывным — множества типа континуума (любой конечный или бесконечный интервал на числовой прямой является примером множества типа континуума). В дальнейшем мы рассматриваем только такие модели экспериментов, для которых множество элементарных исходов Ω либо дискретно, либо непрерывно.

Говорят, что событие A *произошло* (*наступило, осуществилось, реализовалось*), если результатом эксперимента явился элементарный исход ω, принадлежащий A ($\omega \in A$). Событие, совпадающее с пустым множеством $\varnothing$, называется *невозможным событием*, а событие, совпадающее со всем множеством Ω, — *достоверным событием.*

Два события A и B называются *совместными* (*несовместными*), если в результате эксперимента возможно (невозможно) их совместное осуществление. Другими словами, события A и B совместны, если соответствующие множества A и B имеют общие элементы, и несовместны в противном случае.

Построение множества Ω (если оно не задано при описании эксперимента) осуществляется на практике, исходя из требования, чтобы все интересующие нас результаты данного эксперимента могли быть однозначно описаны на основе построенного множества Ω. Другими словами, если нас интересуют события A, B, C и т. д., являющиеся наблюдаемыми событиями в данном эксперименте, то множество Ω должно состоять из таких исходов, чтобы существовали подмножества данного множества, равносильные событиям A, B, C и т.д.

Пример 1. Эксперимент состоит в подбрасывании один раз правильной шестигранной игральной кости. Обозначим X число очков, выпавших на верхней грани кости. Описать множество элементарных исходов Ω и указать состав подмножеств, соответствующих следующим событиям: $A = \{X \text{ кратно трем}\}$, $B = \{X \text{ нечетно}\}$, $C = \{X > 3\}$, $D = \{X < 7\}$, $E = \{X \text{ дробно}\}$, $F = \{0,5 < X < 1,5\}$. Выявить пары совместных событий.

◁ Введем обозначения для следующих наблюдаемых в данном эксперименте событий: $\omega_k = \{X = k\}$, $k = 1, 2, \ldots, 6$; $\omega^{(1)} = \{X \text{ — нечетное число}\}$, $\omega^{(2)} = \{X \text{ — четное число}\}$.

На базе данных исходов можно сконструировать два множества элементарных исходов $\Omega_1 = \{\omega_1, \omega_2, \ldots, \omega_6\}$ и $\Omega_2 = \{\omega^{(1)}, \omega^{(2)}\}$. Какое из них больше подходит для формальной модели данного эксперимента? Ясно, что Ω_2 следует «забраковать», поскольку, например, наблюдаемые события $\omega_1, \omega_2, \ldots, \omega_6$ A, B, D, E не являются подмножествами множества Ω_2. С другой стороны, все перечисленные события могут быть описаны как подмножества множества Ω_1. Действительно, $A = \{\omega_3, \omega_6\}$, $B = \omega^{(1)} = \{\omega_1, \omega_3, \omega_5\}$, $C = \{\omega_4, \omega_5, \omega_6\}$, $D = \{\omega_1, \omega_2, \ldots, \omega_6\} = \Omega_1$, $E = \varnothing$, $F = \omega_1$. Из написанных равенств, в частности, усматриваем, что исходы $\omega^{(1)}$ и $\omega^{(2)}$ разложимы на элементы, которые сами являются исходами данного эксперимента. Таким образом, исходы $\omega_1, \omega_2, \ldots, \omega_6$ более «элементарны», чем исходы $\omega^{(1)}$ и $\omega^{(2)}$.

Сопоставляя попарно события и проверяя наличие общих элементов, находим пары совместных событий: A и B, A и C, A и D, B и C, B и D, B и F, C и D, D и F. ▷

Пример 2. Эксперимент состоит в радиолокационном обнаружении воздушной цели. Наблюдаемый результат — положение светящегося пятна (отраженного импульса от цели) на экране индикатора цели, имеющего форму круга радиуса 10 см, в декартовой системе координат с началом, совпадающим с центром экрана. Описать множество элементарных исходов и состав подмножеств, соответствующих следующим событиям: $A =$ {цель находится в первом квадранте}, $B =$ {цель находится в круге радиуса 5 см, центр которого совпадает с центром экрана}, $C =$ {цель находится в круге радиуса 2,5 см, центр которого сдвинут на 5 см вдоль оси Ox в отрицательном направлении}. Совместны ли пары событий A и B, A и C, B и C?

◁ Все интересующие нас в данном эксперименте наблюдаемые события связаны с регистрацией положения светящегося пятна на экране индикатора. Удобной формой математического описания элементарного исхода являются в данном случае координаты случайной точки на плоскости, соответствующей, например, центру пятна (предполагается, что пятно представляет собой круг достаточно малого радиуса).

Таким образом, множество Ω непрерывно и может быть записано в виде

$$\Omega = \{(x,y) \mid x^2 + y^2 \leqslant 100\}.$$

Подмножества, равносильные указанным событиям, имеют вид

$$A = \{(x,\, y) \mid x^2 + y^2 \leqslant 100,\ x > 0,\ y > 0\},$$
$$B = \{(x,\, y) \mid x^2 + y^2 \leqslant 25\},$$
$$C = \{(x,\, y) \mid (x+5)^2 + y^2 \leqslant 6{,}25\}.$$

По определению, события совместны, если соответствующие им подмножества имеют общие элементы (пересекаются), и несовместны в противном случае. Поэтому события A и B, B и C совместны, а события A и C несовместны. ▷

В задачах 18.1–18.10 построить множество элементарных исходов Ω по описанию эксперимента и подмножества, соответствующие указанным событиям.

18.1. Игральная кость подбрасывается дважды. Наблюдаемый результат — пара чисел, соответствующих числам очков, выпавших в первый и второй раз. События: $A =$ {оба раза выпало число очков, кратное трем}, $B =$ {ни разу не выпало число шесть}, $C =$ {оба раза выпало число очков, большее трех}, $D =$ {оба раза выпало одинаковое число очков}.

18.2. Монета подбрасывается три раза. Наблюдаемый результат — появление герба (г) или цифры (ц) на верхней стороне монеты. События: $A =${герб выпал ровно один раз}, $B =$ {ни разу не выпала цифра}, $C =$ {выпало больше гербов, чем цифр}, $D =$ {герб выпал не менее, чем два раза подряд}.

18.3. Монета подбрасывается до первого появления герба. Наблюдаемый результат — общее число подбрасываний. События: $A =$ {герб выпал при третьем подбрасывании}, $B =$ {герб выпал не ранее, чем при третьем подбрасывании}.

18.4. Эксперимент состоит в раскладывании наудачу трех занумерованных шаров по трем ящикам. В каждый ящик может поместиться любое число шаров. Наблюдаемый результат — тройка чисел (i, j, k), где i, j, k — номера ящиков, в которые попали соответственно первый, второй и третий шары. События: $A =$ $= \{$первый ящик пустой$\}$, $B = \{$в каждый ящик попало по одному шару$\}$, $C = \{$все шары попали в один ящик$\}$.

18.5. Производится стрельба по плоской прямоугольной мишени: $-2 \leqslant x \leqslant 2$, $-1 \leqslant y \leqslant 1$. Наблюдаемый результат — координаты точки попадания в декартовой системе координат. По условиям стрельбы непопадание в указанный прямоугольник исключено. События: $A = \{$абсцисса точки попадания не меньше ординаты$\}$, $B = \{$произведение координат точки неотрицательно$\}$, $C = \{$сумма абсолютных величин координат точки превышает единицу$\}$. Выявить пары совместных событий.

18.6. На отрезке $[a, b]$ наудачу ставится точка. Пусть x — координата этой точки. Затем на отрезке $[a, x]$ наудачу ставится еще одна точка с координатой y. Наблюдаемый результат — пара чисел (x, y). События: $A = \{$вторая точка ближе к правому концу отрезка $[a, b]$, чем к левому$\}$, $B = \{$расстояние между двумя точками меньше половины длины отрезка$\}$, $C = \{$первая точка ближе ко второй, чем к правому концу отрезка $[a, b]\}$. Выявить пары несовместных событий.

18.7. Иван и Петр договорились о встрече в определенном месте между одиннадцатью и двенадцатью часами. Каждый приходит в случайный момент указанного промежутка и ждет появления другого до истечения часа, но не более 15 минут, после чего уходит. Наблюдаемый результат — пара чисел (x, y), где x — время прихода Петра, y — время прихода Ивана (время исчисляется в минутах, начиная от 11 часов). Событие $A = \{$встреча состоялась$\}$.

18.8 (продолжение). В условиях эксперимента задачи 18.7 рассмотреть следующие события: $B = \{$Петр ждал Ивана все обусловленное время и не дождался$\}$, $C = \{$Ивану не пришлось ждать Петра$\}$.

18.9 (продолжение). В условиях эксперимента задачи 18.7 рассмотреть события: $D = \{$встреча состоялась после 11 ч 30 мин$\}$, $E = \{$Иван опоздал на встречу$\}$, $F = \{$встреча состоялась, когда до истечения часа оставалось меньше пяти минут$\}$.

18.10*. Проводится матч на первенство страны по футболу между командами «Динамо» и «Спартак». Интересующие нас события: $A = \{$выиграла команда «Динамо»$\}$, $B = \{$игра закончилась победой одной из команд$\}$, $C = \{$игра закончилась со счетом 3 : 1 в пользу «Спартака»$\}$, $D = \{$в игре забито не меньше трех голов$\}$.

18.11*. С помощью специального прибора регистрируется направление φ и скорость ветра v в данном месте Земли. Прибор устроен таким образом, что позволяет определять скорость ветра сколь угодно точно, а регистрация направления ветра возможна лишь с точностью до 2°. Установить, наблюдаемы ли в данном эксперименте события: $A = \{(v, \varphi) \,|\, v < 12$ км/ч, $\varphi = 343^\circ 35'\}$, $B = \{(v, \varphi) \,|\, v = 15{,}5$ км/ч, $340^\circ \leqslant \varphi < 350^\circ\}$, $C = \{(v, \varphi) \,|\, v \geqslant 3{,}25$ км/ч, $48^\circ \leqslant \varphi \leqslant 51^\circ\}$.

2. Алгебраические операции над событиями. Поскольку событие отождествляется с множеством, то над событиями можно совершать все операции, выполнимые над множествами. В частности, определены следующие операции и отношения между событиями:

$A \subset B$ (отношение включения множеств: множество A является подмножеством множества B)— *событие A влечет за собой событие B.* Иначе говоря, событие B происходит всякий раз, как происходит событие A.

$A = B$ (отношение эквивалентности множеств) — *событие A тождественно событию B.* Это возможно в том и только в том случае, когда $A \subset B$ и одновременно $B \subset A$.

$A + B$ (объединение множеств) — *сумма событий.* Это событие, состоящее в том, что произошло хотя бы одно из двух событий A или B (не исключающее логическое «или»).

AB (пересечение множеств) — *произведение событий.* Это событие, состоящее в совместном осуществлении событий A и B (логическое «и»). Таким образом, события A и B несовместны, если $AB = \varnothing$.

$A - B$ (множество элементов, принадлежащих A, но не принадлежащих B)— *разность событий.* Это событие, состоящее в том, что A происходит, а B не происходит.

$\overline{A} = \Omega - A$ (дополнение множества A до Ω)— *противоположное событие.* Это событие, состоящее в том, что A не происходит (логическое отрицание).

Система $\mathcal{F}$ подмножеств множества Ω такая, что в результате применения любой из описанных операций к любым двум элементам системы снова получается элемент данной системы, называется *алгеброй* (или *булевой алгеброй* — по имени английского математика Дж. Буля (1815 - 1864)). Под *наблюдаемым событием* понимается такое подмножество множества Ω, которое одновременно принадлежит и булевой алгебре $\mathcal{F}$. Таким образом, класс наблюдаемых в данном эксперименте событий, вообще говоря, у́же класса всех подмножеств множества Ω. Если, например, $R \subset \Omega$, но $R \notin \mathcal{F}$, то событие R по определению не наблюдаемо в данном эксперименте. Такое определение наблюдаемого события согласуется с введенным ранее эмпирическим понятием случайного события как наблюдаемого результата эксперимента.

Для экспериментов с конечным числом исходов множество всех подмножеств Ω, включающее пустое множество $\varnothing$, составляет алгебру. Поэтому для таких экспериментов любое подмножество множества Ω может интерпретироваться как наблюдаемое событие. В тех случаях, когда Ω — счетное или непрерывное множество, необходимо потребовать, чтобы поле событий, состоящее из бесконечного числа наблюдаемых событий, было замкнуто относительно алгебраических операций над счетным числом событий. Система подмножеств множества Ω, удовлетворяющая этому условию, называется σ-алгеброй, а соответствующее поле событий — *борелевским полем событий.*

Пример 3. Игральная кость подбрасывается один раз. Наблюдаемый результат — число очков на верхней грани. События A, B, C, D, E, F описаны в примере 1. Описать состав и выяснить смысл следующих событий: $E_1 = \overline{B}$, $E_2 = \overline{C}$, $E_3 = AB$, $E_4 = A + B$, $E_5 = A - \overline{B}$, $E_6 = E + D$, $E_7 = EF$.

◁ В обозначениях примера 1 напишем состав указанных событий, используя определение соответствующей алгебраической операции: $E_1 = \{\omega_2, \omega_4, \omega_6\}$ — выпало четное число очков; $E_2 = \{\omega_1, \omega_2, \omega_3\}$ — выпало число очков, не большее трех; $E_3 = \{\omega_3\}$ — выпавшее число очков нечетно и кратно трем; $E_4 = \{\omega_1, \omega_3, \omega_5, \omega_6\}$ — выпавшее число очков или нечетно, или кратно трем; $E_5 = AB = E_3$; $E_6 = \varnothing + D = D = \Omega$. ▷

Пример 4. Рассмотрим снова случайный эксперимент, описанный в примере 2. Составляет ли множество всех квадрируемых подмножеств множества Ω алгебру событий?

◁ Будем говорить, что область $S \subset \Omega$ квадрируема, если она имеет площадь, понимаемую в смысле меры Лебега. Из свойств меры Лебега вытекает, что объединение, пересечение и дополнение конечного или счетного числа квадрируемых подмножеств некоторого квадрируемого множества на плоскости являются квадрируемыми множествами. Отсюда следует, что система $\mathcal{F}$ квадрируемых подмножеств множества Ω образует σ-алгебру для данного эксперимента [1]). ▷

При доказательстве некоторых утверждений в алгебре событий очень полезной оказывается геометрическая иллюстрация событий, трактуемых как попадание точки в область, соответствующую этому событию. Условно изображая события в виде различных областей на плоскости, получаем так называемые *диаграммы Венна.*

Пример 5. Доказать, что операции сложения и умножения событий обладают следующими свойствами:

а) $A + B = B + A$, $AB = BA$ (*коммутативность*);

б) $(A+B)+C = A+(B+C)$, $A\,(BC) = (AB)\,C$ (*ассоциативность*);

в) $(A + B)\,C = AC + BC$ (*дистрибутивность умножения относительно сложения*).

◁ Свойства коммутативности а) и ассоциативности б) непосредственно вытекают из определения операций сложения и умножения, поскольку события «произошло хотя бы одно из ...» и «произошли все вместе ...» не зависят от порядка участвующих событий. Докажем, что событие $(A + B)\,C$ тождественно событию $AC + BC$. Пусть $\omega \in (A + B)\,C$. Это

[1]) Если квадрируемость области на плоскости понимать в смысле меры Жордана, как это обычно делается в курсе математического анализа, то система квадрируемых по Жордану подмножеств множества Ω может не составить σ-алгебру. Однако, если рассматриваются такие простейшие области на плоскости, для которых обе меры совпадают, то использование для определения площадей этих областей интеграла Римана вместо интеграла Лебега не приводит к ошибкам.

значит, что $\omega \in C$ и принадлежит по крайней мере одному из событий A или B. Но тогда ω принадлежит хотя бы одному из событий AC или BC, т. е. $\omega \in AC + BC$. Наоборот, пусть $\omega \in AC + BC$. Тогда $\omega \in AC$ или $\omega \in BC$. Следовательно, $\omega \in C$ и, кроме того, принадлежит по крайней мере одному из событий A или B, т.е. $\omega \in (A + B)\,C$. Тем самым тождественность левой и правой частей доказана.

Проиллюстрируем свойство в) диаграммами Венна. Пусть события означают: $A = \{$попадание в квадрат$\}$, $B = \{$попадание в круг$\}$, $C =$ $= \{$попадание в треугольник$\}$. Соответствующие области изображены на рис. 1. Горизонтальной штриховкой отмечена область, соответствующая событию AC, вертикальной — событию BC, косая штриховка соответствует событию $AC + BC$. $\triangleright$

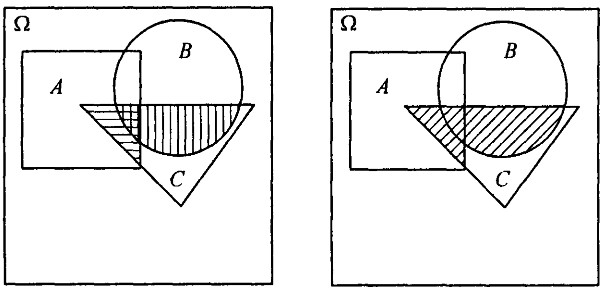


Рис. 1

В задачах 18.12–18.17 доказать справедливость следующих тождеств:

18.12. $A + A = A$, $AA = A$, $A + \varnothing = A$, $A\varnothing = \varnothing$, $A\Omega = A$, $A + \Omega = \Omega$.

18.13. $A + \overline{A} = \Omega$, $\overline{\Omega} = \varnothing$, $\overline{\varnothing} = \Omega$, $A\overline{A} = \varnothing$.

18.14. а) $\overline{A + B} = \overline{A}\,\overline{B}$, $\overline{AB} = \overline{A} + \overline{B}$ (правила де Моргана). б) Обобщить правила де Моргана на произвольное число n событий.

18.15.** $AB + C = (A + C)(B + C)$ (дистрибутивность сложения относительно умножения).

18.16. $A - B = A\overline{B}$.

18.17*. $(A + B) - B = A - AB = A\overline{B} = A - B$.

Замечание. Этот пример показывает, что «приведение подобных членов» в алгебре событий недопустимо.

18.18. Пусть A, B и C — события, наблюдаемые в эксперименте, причем A и B несовместны. Показать, что события AC и BC также несовместны.

18.19. Показать, что:

а) если $A \subset B$, то выполняются соотношения

$$AB = A, \quad A + B = B; \tag{1}$$

б) из справедливости любого из соотношений (1) следует

$$A \subset B.$$

18.20. Пусть A и B — наблюдаемые события в эксперименте. Показать, что событие $A + B$ можно разложить на сумму несовместных событий следующими способами: а) $A + B = A + (B - AB)$; б) $A + B = AB + A\overline{B} + \overline{A}B$; в) $A + B = A + B\overline{A}$.

18.21*. Показать, что если $A \subset B$, то $\overline{B} \subset \overline{A}$.

18.22. Показать, что если $B \subset A$, то $(A - B) + B = A$.

Доказать тождества:

18.23. $(A + B)(A + \overline{B}) = A$.

18.24. $(A + B)(\overline{A} + B)(A + \overline{B}) = AB$.

18.25. $(\overline{A} + BC)(\overline{B} + AC)(\overline{C} + AB) = ABC + \overline{A}\,\overline{B}\,\overline{C}$.

18.26*. $AC - B = AC - BC$.

18.27*. $(A - B) + (A - C) = A - BC$.

Симметрическая разность двух событий $A \triangle B$ определяется следующим образом:

$$A \triangle B = (A - B) + (B - A).$$

Доказать следующие тождества:

18.28. $A \triangle B = (A + B) - AB$.

18.29. $\overline{A \triangle B} = AB + \overline{A}\,\overline{B}$.

18.30. $A \triangle B = (\overline{A\overline{B}}) \triangle (\overline{\overline{A}B})$.

18.31. Пусть $C = A \triangle B$. Доказать, что $A \triangle C = B$.

18.32. Найти случайное событие X из равенства

$$\overline{X + A} + \overline{X + \overline{A}} = B.$$

18.33.** Доказать, что $A - B = \varnothing$ тогда и только тогда, когда $A \subset B$.

18.34*. Очередной посетитель входит в зал музея, где уже собралось $2n$ человек, и начинает отыскивать знакомых среди собравшихся. Интересующие нас события: $A = \{$среди собравшихся найдется n человек, знакомых посетителю$\}$, $B = \{$среди собравшихся найдется n человек, не знакомых посетителю$\}$. Доказать, что события $A + B$ и $\overline{A - B} + \overline{B}$ достоверные.

Пусть A, B, C — три события, наблюдаемые в данном эксперименте. В задачах 18.35–18.37 выразить указанные события в алгебре событий.

18.35. $E_1 = \{$из трех событий A, B, C произойдет ровно одно$\}$, $F_1 = \{$из трех событий A, B, C произойдет ровно два$\}$.

18.36. $E_2 = \{$из трех событий A, B, C произойдет хотя бы одно$\}$, $F_2 = \{$из трех событий A, B, C произойдет не меньше двух$\}$.

18.37. $E_3 = \{$из трех событий A, B, C не произойдет ни одного$\}$, $F_3 = \{$из трех событий A, B, C произойдет хотя бы два$\}$, $G = \{$из трех событий A, B, C не произойдет хотя бы одно$\}$.

18.38. Поражение боевого самолета может наступить или в результате поражения обоих двигателей (события D_1 и D_2), или в результате попадания в кабину пилота (событие K). Производится длительный обстрел самолета из зенитного орудия. Любое попадание в соответствующий агрегат приводит к его поражению. Пусть событие $A = \{$поражение самолета$\}$.

а) Описать множество элементарных исходов.

б) Записать A в алгебре событий как непосредственно с помощью событий D_1, D_2 и K, так и через элементарные исходы.

в)** Получить из второй записи первую путем допустимых алгебраических преобразований.

18.39. Электрическая цепь составлена по схеме, приведенной на рис. 2. Событие $A_k = \{$элемент с номером k вышел из строя$\}$, $k = 1, 2, 3, 4$. Событие $B = \{$разрыв цепи$\}$. Выразить событие B в алгебре событий A_1, A_2, A_3, A_4.

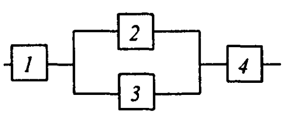


Рис. 2

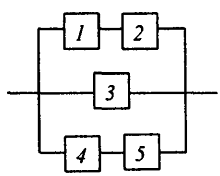


Рис. 3

18.40. Электрическая цепь составлена по схеме, приведенной на рис. 3. Событие $A_k = \{$элемент с номером k вышел из строя$\}$, $k = 1, 2, 3, 4, 5$. Событие $B = \{$разрыв цепи$\}$. Выразить событие B в алгебре событий A_1, A_2, A_3, A_4, A_5.

18.41. На отрезке $[a, b]$ наудачу ставятся две точки. Пусть x и y — координаты этих точек. Изобразить на плоскости Oxy области, соответствующие событиям Ω, A, B, AB, $A - B$, $A + B$, где $A = \{$вторая точка ближе к левому концу отрезка, чем первая точка к правому концу$\}$, $B = \{$расстояние между точками меньше половины длины отрезка$\}$.

Пусть заданы n множеств $\Omega_1, \Omega_2, \ldots, \Omega_n$, содержащих, вообще говоря, различное число элементов. Множество вида

$$\Omega = \{(\omega_1, \omega_2, \ldots, \omega_n,) \,|\, \omega_1 \in \Omega_1,\ \omega_2 \in \Omega_2, \ldots, \omega_n \in \Omega_n\}$$

называется *прямым произведением* множеств $\Omega_1, \ldots, \Omega_n$ и обозначается

$$\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n.$$

Пусть каждое из множеств Ω_i, $i = 1, 2, \ldots$, является множеством элементарных исходов некоторого эксперимента E_i. Тогда составной эксперимент, состоящий в проведении последовательно экспериментов $E_1, E_2, \ldots, E_n$ и наблюдений совместного результата, имеет множество элементарных исходов $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$ и называется *последовательностью испытаний*.

18.42. Произведено три выстрела из орудия по цели. Событие $A_k = \{$попадание при k-м выстреле$\}$ $(k = 1, 2, 3)$.

а) Выяснить состав множества Ω, выразив каждый элементарный исход ω_i через события A_k.

б) Записать в алгебре событий следующие события:

$A = \{$ровно одно попадание$\}$, $B = \{$хотя бы одно попадание$\}$, $C = \{$хотя бы один промах$\}$, $D = \{$не меньше двух попаданий$\}$, $E = \{$попадание не раньше, чем при третьем выстреле$\}$.

18.43. Из ящика, содержащего 10 деталей, из которых 3 бракованных, наудачу последовательно и без возвращения извлекается по одной детали до появления бракованной, после чего опыт прекращается. Обозначим исход i-го испытания $\omega_i = \{$бракованная деталь появится при i-м испытании$\}$. Рассмотрим событие $A = \{$придется производить третье по счету извлечение детали$\}$.

а) Сконструировать элементарные исходы данного опыта с помощью алгебраических операций над исходами ω_i, $i = 1, 2, \ldots$

б) Записать событие A через элементарные исходы и упростить запись путем алгебраических преобразований.

18.44. Два баскетболиста по очереди бросают мяч в корзину до первого попадания. Выигрывает тот, кто первый забросит мяч. События: $A_k = \{$первый баскетболист попадает при своем k-м броске$\}$, $B_k = \{$второй баскетболист попадает при своем k-м броске$\}$; $A = \{$выигрывает первый баскетболист$\}$, $B = \{$выигрывает второй$\}$. Первый баскетболист бросает первым. Определить состав множества элементарных исходов и записать события A и B в алгебре событий.

Система множеств $\{E_1, E_2, \ldots, E_l\}$ называется *разбиением* множества E, если выполняются следующие условия:

$$E_i \neq \varnothing, \qquad i = 1, 2, \ldots, l,$$
$$E_i E_j = \varnothing, \qquad i \neq j,$$
$$E_1 + E_2 + \ldots + E_l = E.$$

Если разбиению подвергается множество элементарных исходов Ω некоторого эксперимента, то говорят, что система подмножеств (событий), осуществляющих разбиение, составляет *полную группу несовместных событий.*

18.45. Показать, что совокупность элементарных исходов любого эксперимента с конечным множеством Ω образует разбиение множества Ω.

18.46. Образуют ли события A, B и C из задачи 18.4 полную группу событий?

18.47. Показать, что система событий $\{D_1 + D_2,\ K\overline{D}_1\overline{D}_2,\ \overline{D_1 + D_2 + K}\}$, где D_1, D_2 и K — наблюдаемые события в эксперименте, описанном в задаче 18.38, образует разбиение множества Ω для данного эксперимента.

18.48. Множество элементарных исходов некоторого эксперимента состоит из четырех исходов. Сколько различных разбиений можно составить для данного множества?

18.49. Пусть A — произвольное наблюдаемое в некотором эксперименте событие такое, что $A \neq \Omega$ и $A \neq \varnothing$. Показать, что система множеств $\{A, \overline{A}\}$ образует разбиение множества Ω.

18.50. Пусть $\Omega = \{1, 2, 3, \ldots\}$ — множество натуральных чисел. Показать, что система $\{S_1, S_2, S_3\}$, где $S_1 = \{x \mid x = 3n;\ n = 1, 2, 3, \ldots\}$, $S_2 = \{x \mid x = 3n-1;\ n = 1, 2, 3, \ldots\}$, $S_3 = \{x \mid x = 3n - 2;\ n = 1, 2, 3, \ldots\}$, образует разбиение множества Ω.

18.51. Для некоторого эксперимента множество Ω содержит ровно n элементарных исходов. Показать, что число всех наблюдаемых событий, содержащихся в поле событий для данного эксперимента, равно 2^n.

3. Аксиоматическое определение вероятности события. Для количественного описания степени объективной возможности наступления того или иного наблюдаемого в эксперименте события вводится специальная числовая функция $\mathrm{P}(A)$, называемая вероятностью события A.

Пусть $\mathcal{F}$ — поле событий для данного эксперимента. *Вероятностью* $\mathrm{P}(A)$ называется числовая функция, определенная для всех $A \in \mathcal{F}$ и удовлетворяющая трем условиям (аксиомам вероятностей):

1) $\mathrm{P}(A) \geqslant 0$.

2) $\mathrm{P}(\Omega) = 1$.

3) Для любой конечной или бесконечной последовательности наблюдаемых событий $A_1, A_2, \ldots, A_n, \ldots$ таких, что $A_iA_j = \varnothing$ при $i \neq j$,

$$\mathrm{P}\Big(\sum_k A_k\Big) = \sum_k \mathrm{P}(A_k).$$

Так как событие есть множество, то вероятность является также функцией множества. Указанные аксиомы выделяют специальный класс числовых функций, являющихся вероятностями. В соответствии со смыслом этих аксиом вероятность есть неотрицательная, нормированная и аддитивная функция множеств, принадлежащих алгебре $\mathcal{F}$.

Тройку $\{\Omega, \mathcal{F}, \mathrm{P}\}$, где $\mathcal{F}$ — алгебра подмножеств множества Ω, P — числовая функция, удовлетворяющая аксиомам 1–3, называют *вероятностным пространством* случайного эксперимента.

Аксиоматическая теория вероятностей в ее современном виде была создана А. Н. Колмогоровым в 1933 г.

Пример 6. Доказать, что если для некоторого эксперимента $A \subset B$, то $\mathrm{P}(A) \leqslant \mathrm{P}(B)$.

$\lhd$ Так как $A \subset B$, то $AB = A$ (см. задачу 18.19, а)), поэтому $B = B(A+\overline{A}) = BA+B\overline{A} = A+B\overline{A}$, причем оба слагаемых несовместны. В силу аксиомы аддитивности $\mathrm{P}(B) = \mathrm{P}(A) + \mathrm{P}(B\overline{A})$, а в силу аксиомы 1) $\mathrm{P}(B\overline{A}) \geqslant 0$. Следовательно, $\mathrm{P}(B) \geqslant \mathrm{P}(A)$. $\rhd$

Доказать справедливость следующих следствий из определения вероятности события:

18.52*. $\mathrm{P}(\varnothing) = 0$.

18.53. $\mathrm{P}(\overline{A}) = 1 - \mathrm{P}(A)$.

18.54*. $\mathrm{P}(A) \leqslant 1$.

18.55.** $\mathrm{P}(A+B) = \mathrm{P}(A)+\mathrm{P}(B)-\mathrm{P}(AB)$ (формула сложения вероятностей).

18.56. $\mathrm{P}(A+B) \leqslant \mathrm{P}(A) + \mathrm{P}(B)$.

18.57. $\mathrm{P}(AB) \leqslant \mathrm{P}(A) \leqslant \mathrm{P}(A+B)$.

18.58. Доказать, что, если $A = B$, то $\mathrm{P}(A) = \mathrm{P}(B)$.

18.59*. Пусть A, B и C — три события из поля событий для данного эксперимента. Показать, что

$$\mathrm{P}((A-B)+(B-C)+(C-A)) = 1 - \mathrm{P}(ABC) - \mathrm{P}(\overline{A}\,\overline{B}\,\overline{C}).$$

18.60*. Пусть A и B — наблюдаемые события в эксперименте. Показать, что

$$\mathrm{P}(A\triangle B) = \mathrm{P}(A) + \mathrm{P}(B) - 2\mathrm{P}(AB) = \mathrm{P}(A+B) - \mathrm{P}(AB).$$

18.61. Доказать, что если $A \supset B$, то

$$\mathrm{P}(A-B) = \mathrm{P}(A) - \mathrm{P}(B).$$

18.62. Доказать, что если A и B — наблюдаемые события в эксперименте, то справедливо равенство

$$\mathrm{P}(A\overline{B}) - \mathrm{P}(\overline{A}B) = \mathrm{P}(A) - \mathrm{P}(B).$$

18.63. Показать, что для любых наблюдаемых в эксперименте событий A, B и C справедливо неравенство

$$\mathrm{P}(AB+BC+AC) \leqslant \frac{2}{3}(\mathrm{P}(A)+\mathrm{P}(B)+\mathrm{P}(C)).$$

18.64*. Показать, что для трех наблюдаемых в эксперименте событий A, B и C справедлива следующая формула сложения вероятностей:

$$P(A+B+C) =$$
$$= P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC).$$

18.65.** Известно, что для данного эксперимента совместное наступление событий A и B с необходимостью влечет за собой наступление события C. Доказать, что

$$P(C) \geqslant P(A) + P(B) - 1.$$

4. Классическая вероятностная схема — схема урн. Во многих случаях вероятностное пространство строится на основе проведения аналогии между описываемым экспериментом и какой-либо хорошо изученной моделью случайного эксперимента с известным распределением вероятностей. Таковы, например, опыты, сводящиеся к классической или геометрической схеме, которые подробно рассматриваются далее.

Всякий эксперимент, удовлетворяющий тому условию, что соответствующее ему множество Ω представляет собой конечное множество равновероятных исходов $\left(\text{т.е. } P(\omega_1) = P(\omega_2) = \dots = P(\omega_n) = \dfrac{1}{n}\right)$, называется *классической схемой* или *схемой урн.* В силу конечности Ω алгебра событий $\mathcal{F}$ совпадает с множеством всех подмножеств множества Ω (включая и пустое множество). Поэтому любое событие вида $A = \{\omega_{k_1}, \omega_{k_2}, \dots, \omega_{k_m}\} \subset \Omega$ наблюдаемо в таком эксперименте, и вероятность его осуществления определяется по *формуле классической вероятности*

$$P(A) = \frac{N(A)}{N(\Omega)} = \frac{m}{n},$$

где $N(A) = m$ — число элементов множества A (число всех благоприятствующих событию A исходов), $N(\Omega) = n$ — число элементов множества Ω (число всех исходов эксперимента).

Классическая схема является математической формализацией опытов, в которых элементарные исходы обладают определенной симметрией по отношению к условиям опыта, так что нет оснований считать какой-либо из исходов более вероятным, чем другой. Таким свойством, например, обладают опыты по извлечению наудачу определенного числа шаров из урны, содержащей заданное количество неразличимых на ощупь шаров. (Отсюда и название — схема урн.)

18.66. В магазин поступило 30 новых цветных телевизоров, среди которых 5 имеют скрытые дефекты. Наудачу отбирается один телевизор для проверки. Какова вероятность, что он не имеет скрытых дефектов?

18.67. Автомат изготавливает однотипные детали, причем технология изготовления такова, что 5 % произведенной продукции оказывается бракованной. Из большой партии взята наудачу одна деталь для контроля. Найти вероятность события $A = \{$деталь бракованная$\}$.

18.68. Игральная кость подбрасывается один раз. Найти вероятности следующих событий: $A = \{$число очков равно 6$\}$, $B = \{$число очков кратно трем$\}$, $C = \{$число очков четно$\}$, $D = \{$число очков меньше пяти$\}$, $E = \{$число очков больше двух$\}$.

Подбрасываются две игральные кости. В задачах 18.69–18.71 найти вероятности указанных событий.

18.69. $A = \{$числа очков на обеих костях совпадают$\}$, $B = \{$число очков на первой кости больше, чем на второй$\}$.

18.70. $C = \{$сумма очков четна$\}$, $D = \{$сумма очков больше двух$\}$.

18.71. $E = \{$сумма очков не меньше пяти$\}$, $F = \{$хотя бы на одной кости появится цифра 6$\}$, $G = \{$произведение выпавших очков равно 6$\}$.

Подсчет числа элементов тех или иных подмножеств множества Ω часто облегчается благодаря следующей формуле. Число элементов прямого произведения множеств равно произведению числа элементов составляющих множеств, т.е.

$$N(\Omega_1 \times \Omega_2 \times \cdots \times \Omega_s) = N(\Omega_1)\, N(\Omega_2) \, \ldots \, N(\Omega_s).$$

18.72. Наудачу выбирается пятизначное число. Какова вероятность следующих событий: $A = \{$число одинаково читается как слева направо, так и справа налево (как, например, 13531)$\}$, $B = \{$число кратно пяти$\}$, $C = \{$число состоит из нечетных цифр$\}$.

18.73. 1 сентября на первом курсе одного из факультетов запланировано по расписанию три лекции по разным предметам. Всего на первом курсе изучается 10 предметов. Студент, не успевший ознакомиться с расписанием, пытается его угадать. Какова вероятность успеха в данном эксперименте, если считать, что любое расписание из трех предметов равновозможно?

18.74. Зенитная батарея, состоящая из n орудий, производит залп по группе, состоящей из m самолетов. Каждое из орудий выбирает себе цель наудачу независимо от остальных. Найти вероятность того, что все орудия выстрелят по одному самолету.

18.75. Пяти полевым радиостанциям разрешено во время учений работать на шести радиоволнах. Выбор волны на каждой станции производится наудачу. Найти вероятности следующих событий: A = {при одновременной работе всех пяти радиостанций хотя бы две волны не совпадут}; B = {будут использованы различные радиоволны}.

18.76. На шахматную доску случайным образом ставят две ладьи — белую и черную. Какова вероятность того, что ладьи не побьют друг друга?

18.77. Каждое из 8 вычислительных устройств обслуживается одним оператором. В штатном составе вычислительного центра имеется 6 операторов. Назначение оператора на данное вычислительное устройство производится наудачу. Найти вероятность того, что первые шесть вычислительных устройств будут обслужены.

5. Комбинаторный метод вычисления вероятностей в классической схеме. Решение вероятностных задач на классическую схему часто облегчается использованием комбинаторных формул. Каждая из комбинаторных формул определяет общее число элементарных исходов в некотором опыте, состоящем в выборе наудачу m элементов из n различных элементов исходного множества $E = \{e_1, e_2, \ldots, e_n\}$. При этом в постановке каждого такого опыта строго оговорено, каким способом производится выбор и что понимается под различными выборками.

Существуют две принципиально различные схемы выбора. В первой схеме выбор осуществляется без возвращения элементов (это значит, что отбираются либо сразу все m элементов, либо последовательно по одному элементу, причем каждый отобранный элемент исключается из исходного множества). Во второй схеме выбор осуществляется поэлементно с обязательным возвращением отобранного элемента на каждом шаге и тщательным перемешиванием исходного множества перед следующим выбором. После того как выбор тем или иным способом осуществлен, отобранные элементы (или их номера) могут быть либо упорядочены (т.е. выложены в последовательную цепочку), либо нет. В результате получаются следующие четыре различные постановки эксперимента по выбору наудачу m элементов из общего числа n различных элементов множества E.

Схема выбора, приводящая к сочетаниям. Если опыт состоит в выборе m элементов без возвращения и без упорядочивания, то различными исходами следует считать m-элементные подмножества множества E, имеющие различный состав. Получаемые при этом комбинации элементов (элементарные исходы) носят название *сочетания из n элементов по m*, а их общее число $N(\Omega)$ определяется по формуле

$$C_n^m = \frac{n!}{m!\,(n-m)!} = \frac{n(n-1)\ldots(n-m+1)}{m!}. \tag{2}$$

Для чисел C_n^m, называемых также *биномиальными коэффициентами*, справедливы следующие тождества, часто оказывающиеся полезными при решении задач:

$C_n^m = C_n^{n-m}$ (свойство симметрии),

$C_{n+1}^k = C_n^k + C_n^{k-1}$, $C_n^0 = 1$ (рекуррентное соотношение),

$C_n^0 + C_n^1 + \dots + C_n^n = 2^n$ (следствие биномиальной формулы Ньютона).

Пример 7. Из партии, содержащей 10 изделий, среди которых 3 бракованных, наудачу извлекают три изделия для контроля. Найти вероятности следующих событий: $A = \{$в полученной выборке ровно одно изделие бракованное$\}$, $B = \{$в полученной выборке нет ни одного бракованного изделия$\}$.

$\lhd$ Занумеруем изделия числами от 1 до 10, и пусть множество номеров $E_1 = \{1, 2, \ldots, 7\}$ соответствует годным изделиям, а множество номеров $E_2 = \{8, 9, 10\}$ — бракованным изделиям.

Согласно описанию эксперимента производится выбор без возвращения и без упорядочивания трех элементов из множества $E = E_1 \cup E_2 = \{1, 2, \ldots, 10\}$. Поэтому $N(\Omega) = C_{10}^3 = 120$.

Событию A благоприятствуют только такие исходы, когда один элемент выборки принадлежит E_2, а остальные два элемента — множеству E_1. По формуле прямого произведения множеств получаем, что число всех таких исходов $N(A) = C_3^1 \cdot C_7^2 = 63$, поэтому

$$\mathrm{P}(A) = \frac{N(A)}{N(\Omega)} = \frac{21}{40}.$$

Событию B благоприятствуют только такие исходы, когда все три отобранных элемента принадлежат множеству E_1, поэтому $N(B) = C_7^3 = 35$. Отсюда следует, что

$$\mathrm{P}(B) = \frac{N(B)}{N(\Omega)} = \frac{7}{24}. \ \rhd$$

18.78. Из полного набора домино (28 штук) наудачу выбирают 7 костей. Какова вероятность, что среди них окажется по крайней мере одна кость с шестью очками?

18.79. Из десяти первых букв русского алфавита наудачу составляется новый алфавит, состоящий из пяти букв. Найти вероятности следующих событий: $A = \{$в состав нового алфавита входит буква *а*$\}$, $B = \{$в состав нового алфавита входят только согласные буквы$\}$.

18.80. Среди кандидатов в студенческий совет факультета 3 первокурсника, 5 второкурсников и 7 третьекурсников. Из этого состава наудачу выбирают пять человек на предстоящую конференцию. Найти вероятности следующих событий: $A = \{$будут выбраны одни третьекурсники$\}$, $B = \{$все первокурсники попадут на конференцию$\}$, $C = \{$не будет выбрано ни одного второкурсника$\}$.

18.81. Из урны, содержащей m_1+m_2 шаров, из которых m_1 белых и m_2 черных, наудачу отбирают m шаров ($m \leqslant \min(m_1, m_2)$) и откладывают в сторону. Найти вероятности следующих событий: $A = \{$все отложенные шары белые$\}$, $B = \{$среди отложенных шаров ровно k белых; $k \leqslant m\}$.

18.82 (продолжение). В условиях предыдущей задачи найти вероятности событий: $C = \{$вынут хотя бы один белый шар$\}$, $D = \{$вынуто не менее k белых шаров; $k \leqslant m\}$.

18.83. Для уменьшения числа игр $2n$ футбольных команд, среди которых 2 призера предыдущего чемпионата, путем жеребьевки разбиваются на две подгруппы (первую и вторую) по n команд в каждой. Какова вероятность q_n того, что обе команды-призеры попадут в разные группы?

18.84. Из колоды в 52 карты извлекаются наудачу 4 карты. Найти вероятности следующих событий: $A = \{$в полученной выборке все карты бубновой масти$\}$, $B = \{$окажется хотя бы один туз$\}$, $C = \{$появятся ровно 2 пики$\}$.

18.85. Из урны, содержащей m_1 шаров с номером 1, m_2 шаров с номером 2, …, m_s шаров с номером s, наудачу без возвращения извлекается n шаров. Найти вероятности событий: $A = \{$появится n_1 шаров с номером 1, n_2 шаров с номером 2, …, n_s шаров с номером $s\}$; $B = \{$не появятся шары с номерами 1 или 2$\}$.

18.86. Два равных по силе противника играют матч из n партий в теннис. Каждая партия заканчивается выигрышем, либо проигрышем одного из участников. Все исходы данного матча считаются равновероятными. Найти вероятность того, что первый игрок выиграет ровно m партий ($m \leqslant n$).

Схема выбора, приводящая к размещениям. Если опыт состоит в выборе m элементов без возвращения, но с упорядочиванием их по мере выбора в последовательную цепочку, то различными исходами данного опыта будут упорядоченные m-элементные подмножества множества E, отличающиеся либо набором элементов, либо порядком их следования. Получаемые при этом комбинации элементов (элементарные исходы) называются *размещениями из n элементов по m*, а их общее число определяется формулой

$$N(\Omega) = A_n^m = C_n^m \cdot m! = \frac{n!}{(n-m)!} = n(n-1)\ldots(n-m+1). \qquad (3)$$

В частном случае $m = n$ опыт фактически состоит в произвольном упорядочивании множества E, т.е. сводится к случайной перестановке элементов всего множества. При этом $N(\Omega) = A_n^n = n!$.

Пример 8. Множество E состоит из 10 первых букв русского алфавита. Опыт состоит в выборе без возвращения 4 букв и записи слова в порядке поступления букв. Сколько 4-буквенных слов может быть получено в данном опыте? Какова вероятность того, что наудачу составленное слово будет оканчиваться буквой a?

$\triangleleft$ $N(\Omega)$ — число всех 4-буквенных слов в данном опыте — равно числу 4-элементных упорядоченных подмножеств из 10 элементов, т.е.

$$N(\Omega) = A_{10}^4 = 10 \cdot 9 \cdot 8 \cdot 7 = 5040.$$

Пусть событие $A = \{$наудачу составленное слово из 4 букв множества E оканчивается буквой $a\}$. Число элементов множества A равно числу способов разместить на три оставшиеся места по одному символу из 9 (символ a исключен из рассмотрения, поскольку его место уже определено); таким образом,

$$N(A) = A_9^3 = 9 \cdot 8 \cdot 7 = 504$$

и

$$\mathrm{P}(A) = \frac{N(A)}{N(\Omega)} = \frac{504}{5040} = \frac{1}{10}. \ \triangleright$$

Числа 1, 2, ..., 9 записываются в случайном порядке. В задачах 18.87–18.89 найти вероятности указанных событий.

18.87. $A = \{$числа будут записаны в порядке возрастания$\}$.

18.88. $B = \{$числа 1 и 2 будут стоять рядом и в порядке возрастания$\}$, $C = \{$числа 3, 6 и 9 будут стоять рядом$\}$.

18.89. $D = \{$на четных местах будут стоять четные числа$\}$, $E = \{$сумма каждых двух чисел, стоящих на одинаковом расстоянии от концов, равна 10$\}$.

18.90. Группа, состоящая из 8 человек, занимает места за круглым столом в случайном порядке. Какова вероятность того, что при этом два определенных лица окажутся сидящими рядом?

18.91. Группа, состоящая из 8 человек, занимает места с одной стороны прямоугольного стола. Найти вероятность того, что два определенных лица окажутся рядом, если

а) число мест равно 8;

б) число мест равно 12.

18.92. На пяти карточках написаны цифры от 1 до 5. Опыт состоит в случайном выборе трех карточек и раскладывании их в порядке поступления в ряд слева направо. Найти вероятности следующих событий: $A = \{$появится число 123$\}$, $B = \{$появится число, не содержащее цифры 3$\}$.

18.93 (продолжение). В условиях предыдущей задачи найти вероятности событий: $C = \{$появится число, состоящее из последовательных цифр$\}$, $D = \{$появится четное число$\}$, $E = \{$появится число, содержащее хотя бы одну из цифр 2 или 3$\}$.

18.94. n человек входят в комнату, где имеется всего m стульев ($m \leqslant n$), и рассаживаются случайным образом, но так, что все стулья оказываются занятыми.

а) Показать, что число всех способов рассаживания определяется формулой (3).

б) Какова вероятность того, что два определенных лица окажутся без места?

в) Какова вероятность того, что k определенных лиц будут сидеть ($k \leqslant m$)?

18.95. n мужчин и n женщин случайным образом рассаживаются в ряд на $2n$ мест. Найти вероятности следующих событий: $A = \{$никакие два мужчины не будут сидеть рядом$\}$, $B = \{$все мужчины будут сидеть рядом$\}$.

18.96. 10 вариантов контрольной работы, написанные каждый на отдельной карточке, перемешиваются и распределяются случайным образом среди восьми студентов, сидящих в одном ряду, причем каждый получает по одному варианту. Найти вероятности следующих событий: $A = \{$варианты с номерами 1 и 2 останутся неиспользованными$\}$, $B = \{$варианты 1 и 2 достанутся рядом сидящим студентам$\}$, $C = \{$будут распределены последовательные номера вариантов$\}$.

18.97. 12 студентов, среди которых Иванов и Петров, случайным образом занимают очередь за учебниками в библиотеку. Какова вероятность, что между Ивановым и Петровым в образовавшейся очереди окажутся ровно 5 человек?

Схема выбора, приводящая к сочетаниям с повторениями. Если опыт состоит в выборе с возвращением m элементов множества $E = \{e_1, e_2, \ldots, e_n\}$, но без последующего упорядочивания, то различными исходами такого опыта будут всевозможные m-элементные наборы, отличающиеся составом. При этом отдельные наборы могут содержать повторяющиеся элементы. Например, при $m = 4$ наборы $\{e_1, e_1, e_2, e_1\}$ и $\{e_2, e_1, e_1, e_1\}$ неразличимы для данного эксперимента, а набор $\{e_1, e_1, e_3, e_1\}$ отличен от любого из предыдущих. Получающиеся в результате данного опыта комбинации называются *сочетаниями с повторениями*, а их общее число определяется формулой

$$N(\Omega) = C_{n+m-1}^{m}.$$

Пример 9. В технической библиотеке имеются книги по математике, физике, химии и т.д.; всего по 16 разделам науки. Поступили очередные четыре заказа на литературу. Считая, что любой состав заказанной литературы равновозможен, найти вероятности следующих событий: $A = \{$заказаны книги из различных разделов науки$\}$, $B = \{$заказаны книги из одного и того же раздела науки$\}$.

◁ Число всех равновероятных исходов данного эксперимента равно, очевидно, числу сочетаний с повторениями из 16 элементов по 4, т.е.

$$N(\Omega) = C_{16+4-1}^{4} = C_{19}^{4}.$$

Число исходов, благоприятствующих событию A, равно числу способов отобрать без возвращения четыре элемента из 16, поэтому

$$\mathrm{P}(A) = \frac{N(A)}{N(\Omega)} = \frac{C_{16}^4}{C_{19}^4} = \frac{455}{969} \approx 0{,}47.$$

Число исходов, благоприятствующих событию B, равно числу способов выбрать один элемент из шестнадцати, поэтому

$$\mathrm{P}(B) = \frac{N(B)}{N(\Omega)} = \frac{C_{16}^1}{C_{19}^4} = \frac{4}{969} \approx 0{,}004. \ \triangleright$$

18.98. В кондитерской имеется 7 видов пирожных. Очередной покупатель выбил чек на 4 пирожных. Считая, что любой заказываемый набор пирожных равновероятен, вычислить вероятность того, что покупатель заказал:

а) пирожные одного вида;

б) пирожные разных видов;

в) по два пирожных различных видов.

18.99. Из общего количества костей домино, содержащих числа 0, 1, 2, ..., n, наудачу извлекли одну кость. Оказалось, что это не дубль. Какова вероятность p_n того, что вторую извлеченную также наудачу кость домино можно будет приставить к первой. Найти числовые значения вероятности для $n = 6$ (обычный набор домино) и $n = 9$ (расширенный набор).

Схема выбора, приводящая к размещениям с повторениями. Если выбор m элементов из множества $E = \{e_1, e_2, \ldots, e_n\}$ производится с возвращением и с упорядочиванием их в последовательную цепочку, то различными исходами будут всевозможные m-элементные наборы (вообще говоря, с повторениями), отличающиеся либо составом элементов, либо порядком их следования. Например, при $m = 4$ множества $\omega_1 = \{e_1, e_1, e_2, e_1\}$, $\omega_2 = \{e_2, e_1, e_1, e_1\}$ и $\omega_3 = \{e_1, e_1, e_3, e_1\}$ являются различными исходами данного опыта. Получаемые в результате комбинации называются *размещениями с повторениями*, а их общее число определяется формулой

$$N(\Omega) = n^m.$$

Пример 10. 7 одинаковых шариков случайным образом рассыпаются по 4 лункам (в одну лунку может поместиться любое число шариков). Сколько существует различных способов распределения 7 шариков по 4 лункам? Какова вероятность того, что в результате данного опыта первая лунка окажется пустой (при этом может оказаться пустой и еще какая-либо лунка)?

$\triangleleft$ Занумеруем лунки и шарики. Можно считать, что опыт состоит в 7-кратном выборе с возвращением номера лунки и записи 7-буквенного слова. При этом каждому порядковому номеру буквы (номеру шарика) будет поставлена в соответствие одна из четырех букв алфавита (номер лунки).

Так, например, слово

1	1	3	1	4	4	2
1	2	3	4	5	6	7

означает, что в первую лунку попали шары № 1, № 2 и № 4, во вторую лунку — шар № 7, в третью — шар № 3, в четвертую — шары № 5 и № 6. Таким образом, число всех способов распределить 7 шариков по 4 лункам равно числу различных 7-буквенных слов из алфавита в 4 буквы, т.е. $N(\Omega) = 4^7$.

Событие $A = \{$первая лунка окажется пустой$\}$ соответствует такому выбору, когда символ 1 (номер первой лунки) удален из алфавита. Поэтому $N(A) = 3^7$ и

$$\mathrm{P}(A) = \frac{N(A)}{N(\Omega)} = \left(\frac{3}{4}\right)^7 \approx 0{,}133. \ \triangleright$$

18.100. Бросается 10 одинаковых игральных костей. Вычислить вероятности следующих событий: $A = \{$ни на одной кости не выпадет 6 очков$\}$, $B = \{$хотя бы на одной кости выпадет 6 очков$\}$, $C = \{$ровно на 3 костях выпадет 6 очков$\}$.

18.101. Опыт состоит в четырехкратном выборе с возвращением одной из букв алфавита $E = \{а, б, к, о, м\}$ и выкладывании слова в порядке поступления букв. Какова вероятность того, что в результате будет выложено слово *мама*?

18.102. В подъезде дома установлен замок с кодом. Дверь автоматически отпирается, если в определенной последовательности набрать три цифры из имеющихся десяти. Некто вошел в подъезд и, не зная кода, стал наудачу пробовать различные комбинации из трех цифр. На каждую попытку он тратит 20 секунд. Какова вероятность события $A = \{$вошедшему удастся открыть дверь за один час$\}$?

18.103. Телефонная книга раскрывается наудачу и выбирается случайный номер телефона. Считая, что телефонные номера состоят из 7 цифр, причем все комбинации цифр равновероятны, найти вероятности следующих событий: $A = \{$четыре последние цифры телефонного номера одинаковы$\}$, $B = \{$все цифры различны$\}$.

18.104 (продолжение). В условиях предыдущей задачи найти вероятности событий: $C = \{$номер начинается с цифры 5$\}$, $D = \{$номер содержит три цифры 5, две цифры 1 и две цифры 2$\}$.

18.105. Шесть человек вошли в лифт на первом этаже семиэтажного дома. Считая, что любой пассажир может с равной вероятностью выйти на 2-м, 3-м, ..., 7-м этажах, найти вероятности следующих событий: $A = \{$на втором, третьем и четвертом этажах не выйдет ни один из пассажиров$\}$, $B = \{$трое пассажиров выйдут на седьмом этаже$\}$, $C = \{$на каждом этаже выйдет по одному пассажиру$\}$, $D = \{$все пассажиры выйдут на одном этаже$\}$.

18.106. К четырехстороннему перекрестку с каждой стороны подъехало по одному автомобилю. Каждый автомобиль может с равной вероятностью совершить один из четырех маневров на перекрестке: развернуться и поехать обратно, поехать прямо, налево или направо. Через некоторое время все автомобили покинули перекресток. Найти вероятности следующих событий: $A = \{$все автомобили поедут по одной и той же улице$\}$, $B = \{$по определенной улице поедут ровно три автомобиля$\}$, $C = \{$по крайней мере по одной из улиц не поедет ни один автомобиль$\}$.

18.107*. n различных предметов случайным образом распределяются по s занумерованным ящикам таким образом, чтобы k-й по счету ящик содержал ровно $n_k \geqslant 0$ предметов $(n_1 + n_2 + \dots + n_s = n)$. Показать, что число всех элементарных исходов данного опыта (*число упорядоченных разбиений из n по s*) определяется формулой

$$N(\Omega) = \frac{n!}{n_1! n_2! \dots n_s!}. \tag{4}$$

18.108. Десять приезжих мужчин, среди которых Петров и Иванов, размещаются в гостинице в два трехместных и один четырехместный номер. Сколько существует способов их размещения? Какова вероятность события A, состоящего в том, что Петров и Иванов попадут в четырехместный номер?

18.109. В условиях задачи 18.80 найти вероятность события $D = \{$будут выбраны: 1 первокурсник, 2 второкурсника и 2 третьекурсника$\}$.

18.110. 20 футбольных команд, среди которых 4 призера предыдущего первенства, по жеребьевке разбиваются на 4 занумерованные подгруппы по 5 команд. Найти вероятности событий: $A = \{$в первую и вторую подгруппы не попадет ни один из призеров$\}$, $B = \{$в каждую подгруппу попадет один из призеров$\}$.

18.111. Множество E состоит из n символов, среди которых n_1 символов e_1, n_2 символов e_2, ..., n_s символов e_s $(n_1+n_2+\dots+n_s = n)$. Опыт состоит в поэлементном выборе без возвращения всех n элементов множества E и записи слова.

а)** Показать, что число всех различных слов, полученных в данном эксперименте, определяется формулой (4).

б) Какова вероятность, что в полученном слове первыми n_1 буквами является символ e_1?

18.112. Бросается 6 игральных костей. Найти вероятности следующих событий: $A = \{$выпадут 3 единицы, две тройки и одна шестерка$\}$, $B = \{$выпадут различные цифры$\}$, $C = \{$выпадут три и только три одинаковые цифры$\}$, $D = \{$выпадут только нечетные цифры$\}$, $E = \{$выпадут три четные и три нечетные цифры$\}$.

18.113. Из разрезной азбуки выкладывается слово *математика*. Затем все буквы этого слова тщательно перемешиваются и снова выкладываются в случайном порядке. Какова вероятность того, что снова получится слово *математика*?

18.114. 52 карты раздаются четырем игрокам (каждому по 13 карт). Найти вероятности следующих событий: $A = \{$каждый игрок получит туза$\}$, $B = \{$первый игрок получит все 13 карт одной масти$\}$.

18.115 (продолжение). В условиях предыдущей задачи найти вероятности следующих событий: $C = \{$все тузы попадут к одному из игроков$\}$, $D = \{$двое определенных игроков не получат ни одного туза$\}$.

Решить задачи 18.116–18.138, используя подходящие комбинаторные схемы.

18.116. Множество E состоит из трех различных элементов: $E = \{a, b, c\}$. Выписать состав Ω во всех четырех опытах по выбору двух элементов из множества E без возвращения и с возвращением, без упорядочивания и с упорядочиванием. Определить число элементов множества Ω (число различных выборок) в каждом из четырех случаев и сравнить результат с тем, который получается по соответствующей комбинаторной формуле.

18.117. Опыт состоит в случайном выборе одного элемента из множества $E_1 = \{a, b\}$ и одного элемента из множества $E_2 = \{a, b, c\}$. Перечислить состав множества $E = E_1 \times E_2$. Какова вероятность того, что выборка будет состоять из одинаковых элементов?

18.118. Из урны, содержащей шары с номерами $1, 2, \ldots, n$, k раз ($k \leqslant n$) вынимается шар и каждый раз возвращается обратно. Найти вероятность того, что номера вынутых шаров образуют возрастающую последовательность.

18.119. В условиях эксперимента, описанного в задаче 18.74, при $m > n$ найти вероятность того, что орудия выстрелят по различным самолетам.

18.120. Из урны, содержащей шары с номерами $1, 2, \ldots, n$, наудачу отбирается k шаров и номера вынутых шаров записываются последовательно. Какова вероятность того, что на фиксированном m-месте ($m \leqslant k$) окажется шар с номером m, если выбор осуществляется: а) без возвращения; б) с возвращением?

18.121. На тренировке детской спортивной школы по футболу роли игроков распределяются случайным образом среди одиннадцати участников. Нужно отобрать одного вратаря, четырех защитников, трех полузащитников и трех нападающих. Какова вероятность того, что два друга-участника Коля и Миша: а) будут играть в нападении; б) получат разные роли, причем один из друзей будет играть в нападении, а другой — в защите?

18.122. В лотерее выпущено n билетов, из которых m выигрышные. Куплено k билетов. Какова вероятность следующих событий: $A = \{$из k билетов хотя бы один выигрышный$\}$, $B = \{$из k билетов ровно один выигрышный$\}$, $C = \{$из k билетов ровно k_1 выигрышных$\}$?

18.123. Какова вероятность p_n того, что в группе из n ($n \leqslant 365$) случайно отобранных студентов хотя бы у двоих окажется один и тот же день рождения? Оценить значение p_n для $n = 24$ и $n = 50$.

18.124. На заводе работает 30 000 рабочих и служащих. Показать, что на данном заводе обязательно найдутся хотя бы два человека с одинаковыми инициалами имени, отчества и фамилии.

18.125. Регистр калькулятора содержит 8 разрядов. Считая, что появление любого числа на регистре равновероятно, определить вероятности следующих событий: $A = \{$во всех разрядах стоят нули$\}$, $B = \{$во всех разрядах стоят одни и те же цифры$\}$.

18.126 (продолжение). В условиях предыдущей задачи найти вероятности следующих событий: $C = \{$регистр содержит ровно две одинаковые цифры$\}$, $D = \{$регистр содержит ровно две пары одинаковых цифр$\}$.

18.127 (продолжение). В условиях задачи 18.125 найти вероятности событий: $E = \{$регистр содержит ровно три одинаковые цифры$\}$, $F = \{$регистр содержит три и только три различные цифры$\}$.

18.128. 7 яблок, 3 апельсина и 5 лимонов раскладываются случайным образом в три пакета, но так, чтобы в каждом было одинаковое количество фруктов. Найти вероятности следующих событий: $A = \{$в каждом из пакетов по одному апельсину$\}$, $B = \{$определенный пакет не содержит апельсинов$\}$.

18.129. Из множества чисел $E = \{1, 2, \ldots, n\}$ выбирается два числа. Какова вероятность, что второе число больше первого, если выбор осуществляется: а) без возвращения; б) с возвращением?

18.130. Из множества чисел $E = \{1, 2, \ldots, n\}$ выбирается три числа. Какова вероятность того, что второе число заключено между первым и третьим, если выбор осуществляется: а) без возвращения; б) с возвращением?

18.131. Каждая из n палок случайным образом ломается на две части — длинную и короткую. Затем $2n$ полученных обломков наудачу объединяются в n пар, каждая из которых образует новую палку. Найти вероятности следующих событий: $A = \{$все обломки объединятся в первоначальном порядке$\}$, $B = \{$все длинные части объединятся с короткими$\}$.

18.132. Путем жеребьевки разыгрывается шесть подписных изданий среди десяти участников.

Сколько различных распределений подписок возможно, если каждое очередное наименование разыгрывается между всеми участниками? Найти вероятность того, что первые шесть человек получат каждый по одной подписке.

18.133 (продолжение). В условиях предыдущей задачи ответить на те же вопросы, если каждый участник, получивший подписку, выбывает из игры.

18.134. Опыт состоит в том, что n различных предметов случайным образом распределяются среди m человек ($m < n$), причем таким образом, что каждый может получить любое число предметов из имеющихся. Какова вероятность следующих событий: $A = \{$все предметы достанутся одному из участников$\}$, $B = \{$определенное лицо не получит ни одного предмета$\}$.

18.135 (продолжение). В условиях эксперимента предыдущей задачи найти вероятности событий: $C = \{$определенные m_1 лиц получат по одному предмету$\}$, $D = \{$определенные n_1 предметов достанутся одному из участников$\}$.

18.136. В условиях эксперимента, описанного в задаче 18.106, найти вероятности следующих событий: $E = \{$автомобили разъедутся по улицам попарно$\}$, $F = \{$по определенной улице поедут два автомобиля$\}$.

18.137. Условия эксперимента, описанного в задаче 18.106, изменены следующим образом: на перекрестке запрещены развороты. Все остальные направления движения для любого из автомобилей равновероятны. Найти вероятности событий E и F, определенных в предыдущей задаче.

18.138*. Из совокупности всех непустых подмножеств множества $E = \{e_1, e_2, \ldots, e_n\}$ по схеме выбора с возвращением отбираются два подмножества E_1 и E_2. Какова вероятность того, что они пересекаются?

6. Геометрические вероятности. Формула классической вероятности следующим образом обобщается на случай непрерывных пространств элементарных исходов Ω.

Пусть Ω-квадрируемая область на плоскости. Рассмотрим систему $\mathcal{F}$ квадрируемых подмножеств множества Ω. Как отмечалось в примере 4, система $\mathcal{F}$ является σ-алгеброй. Пусть условия опыта таковы, что вероятность попадания в произвольную квадрируемую подобласть ω области Ω пропорциональна площади этой подобласти и не зависит от ее местоположения в Ω. При этих условиях для вероятности осуществления любого наблюдаемого в данном эксперименте события $A = \{(x, y) \in \in \omega \,|\, \omega \in \mathcal{F}\}$ справедлива *формула геометрической вероятности*

$$\mathrm{P}(A) = \mathrm{P}\{(x, y) \in \omega \,|\, \omega \in \mathcal{F}\} = \frac{S(\omega)}{S(\Omega)}, \qquad (5)$$

где $S(\omega)$ — площадь подобласти ω.

Формула (5) естественным образом обобщается на случай пространств произвольной размерности:

$$\mathrm{P}(A) = \mathrm{P}\{(x_1, x_2, \ldots, x_n) \in \omega \,|\, \omega \in \mathcal{F}\} = \frac{\mathrm{mes}(\omega)}{\mathrm{mes}(\Omega)},$$

где $\mathrm{mes}(\omega)$ — мера множества ω (длина, площадь, объем и т.д. в зависимости от размерности того пространства, в котором рассматриваются данные множества).

Пример 11. На бесконечную шахматную доску со стороной квадрата a наудачу бросается монета радиуса $r < a/2$. Найти вероятности следующих событий: $A = \{$монета попадет целиком внутрь одного квадрата$\}$, $B = \{$монета пересечет не более одной стороны квадрата$\}$.

◁ Пусть (x, y) — координаты центра упавшей монеты. В силу бесконечности шахматной доски можно считать, что элементарные исходы данного эксперимента полностью определяются положением центра упавшей монеты относительно вершин квадрата, содержащего этот центр. Помещая начало координат в одну из вершин указанного квадрата, можем записать множество элементарных исходов в виде $\Omega = \{(x, y) \,|\, 0 \leqslant x, y \leqslant a\}$. Множество ω_A, соответствующее событию A, имеет вид $\omega_A = \{(x, y) \,|\, r \leqslant x, y \leqslant a - r\}$, т.е. является квадратом со стороной $a - 2r$. По формуле геометрической вероятности (5) находим

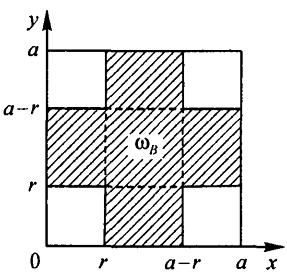


Рис. 4

$$\mathrm{P}(A) = \frac{S(\omega_A)}{S(\Omega)} = \frac{(a - 2r)^2}{a^2}.$$

Множество ω_B, имеющее более сложную структуру, изображено на рис. 4. Так как $S(\omega_B) = a^2 - 4r^2$, то, снова используя формулу (5), находим

$$\mathrm{P}(B) = \frac{S(\omega_B)}{S(\Omega)} = 1 - 4\,\frac{r^2}{a^2}. \ \triangleright$$

18.139. Внутри квадрата с вершинами (0, 0), (1, 0), (1, 1) и (0, 1) наудачу выбирается точка $M(x, y)$. Найти вероятность события $A = \{(x, y) \,|\, x^2 + y^2 \leqslant a^2,\ a > 0\}$.

18.140 (продолжение). В условиях предыдущей задачи найти вероятность события $B = \{(x, y) \mid xy < a, a > 0\}$.

18.141 (продолжение). В условиях задачи 18.139 найти вероятности событий: $C = \{(x, y) \mid \max(x, y) < a, a > 0\}$, $D = \{(x, y) \mid \min(x, y) < a, 0 \leqslant a \leqslant 1\}$.

18.142. На перекрестке установлен автоматический светофор, в котором одну минуту горит зеленый свет и полминуты — красный, затем снова одну минуту — зеленый и полминуты — красный и т. д. В случайный момент времени к перекрестку подъезжает автомобиль. Какова вероятность того, что он проедет перекресток без остановки?

18.143. Какова вероятность того, что сумма трех наудачу взятых отрезков, длина каждого из которых не превосходит l, будет больше l?

18.144. Луч локатора перемещается в горизонтальной плоскости с постоянной угловой скоростью. Какова вероятность того, что цель будет обнаружена в угловом секторе α радиан, если появление цели по любому направлению одинаково возможно?

18.145. Какова вероятность, не целясь, попасть бесконечно малой пулей в прутья квадратной решетки, если толщина прутьев равна a, а расстояние между их осями равно l $(l > a)$?

18.146. На поверхности шара берут наудачу две точки и соединяют меньшей дугой большого круга. Найти вероятность того, что дуга не превзойдет α радиан.

18.147. Какова вероятность того, что случайно выбранная на глобусе точка лежит: а) за полярным кругом ($66°33'$ северной широты); б) между $60°$ и $30°$ северной широты; в) между $10°$ и $40°$ западной долготы?

18.148 (*задача о встрече*). Найти вероятности событий A из задачи 18.7 и C из задачи 18.8.

18.149 (продолжение задачи о встрече). Найти вероятности событий D, E и F из задачи 18.9.

18.150. Два парохода должны подойти к одному и тому же причалу. Время прихода обоих пароходов независимо и равновозможно в течение данных суток. Определить вероятность того, что одному из пароходов придется ожидать освобождения причала, если время стоянки первого парохода — один час, а второго — два часа.

18.151. В случайный момент времени $x \in [0, T]$ появляется радиосигнал длительностью t_1. В случайный момент времени $y \in [0, T]$ включается приемник на время $t_2 < t_1$. Найти вероятность обнаружения сигнала, если: а) приемник настраивается мгновенно; б) время настройки приемника равно t_3 $(t_3 < t_2 < t_1)$.

18.152*. Значения a и b равновозможны в квадрате $|a| \leqslant 1$, $|b| \leqslant 1$. Найти вероятности следующих событий: $A = \{$корни квадратного трехчлена $x^2 + 2ax + b$ действительны$\}$, $B = \{$корни квадратного трехчлена $x^2 + 2ax + b$ положительны$\}$.

18.153*. Однородный прямой круговой цилиндр с высотой h и радиусом основания r случайным образом бросается на горизонтальную плоскость. Найти вероятность того, что цилиндр упадет на боковую поверхность.

18.154*. На отрезке длины l наудачу выбираются две точки M_1 и M_2. Определить вероятность того, что из полученных трех отрезков можно построить треугольник.

18.155*. На плоскость с нанесенной на ней квадратной сеткой многократно бросается монета диаметра d, в результате чего установлено, что в 40 % случаев монета не пересекает ни одной стороны квадрата. Оценить размер сетки.

18.156*. На окружности единичного радиуса наудачу ставятся три точки A, B и C. Какова вероятность того, что треугольник ABC остроугольный?

18.157. Даны две концентрические окружности радиусов $r_2 > r_1$. На большей окружности наудачу ставятся две точки A и B. Какова вероятность того, что отрезок AB не пересечет малую окружность?

18.158. Из отрезка $[-1, 2]$ наудачу взяты два числа. Какова вероятность того, что их сумма больше единицы, а произведение меньше единицы?

18.159* (*задача Бюффона*). На плоскость, разграфленную параллельными прямыми линиями, отстоящими друг от друга на расстояние $2a$, наудачу бросается игла длиной $2l$. Какова вероятность того, что игла пересечет одну из параллельных прямых, если $l \leqslant a$.

18.159 (1) (*парадокс Бертрана*). В круге единичного радиуса наудачу проводится хорда. Какова вероятность, что длина хорды окажется больше длины стороны правильного треугольника, вписанного в данную окружность?

Парадокс заключается в неоднозначности результата, который зависит от истолкования слова «наудачу». Рассмотрим следующие варианты.

1) Случайным образом выбирается точка в круге и через нее единственным способом проводится хорда перпендикулярно радиусу.

2) Как бы ни была проведена хорда, всегда можно сориентировать вписанный треугольник таким образом, чтобы одна из его вершин совпала с одним из концов хорды. Фиксируем этот конец хорды, например, в точке $(1, 0)$. Другую точку выбираем наудачу на окружности.

3) Тот же правильный треугольник можно сориентировать так, чтобы одна из его высот была параллельна проведенной хорде. Поэтому считаем, что хорда проводится параллельно заданному направлению, на-

пример оси Ox. Для этого достаточно выбрать наудачу точку на радиусе, ориентированном вдоль оси Oy, и провести единственную хорду перпендикулярно радиусу.

7. Условные вероятности. Независимость событий. Ранее было введено понятие вероятности как числовой функции, определенной на поле событий для данного эксперимента и удовлетворяющей трем основным аксиомам (п. 3). Такую вероятность называют *безусловной вероятностью*, подчеркивая этим, что она не зависит ни от каких дополнительных условий, кроме фиксированного комплекса условий S, которым характеризуется эксперимент.

Пусть A и B — наблюдаемые события в эксперименте, причем $\mathrm{P}(A) > 0$. *Условной вероятностью* $\mathrm{P}(B/A)$ *осуществления события B при условии, что событие A произошло в результате данного эксперимента*, называется величина, определяемая равенством

$$\mathrm{P}(B/A) = \frac{\mathrm{P}(AB)}{\mathrm{P}(A)}. \tag{6}$$

Для краткости условную вероятность $\mathrm{P}(B/A)$ называют «вероятностью события B при условии A». При $\mathrm{P}(A) = 0$ условная вероятность $\mathrm{P}(B/A)$ не определена.

Нетрудно проверить, что в случае произвольного вероятностного пространства определенная формулой (6) вероятность $\mathrm{P}(B/A)$, рассматриваемая как функция наблюдаемых событий $B \in \mathcal{F}$ при фиксированном событии A, удовлетворяет трем основным аксиомам и всем их следствиям (см. задачу 18.160).

Формула (6), по существу, сводит вопрос о вычислении условной вероятности к вычислению двух безусловных вероятностей, определенных в заданном вероятностном пространстве.

Пример 12. Из урны, содержащей 3 белых и 7 красных шаров, наудачу последовательно и без возвращения извлекаются два шара. События: $A = \{$первый шар белый$\}$, $B = \{$второй шар белый$\}$, $C = \{$по крайней мере один из вынутых шаров белый$\}$. Вычислить вероятности $\mathrm{P}(B/A)$, $\mathrm{P}(A/B)$ и $\mathrm{P}(A/C)$.

$\triangleleft$ Для вычисления искомых условных вероятностей воспользуемся формулой (6). Занумеруем белые шары цифрами 1, 2, 3, а черные — цифрами 4, 5, ..., 10. Согласно описанию эксперимента имеем следующую схему: выбор наудачу, без возвращения пары чисел из множества $\{1, 2, \ldots, 10\}$ с упорядочиванием, поэтому множество элементарных исходов можно записать в виде

$$\Omega = \{\omega_{ij} = (i, j) \mid i = 1, \ldots, 10;\ j = 1, \ldots, 10;\ i \neq j\}.$$

Отсюда следует, что $N(\Omega) = A_{10}^2 = 90$, $\mathrm{P}(\omega_{ij}) = 1/90$ для всех допустимых значений i и j.

Подмножества, соответствующие событиям A и B, имеют следующий состав:

$$\begin{aligned} A &= \{(i, j) \mid i = 1, 2, 3; \quad j = 1, \ldots, 10; \quad i \neq j\}, \\ B &= \{(i, j) \mid i = 1, \ldots, 10; \quad j = 1, 2, 3; \quad i \neq j\}, \end{aligned}$$

причем, как нетрудно подсчитать, $N(A) = 3 \cdot 9$, $N(B) = 9 \cdot 3$,

$$P(A) = \frac{3}{10}, \quad P(B) = \frac{3}{10}.$$

Событию AB соответствует подмножество

$$AB = \{(i, j) \,|\, i = 1, 2, 3;\ j = 1, 2, 3;\ i \neq j\},$$

следовательно, $N(AB) = 3 \cdot 2$,

$$P(AB) = \frac{1}{15}.$$

По формуле (6) отсюда находим

$$P(B/A) = \frac{P(AB)}{P(A)} = \frac{2}{9}, \quad P(A/B) = \frac{P(AB)}{P(B)} = \frac{2}{9}.$$

Далее, по формуле классической вероятности

$$P(C) = 1 - P(\overline{C}) = 1 - \frac{C_7^2}{C_{10}^2} = \frac{8}{15}.$$

Для вычисления вероятности произведения AC заметим, что $AB \subset$ $\subset A$, поэтому $AC = A(A + B) = A + AB = A$ (см. задачу 18.19, формула (1)). Отсюда

$$P(AC) = P(A) = \frac{3}{10}.$$

Наконец, снова используя формулу (6), получаем

$$P(A/C) = \frac{P(A)}{P(C)} = \frac{9}{16}. \ \triangleright$$

18.160. Пусть A, B и C — наблюдаемые события, причем $P(C) > 0$, $P(AC) > 0$. Доказать справедливость следующих формул для условной вероятности:

$P(AB/C) = P(A/C)\,P(B/AC)$ (формула умножения),

$P(A + B/C) = P(A/C) + P(B/C) - P(AB/C)$ (формула сложения).

18.161. Показать, что если A, B и C — такие наблюдаемые в эксперименте события, что $P(A) \neq 0$, $BC \neq \varnothing$ и $A = \overline{BC}$, то справедлива следующая формула сложения:

$$P(B + C/A) = P(B/A) + P(C/A).$$

18.162. Доказать, что $\mathrm{P}(A/B) \geqslant 1 - \dfrac{\mathrm{P}(\overline{A})}{\mathrm{P}(B)}$.

18.163. Один раз подбрасывается игральная кость. События: $A = \{$выпало простое число очков$\}$, $B = \{$выпало четное число очков$\}$. Вычислить вероятность $\mathrm{P}(A/B)$.

18.164. Вероятность попасть в самолет равна 0,4, а вероятность его сбить равна 0,1. Найти вероятность того, что при попадании в самолет он будет сбит.

18.165. Вероятность того, что прибор не откажет к моменту времени t_1, равна 0,8, а вероятность того, что он не откажет к моменту времени t_2 ($t_2 > t_1$), равна 0,6. Найти вероятность того, что прибор, не отказавший к моменту времени t_1, не откажет и к моменту времени t_2.

18.166. Электрическая схема (рис. 5) состоит из элементов, каждый из которых в момент включения с равной вероятностью может либо проводить, либо не проводить ток. Состояние каждого из элементов не влияет на состояния остальных. Введем следующие события: $C = \{$цепь проводит ток$\}$, $A_i = \{i$-й элемент проводит ток, $i = 1, 2, \ldots, 5\}$. Вычислить $\mathrm{P}(A_1/C)$ и $\mathrm{P}(A_2/C)$.

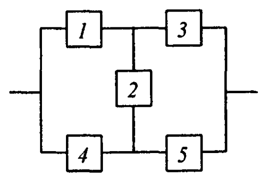


Рис. 5

18.167*. В семье двое детей. Считая, что рождение мальчика и девочки — независимые и равновероятные события, вычислить вероятность того, что оба ребенка — мальчики, если известно, что в семье есть мальчик.

18.168. Из множества чисел $\{1, 2, \ldots, N\}$ по схеме случайного выбора без возвращения выбирают три числа. Найти условную вероятность того, что третье число попадет в интервал, образованный первыми двумя, если известно, что первое число меньше второго.

18.169. Подбрасывают наудачу три игральные кости. Наблюдаемые события: $A = \{$на трех костях выпадут разные грани$\}$, $B = \{$хотя бы на одной из костей выпадет шестерка$\}$. Вычислить $\mathrm{P}(B/A)$ и $\mathrm{P}(A/B)$.

Событие A называется *независимым от события* B, удовлетворяющего условию $\mathrm{P}(B) > 0$, если выполняется равенство

$$\mathrm{P}(A/B) = \mathrm{P}(A).$$

События A и B называются *независимыми*, если

$$\mathrm{P}(AB) = \mathrm{P}(A)\,\mathrm{P}(B). \tag{7}$$

События $A_1, A_2, \ldots, A_n$ называются *независимыми в совокупности*, если для любого набора из m событий ($m = 2, 3, \ldots, n$) выполняется равенство

$$\mathrm{P}(A_{k_1} A_{k_2} \ldots A_{k_m}) = \mathrm{P}(A_{k_1})\mathrm{P}(A_{k_2}) \ldots \mathrm{P}(A_{k_m}), \quad k_m \in \{1, 2, \ldots, n\}. \quad (8)$$

Пример 13. В условиях эксперимента, описанного в примере 12, установить, являются ли независимыми события A и B, A и C, B и C, события A, B и C в совокупности.

$\lhd$ Так как все необходимые вероятности вычислены в примере 12, то для решения задачи достаточно проверить, выполняется ли для каждой пары событий критерий независимости (7), а для трех событий A, B и C — критерий (8). Имеем

$$\mathrm{P}(AB) = \frac{1}{15}, \quad \mathrm{P}(A)\,\mathrm{P}(B) = \frac{9}{100} \neq \mathrm{P}(AB),$$

т.е. события A и B не являются независимыми (в таком случае говорят, что они зависимы). Далее, как установлено в том же примере,

$$\mathrm{P}(AC) = \mathrm{P}(A) \neq \mathrm{P}(A)\,\mathrm{P}(C),$$

так как $\mathrm{P}(C) \neq 1$. Следовательно, события A и C также зависимы. Наконец,

$$\mathrm{P}(BC) = \mathrm{P}(B(A+B)) = \mathrm{P}(AB+B) = \mathrm{P}(B) \neq \mathrm{P}(B)\,\mathrm{P}(C),$$

поэтому и события B и C являются зависимыми.

События A, B и C не являются независимыми в совокупности, так как согласно критерию (8) для этого необходимо, чтобы все три события были попарно независимы. $\rhd$

Формулы (7) и (8) позволяют выделять независимые события в тех случаях, когда модель вероятностного эксперимента формализована и вероятности всех нужных событий полностью определены. Однако в практических задачах, связанных с проведением реальных экспериментов, далеко не всегда возможно использование данных критериев независимости. В таких случаях часто применяют гипотезу о физической независимости событий: считаются независимыми события, не связанные причинно.

Так, например, естественно считать независимыми результаты стрельбы из двух орудий при одновременном выстреле по цели или события, связанные с появлением брака определенного вида изделий, производимых двумя поточными линиями на различных предприятиях, и т. д.

18.170. Пусть A и B — наблюдаемые события в эксперименте, причем $\mathrm{P}(A) > 0$, $\mathrm{P}(B) > 0$ и событие A не зависит от B. Показать, что справедливы следующие утверждения; а) событие B не зависит от A; б) события A и B независимы. (Тем самым устанавливается, что свойство независимости двух событий взаимно.)

18.171. Пусть события A и B несовместны, причем $\mathrm{P}(A) \neq 0$ и $\mathrm{P}(B) \neq 0$. Доказать, что они зависимы. В частности, отсюда следует, что элементарные исходы любого вероятностного эксперимента зависимы.

18.172. Пусть события A и B независимы и не являются невозможными. Доказать, что они обязательно совместны.

18.173. События A и B зависимы. Следует ли из этого, что они несовместны? Привести пример.

18.174*. Пусть события A и B независимы. Показать, что тогда независимы и события $\overline{A}$ и $\overline{B}$.

18.175. Пусть для двух наблюдаемых в эксперименте событий A и B выполняются условия $A \neq \Omega$, $A \neq \varnothing$, $\mathrm{P}(A) > 0$, $\mathrm{P}(B/A) = \mathrm{P}(B/\overline{A})$. Показать, что события A и B независимы.

18.176. Из колоды в 36 карт наудачу извлекается одна карта. События: $A = \{$вынутая карта — туз$\}$, $B = \{$вынута карта черной масти$\}$, $F = \{$вынутая карта — фигура, т.е. является валетом, дамой, королем или тузом$\}$. Установить, зависимы или независимы следующие три пары событий: A и B, A и F, F и B.

18.177. В условиях эксперимента, описанного в задаче 18.169, установить, зависимы или независимы события $C = \{$появится не менее двух единиц$\}$ и $D = \{$появится четное число нечетных цифр$\}$. Вычислить условную вероятность $\mathrm{P}(D/C)$.

18.178. Тетраэдр, три грани которого окрашены соответственно в красный, желтый и синий цвета, а четвертая грань содержит все три цвета, бросается наудачу на плоскость. События K, G и S состоят в том, что тетраэдр упал на грань, содержащую соответственно красный, желтый либо синий цвет. Доказать, что указанные события попарно независимы, но не являются независимыми в совокупности.

18.179. Из 100 студентов, находящихся в аудитории, 50 человек знают английский язык, 40 — французский и 35 — немецкий. Английский и французский языки знают 20 студентов, английский и немецкий — 8, французский и немецкий — 10. Все три языка знают 5 человек. Один из студентов вышел из аудитории. Рассмотрим следующие события: $E = \{$вышедший знает английский язык$\}$, $F = \{$вышедшяй знает французский язык$\}$, $D = \{$вышедший знает немецкий язык$\}$.

а) Указать все пары независимых событий.

б) Установить, являются ли события E, F и D независимыми в совокупности.

18.180. Производится два последовательных извлечения по одному шару без возвращения из урны, содержащей m_1 белых и m_2 черных шаров. События: $A = \{$первый шар белый$\}$,

$B = \{$второй шар белый$\}$. Показать, что

$$\mathrm{P}(B/A) = \frac{m_1 - 1}{m_1 + m_2 - 1} = \mathrm{P}(B'), \qquad (*)$$

где $B' = \{$вынутый шар белый$\}$ — событие, наблюдаемое в новом эксперименте, состоящем в выборе наудачу одного шара из урны, состав которой изменен в соответствии с условием события A.

Указанный метод вычисления условной вероятности называется *методом вспомогательного эксперимента.*

◁ Так как эксперимент представляет собой схему выбора без возвращения и с упорядочиванием, то $N(\Omega) = A^2_{m_1+m_2}$, $N(AB) = A^2_{m_1}$.

Следовательно,

$$\mathrm{P}(AB) = \frac{N(AB)}{N(\Omega)} = \frac{A^2_{m_1}}{A^2_{m_1+m_2}} = \frac{m_1(m_1-1)}{(m_1+m_2)(m_1+m_2-1)}.$$

Аналогично находим

$$\mathrm{P}(A) = \mathrm{P}(AB + A\overline{B}) = \mathrm{P}(AB) + \mathrm{P}(A\overline{B}) =$$
$$= \frac{A^2_{m_1}}{A^2_{m_1+m_2}} + \frac{C^1_{m_1} C^1_{m_2}}{A^2_{m_1+m_2}} = \frac{m_1(m_1+m_2-1)}{(m_1+m_2)(m_1+m_2-1)}.$$

Подставляя полученные выражения в формулу (6) для вычисления условной вероятности, получаем

$$\mathrm{P}(B/A) = \frac{m_1-1}{m_1+m_2-1},$$

и первая часть равенства $(*)$ доказана.

Для доказательства второй части равенства $(*)$ заметим, что согласно условию события A один белый шар удален из урны. Новый (вспомогательный) эксперимент состоит в том, что из оставшихся $m_1 + m_2 - 1$ шаров наудачу извлекают один шар. Вероятность, что он окажется белым (событие B'), определяется по классической формуле

$$\mathrm{P}(B') = \frac{N(B')}{N(\Omega')} = \frac{m_1-1}{m_1+m_2-1},$$

что и доказывает вторую часть равенства $(*)$. ▷

В задачах 18.181–18.185 вычислить указанные условные вероятности методом вспомогательного эксперимента.

18.181. Найти вероятность $\mathrm{P}(B/A)$ в условиях эксперимента, описанного в примере 12.

18.182. В ящике лежат 12 красных, 8 зеленых и 10 синих шаров. Наудачу вынимаются два шара. Найти вероятность того, что будут вынуты шары разного цвета, при условии, что не вынут синий шар.

18.183. На шахматную доску наудачу ставятся два слона — белый и черный. Какова вероятность того, что слоны не побьют друг друга при условии, что белый слон попадет на одно из крайних полей доски?

18.184. Известно, что 5% всех мужчин и 0,25% всех женщин — дальтоники. На обследование прибыло одинаковое число мужчин и женщин. Наудачу выбранное лицо оказалось дальтоником. Какова вероятность, что это мужчина?

18.185. На шахматную доску наудачу ставят две ладьи. Вычислить $\mathrm{P}(B/A)$, если $A = \{$ладьи попали на клетки разного цвета$\}$, $B = \{$ладьи побьют друг друга$\}$.

18.186. Доказать, что для любого эксперимента любое наблюдаемое событие A не зависит от события Ω. Объяснить этот результат.

8. Вероятности сложных событий. *Сложным событием* называется наблюдаемое событие, выраженное через другие наблюдаемые в том же эксперименте события с помощью допустимых алгебраических операций.

Вероятность осуществления того или иного сложного события вычисляется по правилам, основу которых составляют:

формула умножения вероятностей

$$\mathrm{P}(AB\} = \mathrm{P}(A)\,\mathrm{P}(B/A) = \mathrm{P}(B)\,\mathrm{P}(A/B), \tag{9}$$

формула сложения вероятностей

$$\mathrm{P}(A+B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(AB). \tag{10}$$

Формула (9) справедлива, если $\mathrm{P}(A) > 0$, $\mathrm{P}(B) > 0$, и позволяет вычислять вероятность совместного осуществления событий A и B в тех случаях, когда условная вероятность считается известной (из дополнительных опытов) или определяется методом вспомогательного эксперимента.

Формула умножения для произвольного числа событий записывается следующим образом:

$$\mathrm{P}(A_1A_2\ldots A_n) = \mathrm{P}(A_1)\,\mathrm{P}(A_2/A_1)\ldots\mathrm{P}(A_n/A_1A_2\ldots A_{n-1}). \tag{11}$$

Формула (11) справедлива, если все входящие в правую часть условные вероятности определены.

Формула сложения для n слагаемых записывается в виде

$$\mathrm{P}\left(\sum_{k=1}^{n} A_k\right) = \sum_{k=1}^{n}\mathrm{P}(A_k) - \sum_{i<j}\sum \mathrm{P}(A_iA_j) + \\ + \sum_{i<j<k}\sum\sum \mathrm{P}(A_iA_jA_k) + \dots + (-1)^{n-1}\,\mathrm{P}(A_1A_2\ldots A_n). \tag{12}$$

Если события $A_1, A_2, \ldots, A_n$ независимы в совокупности, то вероятность осуществления хотя бы одного из них проще вычисляется не по формуле сложения (12), а с помощью формулы умножения:

$$\mathrm{P}(A_1 + A_2 + \cdots + A_n) = 1 - \mathrm{P}(\overline{A_1 + A_2 + \cdots + A_n}) = \\ = 1 - \mathrm{P}(\overline{A}_1)\,\mathrm{P}(\overline{A}_2)\ldots\mathrm{P}(\overline{A}_n). \quad (13)$$

Пример 14. В продукции завода брак составляет 5 % от общего количества выпускаемых деталей. Для контроля отобрано 20 деталей. Какова вероятность того, что среди них имеется хотя бы одна бракованная?

◁ Для любой детали из продукции завода вероятность быть бракованной равна по условию $p = 0{,}05 = \mathrm{P}(A_k)$, $k = 1, 2, \ldots, 20$, где событие $A_k = \{k\text{-я по счету извлеченная деталь бракованная}\}$. Очевидно, нас интересует событие $A_1 + A_2 + \cdots + A_{20}$. В условиях отлаженного технологического процесса можно считать, что события $A_1, A_2, \ldots, A_{20}$ независимы в совокупности. По формуле (13) получаем

$$\mathrm{P}(A_1 + A_2 + \cdots + A_{20}) = 1 - \prod_{k=1}^{20} \mathrm{P}(\overline{A}_k) = 1 - 0{,}95^{20} \approx 0{,}64. \ ▷$$

Пример 15. В условиях эксперимента, описанного в задаче 18.179, вычислить вероятности следующих событий: $A = \{$вышедший знает или английский или французский язык$\}$, $B = \{$вышедший не знает ни одного языка$\}$.

◁ Так как $A = E + F$, то, используя формулу сложения (10) и данные задачи 18.179, получим

$$\mathrm{P}(A) = \mathrm{P}(E) + \mathrm{P}(F) - \mathrm{P}(EF) = 0{,}5 + 0{,}4 - 0{,}2 = 0{,}7.$$

Для вычисления вероятности события $B = \overline{E}\,\overline{D}\,\overline{F}$ используем правила де Моргана и формулу сложения для трех событий:

$$\mathrm{P}(B) = \mathrm{P}(\overline{E}\,\overline{D}\,\overline{F}) = \mathrm{P}(\overline{E + D + F}) = 1 - \mathrm{P}(E + D + F) = \\ = 1 - (\mathrm{P}(E) + \mathrm{P}(D) + \mathrm{P}(F) - \mathrm{P}(ED) - \mathrm{P}(EF) - \mathrm{P}(DF) + \mathrm{P}(EDF)) = 0{,}08. \ ▷$$

18.187. Из урны, содержащей 6 белых и 4 черных шара, наудачу и последовательно извлекают по одному шару до появления черного шара. Найти вероятность того, что придется производить четвертое извлечение, если выборка производится: а) с возвращением; б) без возвращения.

18.188. В условиях эксперимента, описанного в задаче 18.176, вычислить вероятность событий BF, AF и ABF.

18.189. Только один из n ключей подходит к данной двери. Найти вероятность того, что придется опробовать ровно k ключей ($k \leqslant n$) для открывания данной двери.

18.190. Студент может уехать в институт или автобусом, который ходит через каждые 20 мин, или троллейбусом, который ходит через каждые 10 мин. Какова вероятность того, что студент, подошедший к остановке, уедет в течение ближайших пяти минут?

18.191. Вероятность попадания в мишень при одном выстреле для первого стрелка равна p_1, для второго стрелка равна p_2. Стрелки произвели по одному выстрелу в мишень. Считая попадания в мишень для отдельных стрелков событиями независимыми, найти вероятности следующих событий: $A = \{$ни одного попадания в мишень$\}$, $B = \{$ровно одно попадание в мишень$\}$.

18.192. Радист трижды вызывает корреспондента. Вероятность того, что корреспондент примет первый вызов, равна 0,2, второй — 0,3 и третий — 0,4. По условиям приема события, состоящие в том, что i-й по счету вызов ($i = 1, 2, 3$) услышан, независимы. Найти вероятность того, что корреспондент вообще услышит радиста.

18.193. Известно, что A и B — наблюдаемые события в эксперименте, причем $\mathrm{P}(B) = 0{,}4$, $\mathrm{P}(A/B\} = 0{,}3$, $\mathrm{P}(A/\overline{B}) = 0{,}2$. Найти $\mathrm{P}(A)$, $\mathrm{P}(\overline{A}\,\overline{B})$, $\mathrm{P}(\overline{A}+\overline{B})$ и $\mathrm{P}(A\triangle B)$.

18.194. В условиях эксперимента, описанного в задаче 18.176, вычислить вероятности событий $B + F$, $F - A$, $F - AB$.

18.195. В условиях задачи 18.38 найти вероятность поражения самолета, если $\mathrm{P}(D_1) = \mathrm{P}(D_2) = p_1$, $\mathrm{P}(K) = p_2$.

18.196. Статистика, собранная среди студентов одного из вузов, обнаружила следующие факты: 60 % всех студентов занимаются спортом, 40 % участвуют в научной работе на кафедрах и 20 % занимаются спортом и участвуют в научной работе на кафедрах. Корреспондент местной газеты подошел к наудачу выбранному студенту. Найти вероятности следующих событий: $A = \{$студент занимается по крайней мере одним из двух указанных видов деятельности$\}$, $B = \{$студент занимается одним только спортом$\}$, $C = \{$студент занимается только одним видом деятельности$\}$.

18.197. Студент знает 20 из 25 вопросов программы. Зачет считается сданным, если студент ответит не менее чем на три из четырех поставленных в билете вопросов. Взглянув на первый вопрос билета, студент обнаружил, что он его знает. Какова вероятность того, что студент сдаст зачет?

18.198. Студенты выполняют контрольную работу в классе контролирующих машин. Работа состоит из трех задач. Для получения положительной оценки достаточно решить две. Для каждой задачи зашифровано пять различных ответов, из которых только один правильный. Студент Иванов плохо знает материал и поэтому выбирает ответы для каждой задачи наудачу. Какова вероятность того, что он получит положительную оценку?

18.199. Наудачу подбрасывают две игральные кости. Найти вероятности следующих событий: $A = \{$сумма выпавших очков четна$\}$, $B = \{$произведение очков четно$\}$, $C = \{$на одной из костей число очков четно, а на другой нечетно$\}$, $D = \{$ни на одной из костей не выпало шесть очков$\}$.

18.200. Цех изготовляет кинескопы для телевизоров, причем 70 % всех кинескопов предназначены для цветных телевизоров и 30 % — для черно-белых. Известно, что 50 % всей продукции отправляется на экспорт, причем из общего числа кинескопов, предназначенных для цветных телевизоров, 40 % отправляется на экспорт. Найти вероятность того, что наудачу взятый для контроля кинескоп предназначен для черно-белого телевизора и будет отправлен на экспорт.

18.201. Проводится три повторных независимых измерения некоторой физической величины. Вероятность того, что при одном измерении (любом) ошибка выйдет за пределы допуска, равна 0,1. Найти вероятности следующих событий: $A = \{$во всех проведенных измерениях была достигнута заданная точность$\}$, $B = \{$не более чем в одном измерении ошибка выйдет за пределы допуска$\}$, $C = \{$по крайней мере в двух измерениях подряд была достигнута заданная точность$\}$.

18.202. По каналу связи, состоящему из передатчика, ретранслятора и приемника, передаются два сигнала: единица и нуль. Вследствие воздействия помех сигналы могут искажаться. На участке передатчик — ретранслятор единица переходит в единицу с вероятностью p_1 и в нуль с вероятностью $1 - p_1$; нуль переходит в нуль с вероятностью q_1 и в единицу с вероятностью $1 - q_1$. На участке ретранслятор — приемник вероятности указанных событий соответственно равны p_2, $1 - p_2$, q_2 и $1 - q_2$. Определить вероятность события $A = \{$кодовая комбинация 10, посланная передатчиком, принята без искажений$\}$.

18.203 (*задача де Мере*). Сколько раз нужно бросить пару игральных костей, чтобы с вероятностью, не меньшей 0,5, хотя бы один раз появилась сумма очков, равная 12?

18.204. Вероятность поражения цели при одном выстреле равна p_1. Сколько надо произвести независимых выстрелов в неизменных условиях, чтобы с вероятностью, не меньшей p_2, поразить цель хотя бы один раз? Написать общее выражение для наименьшего числа выстрелов $n(p_1, p_2)$ и найти следующие числовые значения: $n(0{,}3;\ 0{,}9)$, $n(0{,}3;\ 0{,}95)$.

18.205. Самолет состоит из трех различных по уязвимости частей: 1) кабины летчика и двигателей, 2) топливных баков, 3) планера. Для поражения самолета достаточно одного попадания в первую часть, двух попаданий во вторую часть или трех попаданий в третью. При попадании в самолет одного снаряда он с вероятностью p_k и независимо от других попадает в k-ю часть ($k = 1, 2, 3$). Самолет был обстрелян. События: $A = \{$в самолет попало 3 снаряда$\}$, $B = \{$самолет поражен$\}$. Найти условную вероятность $\mathsf{P}(B/A)$.

18.206. В тире имеются мишени двух типов: мелкие (диаметра d) и крупные (диаметра $2d$). Стреляющему обещан приз, если он из трех выстрелов по крайней мере дважды подряд поразит цель, выбирая ее каждый раз по своему усмотрению, но с обязательным условием: не стрелять дважды подряд в мишень одного и того же диаметра. С какой мишени — мелкой или крупной — следует начать состязание стреляющему, если вероятность попадания в мишень пропорциональна ее площади?

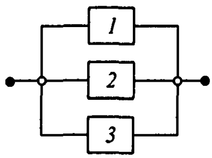
18.206 (1) (*задача Паччоли о разделе ставки*). Двое равносильных игроков играют матч до 6 побед в игру типа тенниса (где нет ничьих). Игра прервана при счете 5 : 3 в пользу одного из игроков. В каком отношении следует по справедливости разделить ставку между ними?

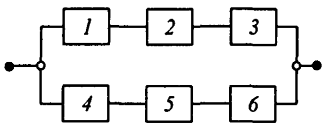
18.207. Производится стрельба из зенитного орудия по воздушной цели. Попадания при отдельных выстрелах независимы н имеют вероятность p. Если снаряд попал в цель, то она поражается с вероятностью p_1. Боевой запас орудия n снарядов. Стрельба ведется до поражения цели или до израсходования всего боезапаса. Найти вероятности следующих событий: $A = \{$не весь боезапас будет израсходован$\}$, $B = \{$останутся неизрасходованными не менее k снарядов$\}$.

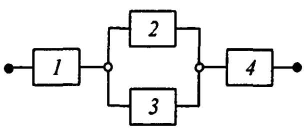
Рассмотрим работу в течение фиксированного интервала времени T некоторой физической системы, состоящей из n определенным образом соединенных элементов (деталей, узлов и т. п.). Надежность такой системы в узком смысле количественно определяется вероятностью безотказной работы в течение интервала времени T. Возникновение отказов системы определяется отказами ее элементов, и, следовательно, надежность системы зависит от надежности элементов.

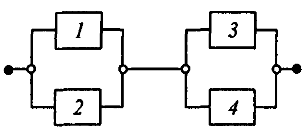
Различают параллельное соединение элементов системы (резервирование или дублирование) и последовательное. При параллельном соединении отказ системы происходит лишь при отказе всех элементов, а при последовательном — при отказе хотя бы одного элемента. Отказы отдельных элементов системы могут быть зависимыми или независимыми (в совокупности) событиями, что необходимо оговаривать при математической формализации соответствующего случайного эксперимента.

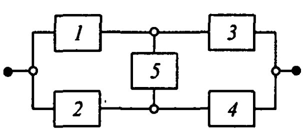
В следующих задачах приведены схемы соединения элементов, образующих цепь с одним входом и одним выходом. Предполагается, что отказы элементов являются независимыми в совокупности событиями. Считается известной надежность p_k k-го элемента (соответственно $q_k = 1 - p_k$ — вероятность его отказа). Отказ любого из элементов приводит к прерыванию сигнала в той ветви цепи, где находится данный элемент. Вычислить надежность p каждой из схем.

18.208. 

18.209. 

18.210. 

18.211. 

18.212. 

18.213. Вероятность отказа прибора после того, как он применялся k раз, равна $p(k)$. Известно, что при первых m применениях прибор не отказал. Какова вероятность того, что при следующих n применениях прибор откажет?

18.214. Иван и Петр поочередно бросают монету. Выигрывает тот, у кого раньше появится герб. Иван бросает первым. Найти вероятности p_1 и p_2 выигрыша для каждого из игроков, считая, что бросание монеты может продолжаться неограниченно долго.

18.215 (продолжение). В условиях предыдущей задачи найти указанные вероятности, считая, что игра ограничена десятью бросаниями для каждого из игроков, причем, если герб не появится у Ивана вплоть до его десятого броска, то выигравшим считается Петр. Как при этом изменились вероятности?

18.216. Установить, можно ли сделать игру из задачи 18.214 более справедливой, если позволить Петру делать большее число бросков, когда наступает его очередь. Вычислить вероятность выигрыша для Ивана, если Петру разрешается делать два броска при его подходе, а общее число бросков не ограничено.

18.217. В театральной кассе к некоторому моменту времени осталось: 1 билет в театр эстрады, 2 билета в драматический театр и 3 билета в театр комедии. Каждый очередной покупатель покупает лишь один билет с равной вероятностью в любой из возможных театров. Два человека из очереди последовательно приобрели билеты. Найти вероятности следующих событий: $A = \{$куплены билеты в разные театры$\}$, $B = \{$куплены билеты в один какой-нибудь театр$\}$.

18.218 (продолжение). В условиях эксперимента предыдущей задачи найти вероятности событий: $C = \{$все билеты в театр эстрады распроданы$\}$, $D = \{$билет в театр комедии куплен раньше, чем в театр эстрады$\}$.

18.219. За некоторый промежуток времени амеба может погибнуть с вероятностью $1/4$, выжить с вероятностью $1/4$ и разделиться на две с вероятностью $1/2$. В следующий такой же промежуток времени с каждой амебой независимо от ее «происхождения» происходит то же самое. Сколько амеб и с какими вероятностями может существовать к концу второго промежутка времени?

18.220* (*задача о рассеянной секретарше*). Некая секретарша написала n деловых писем, вложила их в конверты и по рассеянности написала адреса случайным образом. Какова вероятность p_n того, что хотя бы одно письмо попадет по назначению? Оценить p_n для $n = 5$ и $n = 10$.

18.221. Раскрывается определитель n-го порядка и наудачу выбирается слагаемое. Какова вероятность, что данное слагаемое не содержит элемента главной диагонали?

18.222. Жюри состоит из трех судей. Первый и второй судьи принимают правильное решение независимо друг от друга с вероятностью p, а третий судья для принятия решения бросает монету. Окончательное решение жюри принимает по большинству голосов. Какова вероятность того, что жюри примет правильное решение?

18.223 (продолжение). Все трое членов жюри принимают независимо друг от друга правильное решение с вероятностью p. Каким должно быть p, чтобы данное жюри принимало правильное решение с большей вероятностью, чем жюри из предыдущей задачи?

18.224 (продолжение). Первые двое судей из жюри принимают решение так же, как в условиях задачи 18.222, а третий судья поступает следующим образом: если двое первых судей принимают одинаковые решения, то он к ним присоединяется, если же решения двух первых судей разные, то третий судья бросает монету. Какова вероятность правильного решения у такого жюри?

9. Формула полной вероятности. Пусть $H_1, H_2, \ldots, H_n$ — наблюдаемые события для данного эксперимента, причем система множеств $\{H_1, H_2, \ldots, H_n\}$ образует разбиение множества Ω. Для любого наблюдаемого в эксперименте события A имеет место следующая *формула полной вероятности*:

$$P(A) = \sum_{k=1}^{n} P(H_k)\, P(A/H_k). \tag{14}$$

События H_k принято называть *гипотезами* по отношению к событию A. Безусловные вероятности $P(H_k)$ трактуются как *доопытные* (*априорные*) вероятности гипотез.

Пример 16. Партия транзисторов, среди которых 10% дефектных, поступила на проверку. Схема проверки такова, что с вероятностью 0,95 обнаруживает дефект (если он есть), и существует ненулевая вероятность 0,03 того, что исправный транзистор будет признан дефектным. Какова вероятность того, что случайно выбранный из партии транзистор будет признан дефектным?

◁ Из условия задачи очевидно, что с рассматриваемым событием $A = \{$случайно выбранный транзистор признан дефектным$\}$ тесно связаны две гипотезы: $H_1 = \{$поступивший на проверку транзистор дефектный$\}$, $H_2 = \overline{H}_1 = \{$поступивший на проверку транзистор исправный$\}$. Безусловные априорные вероятности этих гипотез легко вычисляются по классической формуле: $P(H_1) = 0{,}1$, $P(H_2) = 0{,}9$. Условные вероятности определены в условии задачи: $P(A/H_1) = 0{,}95$, $P(A/H_2) = 0{,}03$. Применяя формулу полной вероятности (14), получим

$$P(A) = 0{,}1 \cdot 0{,}95 + 0{,}9 \cdot 0{,}03 = 0{,}122. \ \triangleright$$

18.225. Прибор, установленный на борту самолета, может работать в двух режимах: в условиях нормального крейсерского полета и в условиях перегрузки при взлете и посадке. Крейсерский режим полета осуществляется в 80% всего времени полета, условия перегрузки — в 20%. Вероятность выхода прибора из строя за время полета в нормальном режиме равна 0,1, в условиях перегрузки — 0,4. Вычислить надежность прибора за время полета.

18.226. В продажу поступают телевизоры трех заводов. Продукция первого завода содержит 20% телевизоров со скрытым дефектом, второго — 10% и третьего — 5%. Какова вероятность приобрести исправный телевизор, если в магазин поступило 30% телевизоров с первого завода, 20% — со второго и 50% — с третьего?

18.227. Два цеха штампуют однотипные детали. Первый цех дает $\alpha\,\%$ брака, второй — $\beta\,\%$. Для контроля отобрано n_1 деталей из первого цеха и n_2 из второго. Эти $n_1 + n_2$ деталей смешаны в одну партию, и из нее наудачу извлекают одну деталь. Какова вероятность того, что она бракованная?

18.228. Производится n независимых выстрелов зажигательными снарядами по резервуару с горючим. Каждый снаряд попадает в резервуар с вероятностью p. Если в резервуар попал один снаряд, то горючее воспламеняется с вероятностью p_1, если два снаряда — с полной достоверностью. Найти вероятность того, что при n выстрелах горючее воспламенится.

18.229. При переливании крови надо учитывать группу крови донора и больного. Человеку, имеющему четвертую группу крови, можно перелить кровь любой группы; человеку со второй или третьей группой крови можно перелить кровь либо той же группы, либо первой; человеку с первой группой крови можно перелить только кровь первой группы. Среди населения 33,7 % имеют первую, 37,5 % — вторую, 20,9 % — третью и 7,9 % — четвертую группы крови. Найти вероятность того, что случайно взятому больному можно перелить кровь случайно взятого донора.

18.230. В условиях эксперимента, описанного в задаче 18.202, сигналы 0 и 1 передаются с равной вероятностью. Вычислить вероятность события $C = \{$принято два одинаковых символа$\}$.

18.231. На рис. 6 изображена схема дорог. Туристы выходят из пункта $П_1$, выбирая каждый раз на развилке дорог дальнейший путь наудачу. Какова вероятность, что они попадут в пункт $П_2$?

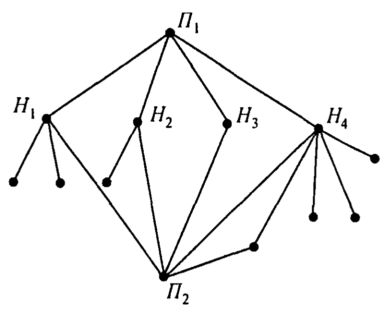


Рис. 6

18.232. Три стрелка, вероятности попадания которых при одном выстреле в мишень в неизменных условиях постоянны и соответственно равны $p_1 = 0{,}8$, $p_2 = 0{,}7$, $p_3 = 0{,}6$, делают по одному выстрелу в одну и ту же мишень. Вычислить вероятность события $A = \{$в мишени окажется ровно две пробоины$\}$, приняв в качестве гипотез элементарные исходы данного эксперимента.

18.233. В ящике лежат 20 теннисных мячей, в том числе 15 новых и 5 играных. Для игры наудачу выбираются два мяча и после игры возвращаются обратно. Затем для второй игры также наудачу извлекаются еще два мяча. Какова вероятность того, что вторая игра будет проводиться новыми мячами?

18.234. Из десяти студентов, пришедших сдавать экзамен по теории вероятностей и взявших билеты, Иванов и Петров знают 20 билетов из 30, Сидоров плохо занимался весь семестр и успел повторить только 15 билетов, остальные студенты знают все 30 билетов. По прошествии отведенного времени на подготовку экзаменатор наудачу вызывает отвечать одного из студентов. Какова вероятность того, что вызванный сдал экзамен, если знание билета гарантирует сдачу экзамена с вероятностью 0,85, а при незнании билета можно сдать экзамен лишь с вероятностью 0,1?

18.235. Шесть шаров, среди которых 3 белых и 3 черных, распределены по двум урнам. Наудачу выбирается урна, а из нее — один шар. Как нужно распределить шары по урнам, чтобы вероятность события $A = \{$вынутый шар белый$\}$ была максимальной?

18.236*. Для поиска месторождения нефти на заданной территории организовано n геологоразведочных партий, каждая из которых независимо от других обнаруживает залежь с вероятностью p. После обработки и анализа сейсмографических записей вся территория была поделена на два района. В первом районе нефть может залегать с вероятностью p_1, а во втором — с вероятностью $1 - p_1$. Как следует распределить n геологоразведочных партий по двум районам, чтобы вероятность обнаружения нефти была максимальной?

В ряде случаев выбор системы гипотез не определяется однозначно условиями эксперимента. В таких случаях предпочтение следует отдавать той системе, для которой условные вероятности вычисляются наиболее просто.

Пример 17. Из полной колоды в 52 карты наудачу последовательно и без возвращения выбирают две карты. Какова вероятность того, что второй картой можно покрыть первую? (Это значит, что вторая карта должна быть более старшей картой той же масти.)

◁ Пусть A — интересующее нас событие. В качестве первой попытки выберем следующие гипотезы: $H_k = \{$в составе двух вынутых карт ровно k картинок$\}$, $k = 0, 1, 2$ (к «картинкам» относятся валет, дама, король и туз каждой масти).

Ясно, что $\{H_k\}$ $(k = 0, 1, 2)$ — разбиение множества Ω, так как выполнены все три условия, сформулированные в определении разбиения множества (проверьте!). Кроме того, нетрудно вычислить безусловные вероятности гипотез $\mathrm{P}(H_k)$.

Однако вычисление условных вероятностей $\mathrm{P}(A/H_k)$ оказывается делом не менее трудным, чем ответ на первоначально поставленный вопрос о вероятности события $\mathrm{P}(A)$. Это объясняется тем, что связь события A с данными гипотезами H_k не может быть достаточно просто описана на языке алгебраических операций.

Рассмотрим более удачный для решения задачи вариант разбиения $\{H_k\}$ $(k = 2, 3, \ldots, 14)$, где $H_k = \{$первая вытянутая карта оценивается в k очков$\}$, при этом значению $k = 2$ соответствует двойка, $k = 3$ — тройка, ..., $k = 11$ — валет, $k = 12$ — дама, $k = 13$ — король и $k = 14$ — туз. Вычислим условные вероятности, применяя метод вспомогательного эксперимента.

$\mathrm{P}(A/H_k) = \mathrm{P}\{$вторая вытянутая карта той же масти, причем ее достоинство оценивается не ниже, чем в $k+1$ очко$\} = \dfrac{14-k}{51}$ в силу формулы классической вероятности. Безусловные вероятности гипотез $\mathrm{P}(H_k) = \dfrac{1}{13}$ в силу равновероятности событий $H_k = \{$вытянуть карту произвольной масти, оцениваемую в k очков$\}$. Применяя формулу полной вероятности (14), получим

$$\mathrm{P}(A) = \sum_{k=2}^{14} \mathrm{P}(H_k)\,\mathrm{P}(A/H_k) = \frac{1}{13}\sum_{k=2}^{14}\frac{14-k}{51} = \frac{2}{17}.$$

Еще более простой путь решения получается, если ввести следующее разбиение множества Ω в данном эксперименте: $\{H_k\}$ $(k = 1, 2)$, где $H_1 = \{$обе вынутые карты одной масти$\}$, $H_2 = \{$две карты разной масти$\}$. Очевидно, что $\mathrm{P}(A/H_2) = 0$, поэтому вторая гипотеза исключается из формулы полной вероятности и $\mathrm{P}(H_1) = \dfrac{12}{51} = \dfrac{4}{17}$ (первая карта может быть произвольной масти, вторая должна быть той же масти, что и первая). Пусть теперь выполнено событие H_1, т.е. обе карты одной масти. Тогда та из них, которую извлекали второй по счету, должна быть старше первой. Но в силу равновероятности исходов $\omega_1 = \{$вторая карта старше первой$\}$ и $\omega_2 = \{$первая карта старше второй$\}$ в этом вспомогательном эксперименте получаем $\mathrm{P}(A/H_1) = 0{,}5$, поэтому

$$\mathrm{P}(A) = \mathrm{P}(H_1)\,\mathrm{P}(\mathrm{A/H_1}) = \frac{4}{17}\cdot\frac{1}{2} = \frac{2}{17}. \ \triangleright$$

18.237. Программа экзамена содержит 30 различных вопросов, из которых студент Иванов знает только 15. Для успешной сдачи экзамена достаточно ответить на 2 предложенных вопроса или на один из них и на дополнительный вопрос. Какова вероятность того, что Иванов успешно сдаст экзамен?

18.238. Из множества чисел $E = \{1, 2, \ldots, n\}$ наудачу последовательно и без возвращения извлекают два числа. Какова вероятность того, что первое число больше второго не менее, чем на m $(0 < m \leqslant n-1)$?

18.239. На шахматную доску ставят наудачу двух слонов, белого и черного. Какова вероятность того, что слоны побьют друг друга?

18.239 (1). Из полного набора домино, содержащего числа 0, 1, 2, ..., n наудачу отбирают 2 кости. Какова вероятность, что их можно приставить друг к другу?

18.240.** В урне находится 7 белых и 3 черных шара. Три игрока по очереди извлекают по одному шару, отмечают цвет и возвращают шар обратно. Выигрывает тот, кто первым достанет черный шар. Найти вероятность выигрыша для каждого из игроков, если игра может продолжаться неограниченно.

18.241. Студент Иванов знает только 10 из 25 экзаменационных билетов. В каком случае шансы Иванова получить известный ему билет выше: когда он подходит тянуть билет первым или вторым по счету?

10. Формула Байеса. Пусть $\{H_1, H_2, \ldots, H_n\}$ — разбиение множества Ω для данного эксперимента, интерпретируемое как совокупность гипотез по отношению к интересующему нас событию A. Пусть эксперимент проведен и стало известно, что событие A осуществилось. Какова *послеопытная* (*апостериорная*) вероятность осуществления гипотезы H_k при условии, что событие A имело место? Ответ дается *формулой Байеса*

$$\mathrm{P}(H_k/A) = \frac{\mathrm{P}(H_k)\,\mathrm{P}(A/H_k)}{\mathrm{P}(A)}, \tag{15}$$

где $\mathrm{P}(A) = \sum\limits_{i=1}^{n} \mathrm{P}(H_i)\,\mathrm{P}(A/H_i)$ — полная вероятность осуществления события A.

Формула Байеса позволяет «переоценить» вероятность каждой из гипотез после поступления новой «информации» относительно осуществления тех или иных наблюдаемых событий.

Пример 18. В условиях эксперимента, описанного в примере 16, случайно выбранный из партии транзистор был признан дефектным. Какова вероятность того, что на самом деле транзистор исправен?

◁ В обозначениях примера 16 требуется вычислить $\mathrm{P}(H_2/A)$ (апостериорную условную вероятность гипотезы H_2). По формуле Байеса

$$\mathrm{P}(H_2/A) = \frac{\mathrm{P}(H_2)\,\mathrm{P}(A/H_2)}{\mathrm{P}(A)} = \frac{0{,}9 \cdot 0{,}03}{0{,}122} \approx 0{,}221.$$

Таким образом, апостериорная условная вероятность того, что транзистор на самом деле исправный, если известно, что он был признан дефектным, существенно меньше априорной вероятности гипотезы H_2, что явилось следствием поступившей информации. ▷

18.242. В урне лежит шар неизвестного цвета — с равной вероятностью белый или черный. В урну опускается один белый шар и после тщательного перемешивания наудачу извлекается один шар. Он оказался белым. Какова вероятность того, что в урне остался белый шар?

18.243. На вход радиолокационного устройства с вероятностью 0,8 поступает смесь полезного сигнала с помехой, а с вероятностью 0,2 — только помеха. Если поступает полезный сигнал с помехой, то устройство регистрирует наличие какого-то сигнала с вероятностью 0,7; если только помеха, — то с вероятностью 0,3. Известно, что устройство зарегистрировало наличие какого-то сигнала. Найти вероятность того, что в его составе есть полезный сигнал.

18.244. Прибор состоит из двух последовательно включенных узлов. Надежность (вероятность безотказной работы в течение времени T) первого узла равна 0,9, второго — 0,8. За время испытания прибора в течение времени T зарегистрирован отказ прибора. Найти вероятности следующих событий: $A_1 = \{$отказал только первый узел$\}$, $A_2 = \{$отказали оба узла$\}$.

18.245. В условиях эксперимента, описанного в примере 12, найти условную вероятность $\mathrm{P}(A/B)$, считая известной условную вероятность $\mathrm{P}(B/A)$ и применяя формулу Байеса.

18.246. В коробке находятся две неотличимые по внешнему виду и по весу игральные кости: одна правильная, с одинаковыми вероятностями выпадения всех шести цифр при случайном подбрасывании; другая неправильная, с неравномерным распределением массы по объему. При случайном подбрасывании неправильной игральной кости шестерка появляется с вероятностью 1/3, единица — с вероятностью 1/9, остальные цифры выпадают с одинаковой вероятностью. Наудачу извлеченная из коробки игральная кость была подброшена, и в результате выпало 6 очков. Найти вероятность того, что была подброшена правильная игральная кость.

18.247. Три стрелка производят по одному выстрелу в одну и ту же мишень. Вероятности попадания в мишень при одном выстреле для каждого из стрелков соответственно равны p_1, p_2 и p_3. Какова вероятность того, что второй стрелок промахнулся, если после выстрелов в мишени оказалось две пробоины?

18.248. Однотипные приборы выпускаются тремя заводами в количественном отношении $n_1 : n_2 : n_3$, причем вероятности брака для этих заводов соответственно равны p_1, p_2 и p_3. Прибор, приобретенный научно-исследовательским институтом, оказался бракованным. Какова вероятность того, что данный прибор произведен первым заводом (марка завода на приборе отсутствует)?

18.249. Число бракованных микросхем на 1000 априори считается равновозможным от 0 до 3. Наудачу опробованы 100 микросхем, оказавшихся исправными. Какова вероятность, что все схемы исправны?

18.250. В группе из 25 человек, пришедших сдавать экзамен по теории вероятностей, имеется 10 отличников, 7 подготовленных хорошо, 5 — удовлетворительно и 3 человека плохо подготовлены. Отличники знают все 25 вопросов программы, хорошо подготовленные — 20, подготовленные удовлетворительно — 15 и плохо подготовленные знают лишь 10 вопросов. Вызванный наудачу студент ответил на два заданных вопроса. Найти апостериорные вероятности гипотез: $H_1 = \{$студент подготовлен отлично или хорошо$\}$, $H_2 = \{$студент подготовлен удовлетворительно$\}$, $H_3 = \{$студент подготовлен плохо$\}$.

Особое значение приобретает формула Байеса для таких экспериментов, в которых гипотезы H_k непосредственно не наблюдаемы, хотя априорные вероятности $\mathrm{P}(H_k)$ и соответствующие условные вероятности $\mathrm{P}(A/H_k)$, $k = 1, \ldots, n$, известны из дополнительных опытов. Такая ситуация может иметь место, например, если отсутствует прибор, позволяющий регистрировать факт осуществления данных гипотез, или же если применение прибора для регистрации осуществления гипотез приводит к разрушению предмета наблюдения (разрушающий контроль). Для подобных экспериментов переоценка вероятностей гипотез после опыта может быть проведена на основании наблюдаемого события A, тесно связанного с гипотезами. Такой подход часто используется в задачах медицинской и технической диагностики.

Пример 19. Изучается три вида дефектов запоминающих устройств, выполненных на интегральных схемах: дефекты схем обрамления (гипотеза H_1, $\mathrm{P}(H_1) = 0,1$), дефекты, вызванные паразитными связями между ячейками (гипотеза H_2, $\mathrm{P}(H_2) = 0,6$), и дефекты адресных шин (гипотеза H_3, $\mathrm{P}(H_3) = 0,3$). Диагностика запоминающих устройств производится с помощью набора тестов $T_1, T_2, \ldots, T_n$, каждый из которых проверяет определенное состояние ячейки памяти. Наблюдаемый результат — состояние выбранной ячейки по отношению к каждому тесту. Пусть диагностика произведена, и наблюдался некоторый результат (произошло событие A). Известно до опыта, что

$$\mathrm{P}(A/H_1) = 0,4, \quad \mathrm{P}(A/H_2) = 0,2, \quad \mathrm{P}(A/H_3) = 0,3.$$

Установить, какая из гипотез имеет наибольшую апостериорную вероятность (т.е. какой из дефектов наиболее вероятен).

$\lhd$ Вычислим сначала полную вероятность осуществления события A. Используя данные задачи и применяя формулу (14), получим

$$\mathrm{P}(A) = \sum_{k=1}^{3} \mathrm{P}(H_k)\,\mathrm{P}(A/H_k) = 0,25.$$

Для ответа на вопрос, какой из дефектов имеет наибольшую апостериорную вероятность, заметим, что в числителе формулы Байеса стоит слагаемое полной вероятности, относящееся к данной переоцениваемой гипотезе. Сравнивая эти слагаемые для трех заданных гипотез, получаем, что наибольшую вероятность имеет гипотеза H_2: $\mathrm{P}(H_2/A) = \dfrac{0{,}12}{0{,}25} = 0{,}48$. ▷

18.251. Астрономический объект, за которым ведется наблюдение, может находиться в одном из двух состояний: H_1 или H_2. Априорные вероятности этих состояний $\mathrm{P}(H_1) = 0{,}6$, $\mathrm{P}(H_2) = 0{,}4$. Наблюдение ведется независимо двумя обсерваториями. Первая обсерватория обычно дает правильные сведения о состоянии наблюдаемого объекта в 90 % случаев, а в 10 % ошибается; вторая дает правильные сведения в 80 % случаев, а в 20 % ошибается. Первая обсерватория сообщила, что объект находится в состоянии H_1, а вторая — что в состоянии H_2. Найти апостериорную вероятность состояния H_1.

18.252. Расследуются причины неудачного запуска космической ракеты, о котором можно высказать четыре предположения (гипотезы) H_1, H_2, H_3 или H_4. По данным статистики $\mathrm{P}(H_1) = 0{,}2$, $\mathrm{P}(H_2) = 0{,}4$, $\mathrm{P}(H_3) = 0{,}3$, $\mathrm{P}(H_4) = 0{,}1$. В ходе расследования обнаружено, что при запуске произошла утечка топлива (событие A). Условные вероятности события A согласно той же статистике равны: $\mathrm{P}(A/H_1) = 0{,}9$, $\mathrm{P}(A/H_2) = 0{,}4$, $\mathrm{P}(A/H_3) = 0{,}2$, $\mathrm{P}(A/H_4) = 0{,}3$. Какая из гипотез наиболее вероятна при данных условиях?

18.253. По каналу связи передается цифровой текст, содержащий только три цифры 1, 2, 3, которые могут появляться в тексте с равной вероятностью. Каждая передаваемая цифра в силу наличия шумов принимается правильно с вероятностью p и с вероятностью $\dfrac{1}{2}(1-p)$ принимается за какую-либо другую цифру. Предполагается, что цифры искажаются независимо. Найти вероятность того, что было передано 111, если принято 123.

18.254. Методом тестирования отыскивается неисправность в арифметическом устройстве вычислительной машины. Можно считать, что есть 4 шанса из 5, что неисправность сосредоточена в одном из 8 микропроцессоров, с равной вероятностью — в любом из них. Были испытаны 7 из этих микропроцессоров, но неисправность не обнаружена. Какова вероятность обнаружить неисправность в последнем из восьми микропроцессоров?

18.255. Предположим, что надежность определения туберкулеза при рентгеновском просвечивании грудной клетки составляет 90 % (т.е. 10 % носителей туберкулеза остаются неопознанными). Вероятность того, что у здорового человека будет ошибочно определен туберкулез, составляет 1 %. Просвечиванию была подвергнута большая группа людей со средним процентом больных, равным 0,1 %. Какова вероятность того, что человек, признанный больным, действительно является носителем туберкулеза?

18.256. Противотанковая батарея состоит из 10 орудий, причем для первой группы из шести орудий вероятности того, что при одном выстреле произойдет недолет, попадание или перелет, равны соответственно 0,1; 0,7; 0,2. Для каждого из остальных четырех орудий вероятности тех же самых событий равны соответственно 0,2; 0,6 и 0,2. Наудачу выбранное орудие произвело три выстрела по цели, в результате чего было зафиксировано одно попадание, один недолет и один перелет. Какова вероятность того, что стрелявшее орудие принадлежит первой группе?

§ 2. Случайные величины

1. Законы распределения и числовые характеристики случайных величин. *Случайной величиной* X называется действительная функция $X = X(\omega)$, определенная на множестве элементарных исходов Ω и такая, что при любом действительном x множество тех ω, для которых $X(\omega) < x$, принадлежит алгебре событий для данного эксперимента.

Случайные величины принято обозначать большими буквами латинского алфавита, а их возможные значения — соответствующими малыми буквами.

Функция $F_X(x)$ действительной переменной x, $-\infty < x < \infty$, определяемая формулой

$$F_X(x) = \mathrm{P}\{X < x\}, \tag{1}$$

называется *функцией распределения* случайной величины X. Функция распределения обладает следующими свойствами:

1. $0 \leqslant F_X(x) \leqslant 1$, $-\infty < x < \infty$;
2. $F_X(-\infty) = 0$, $F_X(+\infty) = 1$;
3. $F_X(x)$ — неубывающая функция на всей оси;
4. $F_X(x)$ непрерывная слева, т.е.

$$\lim_{x \to x_0 - 0} F_X(x) = F_X(x_0).$$

Вероятность попадания случайной величины X на произвольный интервал действительной оси $[x_1, x_2)$ определяется формулой

$$\mathrm{P}\{x_1 \leqslant X < x_2\} = F_X(x_2) - F_X(x_1). \tag{2}$$

Различают случайные величины дискретного типа (сокращенно С. В. Д. Т.) и случайные величины непрерывного типа (сокращенно С. В. Н. Т.).

Случайная величина X называется *случайной величиной дискретного типа*, если множество ее возможных значений $\{x_1, x_2, \ldots\}$ конечно или счетно.

Пусть $\mathrm{P}\{X = x_k\} = p_k > 0$, $\sum\limits_k p_k = 1$, где суммирование распространяется на все возможные значения k.

Перечень всех возможных значений С. В. Д. Т. и соответствующих этим значениям вероятностей называется *законом распределения случайной величины дискретного типа.*

Зная закон распределения С. В. Д. Т., можно вычислить функцию распределения, представляющую собой, в силу определения (1), функцию накопленных вероятностей:

$$F_X(x) = \sum_{x_i < x} \mathrm{P}\{X = x_i\},$$

где суммирование распространяется на все значения индекса i, для которых $x_i < x$. Из этой формулы, в частности, следует, что

$$F_X(x_k + 0) - F_X(x_k) = \mathrm{P}\{X = x_k\}, \quad x_k \in \{x_1, x_2, \ldots\},$$

т.е. функция распределения С. В. Д. Т. испытывает скачки в точках x, для которых существует положительная вероятность события $\{X = x\}$.

Случайная величина X называется *случайной величиной непрерывного типа* (сокращенно С. В. Н. Т.), если существует такая неотрицательная, интегрируемая по Риману в бесконечных пределах функция $f_X(x)$, называемая *плотностью распределения вероятностей*, что при всех $x \in R$

$$F_X(x) = \mathrm{P}\{X < x\} = \int_{-\infty}^{x} f_X(t)\,dt.$$

Плотность распределения вероятностей обладает следующими свойствами:

1. $f_X(x) \geqslant 0$, $-\infty < x < \infty$;
2. $\int\limits_{-\infty}^{+\infty} f_X(x)\,dx = 1$ (условие нормировки);
3. $\dfrac{d}{dx} F_X(x) = f_X(x)$ в точках непрерывности функции $f_X(x)$.

Функция распределения С. В. Н. Т. $F_X(x)$ является непрерывной монотонно возрастающей функцией на всей оси, причем $\mathrm{P}\{X = x\} = \lim\limits_{\delta x \to 0} (F_X(x + \delta x) - F_X(x)) = 0$ при всех $x \in R$. Это значит, что вероятность «попасть в точку» для С. В. Н. Т. равна нулю.

Если X — С. В. Н. Т., то вероятность ее попадания на интервал может быть вычислена как через функцию распределения по формуле (2), так и через плотность распределения вероятностей:

$$\mathrm{P}\{x_1 \leqslant X < x_2\} = \int_{x_1}^{x_2} f_X(x)\,dx. \tag{3}$$

Случайные величины, помимо законов распределения, могут также описываться *числовыми характеристиками*, среди которых различают *характеристики положения* (математическое ожидание, мода, медиана и др.) и *характеристики рассеивания* (дисперсия, среднеквадратичное отклонение, различные моменты распределения порядка выше первого и др.).

Математическим ожиданием (*средним значением по распределению*) называется действительное число, определяемое в зависимости от типа случайной величины X формулой

$$m_x = \mathrm{M}[X] = \begin{cases} \sum\limits_k x_k p_k, & \text{если} \quad X - \text{С. В. Д. Т.}, \\ \int\limits_{-\infty}^{+\infty} x f_x(x)\,dx, & \text{если} \quad X - \text{С. В. Н. Т.} \end{cases} \tag{4}$$

Математическое ожидание существует, в частности, если ряд (соответственно интеграл) в правой части формулы (4) сходится абсолютно.

Модой случайной величины X *непрерывного типа* называется действительное число d_x, определяемое как точка максимума плотности распределения вероятностей $f_x(x)$.

Мода случайной величины дискретного типа определяется как такое возможное значение x_m, для которого

$$\mathrm{P}\{X = x_m\} = \max_k \mathrm{P}(X = x_k\}.$$

Таким образом, мода С. В. Д. Т. есть ее наиболее вероятное значение в случае, если такое значение единственно. Мода может не существовать, иметь единственное значение (*унимодальное распределение*) или иметь множество значении (*мультимодальное распределение*).

Медианой случайной величины X *непрерывного типа* называется действительное число h_x, удовлетворяющее условию

$$\mathrm{P}\{X < h_x\} = \mathrm{P}\{X \geqslant h_x\},$$

т.е. корень уравнения

$$F_x(x) = \frac{1}{2}.$$

Так как данное уравнение может иметь множество корней, то медиана определяется, вообще говоря, неоднозначно.

Дисперсией случайной величины X называется неотрицательное число $\mathrm{D}[X] = D_x$, определяемое формулой

$$D_x = \mathrm{M}[(X - m_x)^2] = \begin{cases} \sum\limits_k (x_k - m_x)^2 p_k, & \text{если} \quad X - \text{С. В. Д. Т.}, \\ \int\limits_{-\infty}^{+\infty} (x - m_x)^2 f_x(x)\,dx, & \text{если} \quad X - \text{С. В. Н. Т.} \end{cases}$$

Дисперсия существует, если ряд (соответственно интеграл) в правой части равенства сходится. Неотрицательное число $\sigma_X = \sqrt{D_X}$ называется *среднеквадратичным отклонением* случайной величины X. Оно имеет размерность случайной величины X и определяет некоторый стандартный среднеквадратичный интервал рассеивания, симметричный относительно математического ожидания. (Величину σ_X иногда называют стандартным отклонением.) Если величина $X = \text{const}$ (т.е. X не случайна), то $\mathrm{D}[X] = 0$.

Случайная величина X называется *центрированной* (обозначается $\mathring{X}$), если $m_X = 0$. Случайная величина X называется *стандартизованной*, если $m_X = 0$ и $\sigma_X = 1$.

Начальным моментом m-го порядка ($m = 0, 1, 2, \dots$) *распределения случайной величины* X (если он существует) называется действительное число α_m, определяемое по формуле

$$\alpha_m = \mathrm{M}[X^m] = \begin{cases} \sum\limits_k x_k^m p_k, & \text{если } X - \text{С. В. Д. Т.,} \\ \int\limits_{-\infty}^{+\infty} x^m f_X(x)\,dx, & \text{если } X - \text{С. В. Н. Т.} \end{cases}$$

Центральным моментом m-го порядка распределения случайной величины X (если он существует) называется число μ_m, определяемое по формуле

$$\mu_m = \mathrm{M}[(X - m_X)^m] = \begin{cases} \sum\limits_k (x_k - m_X)^m p_k, & \text{если } X - \text{С. В. Д. Т.,} \\ \int\limits_{-\infty}^{+\infty} (x - m_X)^m f_X(x)\,dx, & \text{если } X - \text{С. В. Н. Т.} \end{cases}$$

Из определений моментов, в частности, следует, что

$$\alpha_0 = \mu_0 = 1, \quad m_X = \alpha_1, \quad D_X = \sigma_X^2 = \mu_2 = \alpha_2 - m_X^2.$$

Отметим еще две важные характеристики распределения, связанные с моментами высшего порядка:

$a_X = \dfrac{\mu_3}{\sigma_X^3}$ (*коэффициент асимметрии* или *«скошенности»* распределения),

$e_X = \dfrac{\mu_4}{\sigma_X^4} - 3$ (*коэффициент эксцесса* или *«островершинности»* распределения).

Квантилью порядка p (*симметричной квантилью порядка p*) распределения случайной величины X непрерывного типа называется действительное число t_p (действительное число $\hat{t}_p$), удовлетворяющее урав-

нению
$$\mathrm{P}\{X < t_p\} = p \quad (\mathrm{P}\{|X| < \hat{t}_p] = p).$$
В частности, из определения медианы следует, что $h_X = t_{0,5}$.

Критической точкой порядка p (*симметричной критической точкой порядка* p) распределения случайной величины X непрерывного типа называется действительное число $\varkappa_p$ ($\hat{\varkappa}_p$), удовлетворяющее уравнению
$$\mathrm{P}\{X \geqslant \varkappa_p\} = p \quad (\mathrm{P}\{|X| \geqslant \hat{\varkappa}_p] = p).$$

Квантиль и критическая точка одного и того же распределения связаны простым соотношением:
$$\varkappa_p = t_{1-p} \quad (\hat{\varkappa}_p = \hat{t}_{1-p}).$$

Пример 1. Трижды подбрасывается правильная монета. Случайная величина X — число выпавших гербов. Описать закон распределения данной случайной величины, вычислить функцию распределения и числовые характеристики m_X, D_X и d_X.

◁ Очевидно, что X — С. В. Д. Т., причем ее возможные значения составляют множество $\{0, 1, 2, 3\}$. Для вычисления вероятностей событий $\{X = n\}$ воспользуемся тем, что случайная величина X является функцией, определенной на множестве элементарных исходов случайного эксперимента. В данном случае можем записать $\Omega = \{\overline{\omega}_1\overline{\omega}_2\overline{\omega}_3, \omega_1\overline{\omega}_2\overline{\omega}_3, \overline{\omega}_1\omega_2\overline{\omega}_3, \overline{\omega}_1\overline{\omega}_2\omega_3, \overline{\omega}_1\omega_2\omega_3, \omega_1\overline{\omega}_2\omega_3, \omega_1\omega_2\overline{\omega}_3, \omega_1\omega_2\omega_3\}$, где $\omega_i = \{$при i-м подбрасывании монеты выпал герб$\}$, $i = 1, 2, 3$. Заметим, что $N(\Omega) = 2^3$, что соответствует схеме выбора с возвращением и упорядочиванием. Очевидно, что в силу независимости исходов отдельных подбрасываний
$$\mathrm{P}\{X = 0\} = \mathrm{P}\{\overline{\omega}_1\overline{\omega}_2\overline{\omega}_3\} = \mathrm{P}(\overline{\omega}_1)\,\mathrm{P}(\overline{\omega}_2)\,\mathrm{P}(\overline{\omega}_3) = \frac{1}{8}.$$

Для вычисления вероятности события $\{X = 1\}$ используем аксиому сложения и формулу умножения для независимых событий:
$$\mathrm{P}\{X = 1\} = \mathrm{P}\{\omega_1\overline{\omega}_2\overline{\omega}_3 + \overline{\omega}_1\omega_2\overline{\omega}_3 + \overline{\omega}_1\overline{\omega}_2\omega_3\} = 3 \cdot \frac{1}{8}.$$

Аналогично находим
$$\mathrm{P}\{X = 2\} = \mathrm{P}\{\overline{\omega}_1\omega_2\omega_3 + \omega_1\overline{\omega}_2\omega_3 + \omega_1\omega_2\overline{\omega}_3\} = \frac{3}{8}.$$

Наконец,
$$\mathrm{P}\{X = 3\} = \mathrm{P}\{\omega_1\omega_2\omega_3\} = \frac{1}{8}.$$

Для большей наглядности закон распределения случайной величины может быть представлен следующей таблицей:

x_i	0	1	2	3
$\mathrm{P}\{X = x_i\}$	1/8	3/8	3/8	1/8

Вычислим функцию распределения. Согласно определению

$$F_X(x) = \mathrm{P}\{X < x\} = \sum_{\substack{k \\ x_k < x}} \mathrm{P}\{X = x_k\} = \sum_{\substack{k \\ k < x}} \mathrm{P}\{X = k\}.$$

Подставляя сюда найденные выше вероятности, находим

$$F_X(x) = \begin{cases} 0, & \text{если} \quad x \leqslant 0, \\ 1/8, & \text{если} \quad 0 < x \leqslant 1, \\ 1/2, & \text{если} \quad 1 < x \leqslant 2, \\ 7/8, & \text{если} \quad 2 < x \leqslant 3, \\ 1, & \text{если} \quad 3 < x. \end{cases}$$

Найдем среднее значение m_X и дисперсию D_X заданного распределения. По определению математического ожидания С. В. Д. Т.

$$m_X = \sum_{n=0}^{3} n\mathrm{P}\{X = n\} = 3/2.$$

Дисперсию удобнее вычислять через второй начальный момент, используя формулу (см. задачу 18.257) $D_X = \alpha_2 - m_X^2$. Имеем

$$\alpha_2 = \sum_{n=0}^{3} n^2\mathrm{P}\{X = n\} = 3$$

и, таким образом,

$$D_X = 3 - \frac{9}{4} = \frac{3}{4}.$$

Из таблицы распределения усматриваем, что существуют два значения $x_2 = 1$ и $x_3 = 2$ таких, что

$$\mathrm{P}\{X = x_2\} = \mathrm{P}\{X = x_3\} = \max_k \mathrm{P}\{X = x_k\}.$$

Таким образом, данное распределение бимодально. $\triangleright$

Пример 2. Случайная величина X непрерывного типа может принимать ненулевые значения только на отрезке $[-1, 1]$, причем функция распределения вероятностей имеет квадратичную зависимость от x на этом отрезке. Написать выражения для $F_X(x)$ и $f_X(x)$.

◁ Согласно условию задачи и свойству 2 (с. 56) функция распределения вероятностей должна иметь вид

$$F_X(x) = \begin{cases} 0, & \text{если} \quad x \leqslant -1, \\ ax^2 + bx + c, & \text{если} \quad -1 < x \leqslant 1, \\ 1, & \text{если} \quad x > 1. \end{cases}$$

Однако входящие в правую часть числовые константы a, b и c не могут быть произвольными. Они должны быть такими, чтобы выполнялись все свойства функции распределения.

Согласно определению С. В. Н. Т. функция распределения должна быть непрерывна на всей оси. В нашем случае для этого достаточно потребовать выполнения условий непрерывности $F_X(x)$ в точках $x = -1$ и $x = 1$, что приводит к системе уравнений

$$\begin{cases} F_X(-1+0) = a - b + c = F_X(-1-0) = 0, \\ F_X(1-0) = a + b + c = F_X(1+0) = 1. \end{cases}$$

Из этой системы следует, что $b = \dfrac{1}{2}$, $a = \dfrac{1}{2} - c$, $c \in R$. Так как функция $F_X(x)$ дифференцируема всюду на интервале $(-1, 1)$, то требование монотонного неубывания ее на всей оси эквивалентно условию

$$f_X(x) = F_X'(x) \geqslant 0 \quad \text{при} \quad x \in (-1, 1),$$

т.е. условию

$$(1 - 2c)\,x + \frac{1}{2} \geqslant 0 \quad \text{при} \quad x \in (-1, 1).$$

Последнее условие, как нетрудно убедиться, выполняется лишь при

$$\frac{1}{4} \leqslant c \leqslant \frac{3}{4}.$$

Таким образом, функция распределения случайной величины принадлежит однопараметрическому семейству вида

$$F_X(x/c) = \begin{cases} 0, & \text{если} \quad x \leqslant -1, \\ \left(\dfrac{1}{2} - c\right) x^2 + \dfrac{1}{2}x + c, & \text{если} \quad -1 < x \leqslant 1, \\ 1, & \text{если} \quad x > 1, \end{cases}$$

где $1/4 \leqslant c \leqslant 3/4$.

Так как производная функции $F_X(x/c)$ существует всюду, кроме точек $x = -1$ и $x = 1$, где она терпит разрыв первого рода, то полагаем

$$f_X(x) = \frac{dF_X(x/c)}{dx} = \begin{cases} 0, & \text{если} \quad x \leqslant -1, \\ (1-2c)\,x + \dfrac{1}{2}, & \text{если} \quad -1 < x \leqslant 1, \\ 1, & \text{если} \quad x > 1, \end{cases}$$

причем $1/4 \leqslant c \leqslant 3/4$.

Нетрудно убедиться (проверьте!), что все свойства функции плотности распределения в этом случае выполняются. ▷

18.257. Доказать, что для дисперсии σ_X^2 случайной величины X справедлива формула

$$\sigma_X^2 = \alpha_2 - m_X^2.$$

18.258. Закон распределения случайной величины X дискретного типа задан следующей таблицей:

x_i	1	2	3	4
p_i	1/16	1/4	1/2	3/16

Найти m_X и $\mathrm{P}\{X > 2\}$.

18.259 (продолжение). Для случайной величины из предыдущей задачи найти D_X и d_X.

18.260 (продолжение). Для случайной величины из задачи 18.258 построить график функции распределения $F_X(x)$.

18.261. Производится один опыт, в результате которого событие A может появиться с вероятностью p и не появиться с вероятностью $q = 1 - p$. Пусть X — *индикаторная случайная величина* — принимает значение 1, если событие A произошло, и значение 0, если событие A не имело места. Описать закон распределения случайной величины X, функцию распределения, вычислить математическое ожидание.

18.262 (продолжение). Для случайной величины из предыдущей задачи найти дисперсию, третий центральный момент и определить значение вероятности p, при котором дисперсия максимальна.

18.263. В условиях примера 2 найти константу c, если плотность распределения вероятностей непрерывна в точке $x = 1$, и изобразить график функции распределения $F_X(x)$. Вычислить для полученного распределения m_X и h_X.

18.264 (продолжение). Для случайной величины из предыдущей задачи вычислить $\mathrm{P}\{X > 0\}$ и $\mathrm{P}\{-1/2 < X \leqslant 1/2\}$.

18.265. Функция распределения случайной величины X дискретного типа имеет следующий вид:

$$F_X(x) = \begin{cases} 0, & \text{если} \quad x \leqslant 2, \\ 0{,}3, & \text{если} \quad 2 < x \leqslant 3, \\ 0{,}5, & \text{если} \quad 3 < x \leqslant 4, \\ 1, & \text{если} \quad x > 4. \end{cases}$$

Вычислить $\mathrm{P}\,(X \geqslant 3{,}5\}$ и $\mathrm{P}\,\{|X| < 2{,}5\}$.

18.266 (продолжение). Описать закон распределения случайной величины X из предыдущей задачи и найти m_X и D_X.

18.267. Два стрелка независимо друг от друга делают по одному выстрелу в мишень. Вероятность попадания в мишень для первого стрелка p_1, для второго p_2. Случайная величина X — суммарное число попаданий в мишень в данном эксперименте. Описать закон распределения данной случайной величины и найти m_X и D_X.

18.268. Один раз брошены три одинаковые игральные кости. Случайная величина X принимает значение 1, если хотя бы на одной игральной кости выпадет цифра шесть; принимает значение 0, если шестерка не выпала ни на одной грани, но хотя бы на одной из граней появилась цифра 5, и принимает значение -1 в остальных случаях. Описать закон распределения случайной величины X, вычислить функцию распределения и найти математическое ожидание и моду распределения.

18.269. Случайная величина X распределена по закону, определяемому плотностью распределения вероятностей вида

$$f_X(x) = \begin{cases} c\cos x, & \text{если} \quad -\pi/2 \leqslant x \leqslant \pi/2, \\ 0, & \text{если} \quad |x| > \pi/2. \end{cases}$$

Найти константу c, вычислить $\mathrm{P}\,\{|X| < \pi/4\}$, m_X и D_X.

18.270 (продолжение). Квантилью какого порядка для данного в предыдущей задаче распределения является точка $x = \pi/4$?

18.271. Функция распределения С. В. Н. Т. X задана в виде

$$F_X(x) = \begin{cases} 0, & \text{если} \quad x \leqslant 0, \\ x^2/4, & \text{если} \quad 0 < x \leqslant 2, \\ 1, & \text{если} \quad x > 2. \end{cases}$$

Вычислить $\mathrm{P}\,\{X \geqslant 1\}$, m_X, h_X, D_X.

18.272. Производятся последовательные независимые испытания пяти приборов на надежность. Надежность каждого из приборов равна p. Каждый следующий прибор испытывается только в том случае, когда предыдущий оказался надежным. Описать закон распределения случайной величины X — числа испытанных в данном эксперименте приборов — и вычислить d_X и m_X.

18.273. Выразить центральный момент n-го порядка через начальные моменты n-го и меньших порядков.

18.274. Выразить начальный момент n-го порядка через центральные моменты n-го и меньших порядков и математическое ожидание.

18.275*. Случайная величина X принимает только целые неотрицательные значения. Доказать, что

$$m_X = \sum_{k=1}^{\infty} \mathrm{P}\{X \geqslant k\}.$$

18.276*. Случайная величина X непрерывного типа неотрицательна, имеет конечное математическое ожидание и ее закон распределения задан функцией распределения $F_X(x)$. Показать, что математическое ожидание такой случайной величины может быть записано в виде $m_X = \int\limits_0^{+\infty} [1 - F_X(x)]\,dx$.

18.277. Пусть X — С. В. Н. Т. с унимодальным законом распределения. Обозначим $\tau_d^2 = \mathrm{M}[(X - d_X)^2]$ (центральный момент второго порядка относительно моды), $\tau_h^2 = \mathrm{M}[(X - h_X)^2]$ (центральный момент второго порядка относительно медианы). Пусть, кроме того, $d_X \neq h_X$. Показать, что для равенства этих моментов необходимо и достаточно, чтобы $m_X = \dfrac{d_X + h_X}{2}$.

В задачах 18.278–18.311 изучаются некоторые классические распределения дискретного и непрерывного типов.

18.278. Из урны, содержащей 4 белых и 6 черных шаров, случайным образом и без возвращения извлекается 3 шара. Случайная величина X — число белых шаров в выборке. Описать закон распределения.

Полученное распределение относится к семейству *гипергеометрических.*

18.279 (продолжение). Для случайной величины из предыдущей задачи найти m_X и D_X.

18.280. Для сборки прибора требуется 4 однотипных детали. Всего имеется 10 деталей, из которых только 6 доброкачественные. Наудачу отбирают 5 деталей (одну деталь «про запас»). Найти вероятность того, что можно будет произвести сборку прибора.

18.281. Производится тираж спортлото «6 из 45». Некто купил одну карточку и заполнил ее. Какова вероятность того, что он правильно угадал k цифр ($k = 6$; 5)?

Случайная величина X непрерывного типа называется распределенной *равномерно на отрезке* $[a, b]$ (при этом для краткости говорят: X подчиняется закону $R(a, b)$), если ее плотность распределения вероятностей постоянна на данном отрезке:

$$f_X(x) = \begin{cases} 0, & \text{если} \quad x \notin [a, b], \\ \dfrac{1}{b-a}, & \text{если} \quad x \in [a, b]. \end{cases}$$

Равномерное распределение реализуется в экспериментах, в которых наудачу ставится точка на отрезке $[a, b]$ (X — координата поставленной точки), а также в экспериментах по измерению тех или иных физических величин с округлением (X — ошибка округления).

18.282. Автобусы идут с интервалом 5 минут. Считая, что случайная величина X — время ожидания автобуса на остановке — распределена равномерно на указанном интервале, найти среднее время ожидания и дисперсию времени ожидания.

18.283 (продолжение). В условиях предыдущей задачи найти функцию распределения случайной величины X и вычислить вероятность того, что время ожидания превысит 3 мин.

18.284. Случайная величина X распределена по закону $R(a, b)$. Найти выражения для m_X и σ_X через параметры распределения a и b.

18.285. Азимутальный лимб имеет цену делений один градус. Какова вероятность при считывании азимута угла сделать ошибку в пределах ± 10 мин, если отсчет округляется до ближайшего целого числа градусов?

18.286. Шкала рычажных весов, установленных в лаборатории, имеет цену деления 1 г. При измерении массы химических компонентов смеси отсчет делается с точностью до целого деления с округлением в ближайшую сторону. Какова вероятность, что абсолютная ошибка определения массы: а) не превысит величины среднеквадратичного отклонения возможных ошибок определения массы; б) будет заключена между значениями σ_X и $2\sigma_X$?

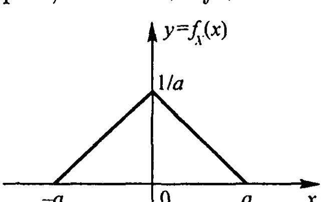


Рис. 7

18.287. Случайная величина X распределена по закону *равнобедренного треугольника в интервале* $(-a, a)$ (*закон Симпсона*), если она непрерывного типа и ее плотность распределения вероятностей имеет вид, изображенный на рис. 7.

Написать выражение для $f_X(x)$, вычислить функцию распределения вероятностей.

18.288 (продолжение). Для случайной величины, распределенной по закону Симпсона, найти математическое ожидание, дисперсию, моду, медиану и коэффициент эксцесса.

Случайная величина X называется распределенной по *показательному (экспоненциальному) закону с параметром* $\lambda > 0$ (при этом для краткости говорят: X подчиняется закону $\mathrm{Ex}(\lambda)$), если она непрерывного типа и ее плотность распределения вероятностей задается формулой

$$f_X(x) = \begin{cases} 0, & \text{если} \quad x \leqslant 0, \\ \lambda e^{-\lambda x}, & \text{если} \quad x > 0. \end{cases}$$

Показательное распределение часто встречается в теории массового обслуживания (например, X — время ожидания при техническом обслуживании или X — длительность телефонных разговоров, ежедневно регистрируемых на телефонной станции) и в теории надежности (например, X — срок службы радиоэлектронной аппаратуры).

18.289. Время безотказной работы радиоаппаратуры является случайной величиной X, распределенной по показательному закону с параметром λ. Вычислить математическое ожидание и дисперсию.

18.290 (продолжение). В условиях предыдущей задачи найти вероятность того, что радиоаппаратура не выйдет из строя в течение времени $t = m_X$. Квантилью какого порядка для данного распределения является значение m_X?

18.291. Время ожидания у бензоколонки автозаправочной станции является случайной величиной X, распределенной по показательному закону со средним временем ожидания, равным t_0. Найти вероятности следующих событий:

$$A = \left\{ \frac{t_0}{2} \leqslant X < \frac{3}{2}\, t_0 \right\}, \quad B = \{X \geqslant 2t_0\}.$$

18.292*. Пусть X — время безотказной работы радиоэлектронной аппаратуры. Примем, что вероятность выхода из строя аппаратуры в течение времени Δx с точностью до величины $o(\Delta x)$ равна $\lambda \Delta x$ $(\lambda > 0)$ независимо от времени x, в течение которого аппаратура уже проработала до рассматриваемого интервала времени Δx. Вычислить функцию распределения случайной величины X.

18.293. Случайная величина X распределена по показательному закону с параметром λ. Вывести рекуррентную формулу, выражающую центральный момент $(k+1)$-го порядка через центральный момент k-го порядка и математическое ожидание, и с ее помощью вычислить коэффициент асимметрии и коэффициент эксцесса показательного распределения.

18.294. Случайная величина X распределена по *закону Коши*, определяемому функцией распределения вероятностей

$$F_X(x) = b + c \operatorname{arctg} \frac{x}{a} \quad \text{при} \quad -\infty < x < +\infty.$$

Выбрать коэффициенты a, b и c таким образом, чтобы данное распределение соответствовало случайной величине непрерывного типа.

18.295 (продолжение). Вычислить плотность вероятности распределения Коши. Существуют ли математическое ожидание и моменты более высокого порядка у данного распределения?

18.296 (продолжение). Найти моду, медиану и квантиль t_p порядка $p = 0{,}75$ распределения Коши.

18.297. Известно, что при стрельбе по плоской мишени в неизменных условиях случайная величина R — расстояние от точки попадания до центра мишени — подчиняется *закону распределения Рэлея* с плотностью распределения вероятностей

$$f_R(x) = \begin{cases} \dfrac{x}{\sigma^2}\, e^{-x^2/(2\sigma^2)} & \text{при} \quad x > 0, \\ 0 & \text{при} \quad x \leqslant 0, \end{cases}$$

где $\sigma > 0$ — параметр, характеризующий распределение. Построить эскиз графика плотности вероятности $f_R(x)$, проверить условие нормировки и вычислить характеристики m_R и D_R.

18.298 (продолжение). Для случайной величины R, распределенной по закону Рэлея, вычислить d_R, h_R и a_R и выяснить взаимное расположение характеристик m_R, d_R и h_R.

18.299. Скорость V молекул идеального газа, находящегося в равновесии при определенной температуре, является случайной величиной, подчиняющейся *закону распределения Максвелла* с плотностью распределения вероятностей

$$f_V(x) = \begin{cases} 0 & \text{при} \quad x \leqslant 0, \\ \sqrt{\dfrac{2}{\pi}}\, \beta^{3/2} x^2 e^{-\beta x^2/2} & \text{при} \quad x > 0, \end{cases}$$

где параметр распределения $\beta > 0$ определяется температурой и массой молекул. Выразить среднее значение и наиболее вероятное значение скорости молекул, а также дисперсию распределения через физический параметр β.

18.300. Случайная величина X подчиняется *закону арксинуса* с плотностью распределения вероятностей

$$f_X(x) = \begin{cases} 0, & \text{если} \quad |x| \geqslant a, \\ \dfrac{1}{\pi\sqrt{a^2 - x^2}}, & \text{если} \quad |x| < a. \end{cases}$$

Найти функцию распределения и вычислить m_X, D_X.

18.301 (продолжение). Для случайной величины, распределенной по закону арксинуса, вычислить d_X, h_X и $\varkappa_{0,75}$.

18.302. Случайная величина X непрерывного типа распределена по *закону Лапласа с параметрами* $m \in \mathbb{R}$ *и* $\sigma > 0$, если ее плотность распределения вероятностей задается формулой

$$f_X(x) = \frac{1}{\sigma\sqrt{2}} \exp\left\{-\frac{|x-m|\sqrt{2}}{\sigma}\right\}, \quad -\infty < x < +\infty.$$

Выразить характеристики m_X и σ_X через параметры распределения.

18.303. Случайная величина X распределена по закону Лапласа с параметрами $m = 0$, $\sigma > 0$. Построить функцию распределения и вычислить вероятности $p_k = \mathrm{P}\{|X| < k\sigma\}$ для $k = 1, 2, 3$.

18.304* (продолжение). Для случайной величины из предыдущей задачи вычислить характеристики a_X и e_X.

18.305. Случайная величина X подчиняется *закону распределения Парето* с параметрами $a > 0$ и $x_0 > 0$, если она С. В. Н. Т. и ее функция распределения вероятностей имеет вид

$$F_X(x) = \begin{cases} 0, & \text{если} \quad x \leqslant x_0, \\ 1 - \left(\dfrac{x_0}{x}\right)^a, & \text{если} \quad x > x_0. \end{cases}$$

Выяснить, при каких значениях параметра a для данного распределения существуют m_X и D_X и вычислить их.

18.306 (продолжение). Вычислить для распределения Парето характеристики d_X и h_X, а также квантиль t_p порядка $p = 0{,}75$.

18.307. В некоторых капиталистических странах действует закон о налогообложении, распространяемый на тех частных предпринимателей, годовой доход которых превосходит некоторый установленный законом уровень x_0. Считая, что годовой доход наудачу выбранного лица, облагаемого налогом, является случайной величиной X, распределенной по закону Парето с параметрами $a = 4$, $x_0 = 1000$, найти вероятности следующих событий:

$$A = \{h_X \leqslant X < m_X\}, \quad B = \{|X - m_X| < \sigma_X\}.$$

Критической точкой какого порядка для данного распределения является математическое ожидание m_X?

Случайная величина X имеет *гамма-распределение с параметрами* $a > 0$ и $b > 0$ (для краткости говорят: X подчиняется закону $\Gamma(a, b)$), если она непрерывного типа и ее плотность распределения вероятностей имеет следующий вид:

$$f_X(x) = \begin{cases} 0 & \text{при} \quad x \leqslant 0, \\ \dfrac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} & \text{при} \quad x > 0, \end{cases}$$

где

$$\Gamma(a) = \int\limits_0^{+\infty} t^{a-1} e^{-t}\, dt \tag{5}$$

— *гамма-функция Эйлера.*

Рассмотренное ранее *показательное распределение с параметром* λ является частным случаем гамма-распределения с параметрами $a = 1$, $b = \lambda > 0$.

Другой частный случай гамма-распределения с параметрами $a = n/2$ (n — натуральное число), $b = 1/2$ называется *распределением хи-квадрат с n степенями свободы* (пишут $\chi^2(n)$). Распределение $\chi^2(n)$ играет большую роль в математической статистике (см. гл. 19, § 6).

Если X подчиняется закону $\chi^2(n)$, то ее плотность распределения вероятностей записывается в виде

$$f_X(x) = \begin{cases} 0, & \text{если} \quad x \leqslant 0, \\ \dfrac{1}{2^{n/2}\,\Gamma\left(\dfrac{n}{2}\right)}\, x^{n/2-1} e^{-x/2}, & \text{если} \quad x > 0. \end{cases} \tag{6}$$

18.308*. Случайная величина X распределена по закону $\Gamma(a, b)$. Найти m_X, D_X, a_X, e_X.

18.309*. Случайная величина X распределена по закону $\chi^2(4)$. Вычислить m_X, D_X и h_X. Критической точкой какого порядка является значение m_X?

Случайная величина X непрерывного типа подчиняется закону *распределения Вейбулла с параметрами* $n \in \mathbb{N}$, $a \in \mathbb{R}$, $b > 0$, если ее плотность распределения вероятностей записывается в виде

$$f_X(x) = \begin{cases} 0, & \text{если} \quad x \leqslant a, \\ \dfrac{n}{b}\left(\dfrac{x-a}{b}\right)^{n-1} \exp\left\{\left(\dfrac{x-a}{b}\right)^{n}\right\}, & \text{если} \quad x > a. \end{cases}$$

Распределение Вейбулла в ряде случаев характеризует срок службы радиоэлектронной аппаратуры и, кроме того, применяется для аппроксимации различных несимметричных распределений в математической статистике.

18.310*. Случайная величина X подчиняется распределению Вейбулла с параметрами n, a, $b > 0$. Вычислить математическое ожидание и моду распределения.

Случайная величина X непрерывного типа имеет *бета-распределение с параметрами $a > 0$ и $b < 0$*, если ее плотность распределения вероятностей записывается в виде

$$f_X(x) = \begin{cases} \dfrac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)}\, x^{a-1}\,(1-x)^{b-1} & \text{при } 0 < x < 1, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Частным случаем бета-распределения при $a = b = 1$ является *равномерное распределение на отрезке* $[0, 1]$.

18.311*. Случайная величина X подчиняется бета-распределению с параметрами $a > 0$ и $b > 0$. Вычислить математическое ожидание и дисперсию.

2. Распределения, связанные с повторными независимыми испытаниями. Ряд классических распределений связан с экспериментом, в котором проводятся последовательные независимые испытания и наблюдается результат совместного осуществления тех или иных исходов каждого испытания.

Последовательные испытания называются *независимыми*, если вероятность осуществления любого исхода в n-м по счету испытании не зависит от реализации исходов предыдущих испытаний.

Простейшим классом повторных независимых испытаний является *последовательность независимых испытаний с двумя исходами* («*успех*» *и* «*неуспех*») *и с неизменными вероятностями* «*успеха*» (p) *и* «*неуспеха*» ($1-p = q$) *в каждом испытании* (*схема испытаний Бернулли*). Наглядным примером таких испытаний является последовательный выбор с возвращением шаров из урны, содержащей m_1 белых и m_2 черных шаров. Если X — число появлений белых шаров в выборке из $n \leqslant m_1 + m_2$ шаров, то

$$\mathrm{P}\{X = k\} = P_{n,k}(p) = C_n^k p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n, \tag{7}$$

где $p = \dfrac{m_1}{m_1 + m_2}$ — вероятность появления белого шара при одном извлечении. Числа $P_{n,k}(p)$ интерпретируются как вероятности получить ровно k успехов в n независимых испытаниях с двумя исходами. Формула (7) называется *формулой Бернулли*, а соответствующее распределение случайной величины X — *биномиальным распределением* (или *распределением Бернулли*). Для краткости говорят, что X распределено по закону $B(n, p)$.

Основные характеристики биномиального распределения:

$$m_X = np, \quad \sigma_X^2 = npq, \quad a_X = \frac{q-p}{\sqrt{npq}}, \quad e_X = \frac{1-6pq}{npq}.$$

Пример 3. Вероятности рождения мальчика и девочки в первом приближении можно считать равными 0,5. Какова вероятность того, что среди $2n$ наудачу отобранных новорожденных будет хотя бы один мальчик (событие A_1); число мальчиков и девочек одинаково (событие A_2); мальчиков будет больше, чем девочек (событие A_3)? Получить числовые значения искомых вероятностей для $2n = 10$; $2n = 100$.

◁ Пусть X — число мальчиков среди $2n$ новорожденных. Случайная величина X подчиняется распределению $B(2n, 1/2)$, т.е. согласно формуле (7)

$$\mathrm{P}\{X = k\} = C_{2n}^k \left(\frac{1}{2}\right)^{2n}, \quad k = 0, 1, \ldots, 2n.$$

Вероятность события A_1 проще всего найти, перейдя к противоположному событию:

$$\mathrm{P}(A_1) = 1 - \mathrm{P}(\overline{A}_1) = 1 - \mathrm{P}\{X = 0\} = 1 - \left(\frac{1}{2}\right)^{2n}.$$

Вероятность события A_2 записывается непосредственно:

$$\mathrm{P}(A_2) = \mathrm{P}\{X = n\} = C_{2n}^n \left(\frac{1}{2}\right)^{2n}.$$

Для подсчета вероятности события A_3 заметим, что распределение Бернулли $B(2n, 1/2)$ симметрично относительно значения $x = n$. Действительно:

$$\mathrm{P}\{X = n - k\} = P_{2n,\, n-k}\left(\frac{1}{2}\right) = C_k^{n-k}\left(\frac{1}{2}\right)^{2n} = \\ = P_{2n,\, n+k}\left(\frac{1}{2}\right) = \mathrm{P}\{X = n + k\}$$

для всех $k = 1, 2, \ldots, n-1$. Кроме того, нетрудно проверить, что это значение является наиболее вероятным, т.е. $d_X = n$ (мода распределения). В силу симметрии распределения выполняется равенство

$$\mathrm{P}\{X > n\} = \mathrm{P}\{X < n\} = \frac{1}{2}(1 - \mathrm{P}\{X = n\}).$$

Таким образом, $\mathrm{P}(A_3) = \dfrac{1}{2}\left(1 - C_{2n}^n \left(\dfrac{1}{2}\right)^{2n}\right)$. Найдем числовые значения полученных вероятностей. При $2n = 10$ результат получается непосредственно:

$$P(A_1) = 1 - \left(\frac{1}{2}\right)^{10} \approx 0{,}9990,$$

$$P(A_2) = C_{10}^5 \cdot \left(\frac{1}{2}\right)^{10} \approx 0{,}2461,$$

$$P(A_3) = \frac{1}{2}(1 - P(A_2)) \approx 0{,}3770.$$

При $2n = 100$ непосредственное вычисление величин $\left(\frac{1}{2}\right)^{100}$ и C_{100}^{50} становится затруднительным.

Для вычисления $\left(\frac{1}{2}\right)^{100}$ применяем логарифмирование по основанию 10: $\left(\frac{1}{2}\right)^{100} = 10^x$, откуда $x = 100 \lg \frac{1}{2} \approx -30$. Таким образом, $P(A_1) = 1 - 10^{-30} \approx 1$.

Для вычисления $C_{100}^{50} = \dfrac{100!}{50!\,50!}$ применяем формулу Стирлинга для $n!$ при больших значениях n:

$$n! \approx \sqrt{2\pi n}\, n^n e^{-n}.$$

Тогда

$$P(A_2) = C_{100}^{50} \left(\frac{1}{2}\right)^{100} \approx \frac{1}{5\sqrt{2\pi}} \approx 0{,}0798.$$

Далее находим

$$P(A_3) = \frac{1}{2}(1 - P(A_2)) \approx 0{,}4601. \ \triangleright$$

18.312. Для стрелка, выполняющего упражнение в тире, вероятность попасть в «яблочко» при одном выстреле не зависит от результатов предшествующих выстрелов и равна $p = 1/4$. Спортсмен сделал 5 выстрелов. Найти вероятности событий: $A = \{$ровно одно попадание$\}$, $B = \{$ровно два попадания$\}$.

18.313 (продолжение). В условиях предыдущей задачи найти вероятности событий: $C = \{$хотя бы одно попадание$\}$, $D = \{$не менее трех попаданий$\}$.

18.314. Десять осветительных лампочек для елки включены в цепь последовательно. Вероятность для любой лампочки перегореть при повышении напряжения в сети равна 0,1. Определить вероятность разрыва цепи при повышении напряжения в сети.

18.315. Пара одинаковых игральных костей бросается 7 раз. Какова вероятность следующих событий: $A = \{$сумма очков, равная 7, выпадет дважды$\}$, $B = \{$сумма очков, равная 7, выпадет по крайней мере 1 раз$\}$.

18.316 (продолжение). В условиях предыдущей задачи найти вероятности событий: $C = \{$каждый раз выпадет сумма очков, бо̀льшая 7$\}$, $D = \{$ни разу не выпадет сумма очков, равная 12$\}$.

18.317. Проводится одно испытание с вероятностью успеха, равной p. Пусть X — случайное число успехов в данном испытании (индикаторная случайная величина согласно определению из задачи 18.261). Описать закон распределения X, вычислить m_X и D_X.

18.318. Вероятность, что покупателю потребуется обувь 40-го размера, равна 0,4. В обувной отдел вошли трое покупателей. Пусть X — число тех покупателей, которым потребовалась обувь 40-го размера. Вычислить $\mathrm{P}\{X \geqslant 2\}$.

18.319 (продолжение). Найти функцию распределения случайной величины X из предыдущей задачи.

18.320. В ячейку памяти ЭВМ записывается 8-разрядное двоичное число. Значения 0 и 1 в каждом разряде появляются с равной вероятностью. Случайная величина X — число единиц в записи двоичного числа. Найти вероятности событий: $A = \{X = 4\}$, $B = \{X > 4\}$.

18.321. Имеется 200 семей, в каждой из которых 4 ребенка. Случайные величины: X — число семей из 200, имеющих одного мальчика и трех девочек, Y — число семей, имеющих двух мальчиков и двух девочек. Считая, что вероятности рождения мальчика и девочки одинаковы, найти $\mathrm{M}[X]$ и $\mathrm{M}[Y]$.

18.322. Два равносильных шахматиста договорились сыграть матч из $2n$ результативных партий. Ничьи не учитываются и считается, что каждый из участников может выиграть очередную партию с вероятностью 0,5. Выигравшим матч считается тот, кто победит в большем числе партий. В каком матче больше шансов выиграть любому из участников: в матче из 8 результативных партий или из 12?

18.323. Устройство состоит из 8 независимо работающих элементов. Вероятности отказов каждого из элементов за время T одинаковы и равны $p = 0{,}2$. Найти вероятность отказа прибора, если для этого достаточно, чтобы отказали хотя бы три элемента из восьми.

18.324. Производится обстрел учебной цели из орудия. Вероятность попадания в цель при одном выстреле равна p. Поражение цели может наступить при k попаданиях ($k = 1, 2, \ldots$) с вероятностью, равной $1 - t^k$ $(0 < t < 1)$. Вычислить вероятность поражения цели при n выстрелах.

18.325. На контроль поступила партия деталей из цеха. Известно, что 5 % всех деталей не удовлетворяет стандарту. Сколько нужно испытать деталей, чтобы с вероятностью не менее 0,95 обнаружить хотя бы одну нестандартную деталь?

18.326. Путем длительных наблюдений установлено, что в данной местности в сентябре в среднем бывает 12 дождливых дней. Что вероятнее: из восьми наудачу взятых дней сентября будет два дождливых или три дождливых дня?

18.327. Доказать *рекуррентную формулу для биномиальных вероятностей*:

$$P_{n,m+1}(p) = \frac{p}{q}\,\frac{n-m}{m+1}\,P_{n,m}(p)$$

и с ее помощью установить, что наиболее вероятное число успехов ($d_X = M$) в серии n независимых испытаний удовлетворяет неравенству

$$np - q \leqslant M \leqslant np + p.$$

18.328. Вероятность отказа каждого прибора при испытании не зависит от отказов остальных приборов и равна 0,2. Испытано 9 приборов. Случайная величина X — число отказавших за время испытаний приборов. Найти наиболее вероятное число отказавших приборов.

18.329 (продолжение). В условиях предыдущей задачи найти вероятность события $A = \{X \geqslant m_X\}$.

18.329 (1). В последовательности n независимых испытаний с вероятностью успеха p в каждом испытании произошло ровно 2 успеха. Какова вероятность, что успехи произошли в соседних испытаниях?

18.330. Испытания по схеме Бернулли с вероятностью успеха p в одном испытании повторяются до тех пор, пока не появится успех, после чего прекращаются. Обозначим X число проведенных испытаний до первого успеха включительно. Описать закон распределения случайной величины X и найти $\mathrm{P}\,\{X \leqslant 3\}$ (полученное распределение называется *геометрическим с параметром* $p > 0$).

18.331 (продолжение). В условиях предыдущей задачи найти m_X и D_X.

18.332. Вероятность появления брака на автоматической линии равна 0,001. Линия работает без переналадки до появления первого бракованного изделия. Сколько изделий в среднем производит данная автоматическая линия между двумя переналадками?

Какова вероятность того, что число произведенных изделий окажется больше $3m_X$?

18.333*. Вероятность попадания стрелка в мишень в неизменных условиях постоянна и равна p. Стрелку выдаются патроны до тех пор, пока он не промахнется. Обозначим X число выданных стрелку патронов в данном эксперименте. Найти m_X, D_X, d_X.

18.334. Проводятся последовательные испытания по схеме Бернулли. Вероятность успеха в одном испытании равна p. Вычислить вероятность события $A = \{$все k успехов в n испытаниях появятся подряд$\}$.

18.335 (продолжение). В условиях предыдущей задачи найти вероятности следующих событий: $B = \{k$-й по счету успех наступит в m-м по счету испытании $(m \geqslant k)\}$, $C = \{k$-й по счету успех наступит прежде, чем наберется m неуспехов$\}$.

18.335 (1). В автобусе едут n пассажиров. Каждый из пассажиров может выйти на следующей остановке с вероятностью p. Кроме того, в автобус могут войти k пассажиров с вероятностью p_k $(k = 0, 1, 2, \ldots, m;\ m \leqslant n)$. Принимая модель независимости поведения каждого из пассажиров от остальных, найти вероятность события $A = \{$после остановки в автобусе снова будут ехать n пассажиров$\}$.

18.336. Испытания по схеме Бернулли с вероятностью успеха в одном испытании p повторяются до получения ровно k успехов. Описать закон распределения и найти среднее значение числа проведенных испытаний в данном эксперименте. (Указанное распределение называется *«отрицательным биномиальным» с параметрами* $k \in \mathbb{N}$ *и* $p > 0$)

18.337. Вероятность получения отметки цели на экране обзорного радиолокатора при одном обороте антенны равна $1/6$. Цель считается обнаруженной, если получены 3 отметки. Какова вероятность, что цель будет обнаружена не более чем за 5 оборотов антенны?

Если вероятность осуществления события A от испытания к испытанию меняется, то формула Бернулли становится неприменимой. Пусть $p_k = P_k(A)$ — вероятность «успеха» в k-м испытании в последовательности независимых испытаний ($q_k = 1 - p_k$ — вероятность «неуспеха» в k-м испытании). Тогда вероятность $P_{n,m}$ осуществления ровно m успехов в n независимых испытаниях равна коэффициенту при x^m в разложении по степеням x *производящей функции*

$$G_n(x) = \prod_{k=1}^{n}(q_k + p_k x) = \sum_{m=0}^{n} P_{n,m}x^m.$$

Искомые коэффициенты $P_{n,m}$ вычисляются дифференцированием по x

производящей функции $G_n(x)$:

$$P_{n,m} = \frac{1}{m!} \frac{d^m G_n(x)}{dx^m}\bigg|_{x=0}. \tag{8}$$

Пример 4. Обозначим X число успехов в последовательности n независимых испытаний с вероятностью успеха в k-м испытании, равной p_k. Вычислить среднее число успехов и дисперсию величины X. Получить аналогичные характеристики биномиального распределения в частном случае $p_k = p$ для всех $k = 0, 1, \ldots, n$.

◁ Пользуясь определением производящей функции, можем написать

$$m_X = \sum_{m=0}^{n} m P_{n,m} = \frac{dG_n(x)}{dx}\bigg|_{x=1} = \sum_{k=1}^{n} p_k.$$

Второй начальный момент находим аналогично:

$$\alpha_2 = \sum_{m=0}^{n} m^2 P_{n,m} = \frac{d}{dx}\left(xG_n'(x)\right)\bigg|_{x=1} = G_n'(1) + G_n''(1),$$

$$G_n''(1) = p_1 \sum_{k\neq 1}^{n} p_k + p_2 \sum_{k\neq 2}^{n} p_k + \cdots + p_n \sum_{k\neq n}^{n} p_k = \left(\sum_{k=1}^{n} p_k\right)^2 - \sum_{k=1}^{n} p_k^2.$$

Отсюда

$$\alpha_2 = \sum_{k=1}^{n} p_k + \left(\sum_{k=1}^{n} p_k\right)^2 - \sum_{k=1}^{n} p_k^2, \quad D_X = \alpha_2 - m_X^2 = \sum_{k=1}^{n} p_k q_k.$$

В частном случае $p_1 = p_2 = \cdots = p_n = p$ из этих формул следует, что $m_X = np$, $D_X = npq$. ▷

18.338. Производится стрельба из орудия по удаляющейся цели. При первом выстреле вероятность попадания равна 0,8, при каждом следующем выстреле вероятность попадания уменьшается в 2 раза. Случайная величина X — число попаданий в цель при двух выстрелах. Описать закон распределения.

18.339 (продолжение). В условиях предыдущей задачи найти m_X и D_X.

18.340. Пусть случайная величина X — число попаданий при 4 выстрелах в условиях задачи 18.338. Найти среднее число попаданий и дисперсию числа попаданий.

18.341 (продолжение). В условиях предыдущей задачи вычислить вероятности событий: $A = \{$ровно одно попадание$\}$, $B = \{$по крайней мере одно попадание$\}$.

18.342. Прибор состоит из пяти элементов. Отказ k-го элемента за время T независимо от остальных элементов происходит с вероятностью $p_k = 0{,}2 + (k-1)\cdot 0{,}1$. Определить: а) математическое ожидание и дисперсию числа отказавших за время T элементов; б) вероятность того, что за время T откажет хотя бы один из элементов прибора.

18.343. Вероятность перегорания первой, второй и третьей лампы соответственно равна 0,1, 0,2 и 0,3. Если перегорает одна лампа, то прибор выходит из строя с вероятностью 0,5, а если две или три — то прибор заведомо выйдет из строя. Найти вероятность выхода прибора из строя.

18.344*. Проводятся последовательные независимые испытания с двумя исходами, причем вероятность успеха в k-м по счету испытании равна p_k ($q_k = 1 - p_k$ — вероятность неуспеха). Доказать *рекуррентную формулу* для вероятности осуществления k успехов в m испытаниях:

$$\begin{aligned} &P_{m,k} = P_{m-1,k}\, q_m + P_{m-1,k-1}\, p_m, \\ &m = 2, 3, \ldots; \quad k = 1, 2, \ldots, m-1, \\ &P_{m,m} = p_1 p_2 \ldots p_m, \quad m = 1, 2, \ldots, \\ &P_{m,0} = q_1 q_2 \ldots q_m, \quad m = 1, 2, \ldots \end{aligned}$$

18.345*. Последовательно посылается 4 радиосигнала. Вероятность приема каждого из них не зависит от того, приняты или нет остальные сигналы, и равна соответственно 0,1; 0,2; 0,3 и 0,4. Вычислить вероятность того, что будет принято ровно два радиосигнала.

Пусть каждое из n независимых испытаний имеет N взаимоисключающих исходов $\omega_1, \omega_2, \ldots, \omega_N$ соответственно с вероятностями p_1, $p_2, \ldots, p_N$ $\left(\sum\limits_{k=1}^{N} p_k = 1\right)$, не меняющимися от испытания к испытанию. Обозначим $X_{n,k}$ ($k = 1, 2, \ldots, N$) число появлений исхода ω_k в n испытаниях. Тогда вероятность совместного осуществления составного исхода всех n испытаний, состоящего в том, что исход ω_1 появится m_1 раз, исход ω_2 — m_2 раз, ..., исход ω_N — m_N раз ($m_1 + m_2 + \cdots + m_N = n$), выражается формулой

$$\mathrm{P}\{X_{n,1} = m_1,\ X_{n,2} = m_2,\ \ldots,\ X_{n,N} = m_N\} = \\ = P_{n;\, m_1, m_2, \ldots, m_N} = \frac{n!}{m_1!\, m_2! \ldots m_N!}\, p_1^{m_1}\, p_2^{m_2} \ldots p_N^{m_N}. \quad (9)$$

Описанная схема последовательности испытаний с N исходами называется *полиномиальной схемой*, а формула (9) определяет вероятности *полиномиального распределения.* Распределение Бернулли (7) является частным случаем полиномиального распределения при $N = 2$, $p_2 = 1 - p_1 = q_1$.

Пример 5. В урне содержится 8 белых, 5 красных и 2 черных шара. Производится 5 извлечений с возвращением по одному шару. Рассматриваются события $A = \{$появился следующий состав шаров: 3 белых и по одному остальных цветов$\}$, $B = \{$появилось ровно 3 белых шара$\}$, $C = \{$появилось 3 белых шара и по одному остальных цветов, причем белые шары появились подряд$\}$. Определить их вероятности.

$\triangleleft$ Событие A соответствует полиномиальной схеме при $n = 5$, $N = 3$, $p_1 = 8/15$, $p_2 = 5/15$, $p_3 = 2/15$, поэтому

$$\mathrm{P}(A) = P_{5;\,3,\,1,\,1} = \frac{5!}{3!\,1!\,1!} \cdot \left(\frac{8}{15}\right)^3 \cdot \frac{5}{15} \cdot \frac{2}{15} = \frac{8}{9} \cdot \left(\frac{8}{15}\right)^3 \approx 0{,}1348.$$

Событие B соответствует биномиальной схеме со значениями $N = 2$, $p_1 = 8/15$, $q_1 = 7/15$, поэтому

$$\mathrm{P}(B) = P_{5,\,3}\left(\frac{8}{15}\right) = C_5^3 \cdot \left(\frac{8}{15}\right)^3 \cdot \left(\frac{7}{15}\right)^2 = 10 \cdot \left(\frac{8}{15}\right)^3 \cdot \left(\frac{7}{15}\right)^2 \approx 0{,}3304.$$

Событие C соответствует комбинированной схеме, в которой в каких-либо трех последовательных испытаниях белый шар выпал трижды, а в остальных двух испытаниях по одному разу выпали черный и красный шары, поэтому

$$\mathrm{P}(C) = 3P_{3,\,3}\left(\frac{8}{15}\right) P_{2;\,0,\,1,\,1} =$$

$$= 3 \cdot \left(\frac{8}{15}\right)^3 \cdot \frac{2!}{0!\,1!\,1!} \cdot \left(\frac{8}{15}\right)^0 \cdot \frac{5}{15} \cdot \frac{2}{15} = \frac{1}{2} \cdot \left(\frac{8}{15}\right)^4 \approx 0{,}0404. \ \triangleright$$

18.346. Произведено три независимых выстрела по мишени в неизменных условиях. Вероятность при одном выстреле попасть в «десятку» равна $p_{10} = 0{,}3$, вероятность попасть в «девятку» равна $p_9 = 0{,}4$, вероятность не попасть ни в девятку, ни в десятку равна $p_0 = 1 - p_{10} - p_9 = 0{,}3$. Найти вероятности следующих событий: $A = \{$одно попадание в «десятку» и одно в «девятку»$\}$, $B = \{$ровно два попадания в «десятку»$\}$.

18.347 (продолжение). В условиях предыдущей задачи найти вероятность события $C = \{$будет набрано не менее 29 очков$\}$.

18.348. Каждый из десяти аспирантов группы случайным образом и независимо от остальных выбирает один из четырех дней наступающей недели (понедельник, вторник, среду или четверг) для работы в библиотеке в отделе текущей периодики. Найти вероятность следующих событий: $A = \{$в понедельник в библиотеку явится один аспирант, во вторник — два, в среду — три, в четверг — четыре аспиранта}, $B =$ {в понедельник появятся 3 аспиранта, а во вторник 7}.

18.349 (продолжение). В условиях предыдущей задачи найти вероятности событий: $C =$ {пятеро из аспирантов появятся в библиотеке в первые два дня недели и пятеро — в следующие два дня}, $D =$ {в понедельник и вторник не появится ни один аспирант}.

18.350. Два равносильных шахматиста играют матч из 12 партий. В каждой партии возможно три исхода: $\omega_1 =$ {выиграл первый игрок (проиграл второй)}, $\omega_2 =$ {ничья}, $\omega_3 =$ {выиграл второй (проиграл первый)}. Пусть

$$\mathrm{P}(\omega_1) = \mathrm{P}(\omega_3) = 0{,}2, \quad \mathrm{P}(\omega_2) = 1 - \mathrm{P}(\omega_1) - \mathrm{P}(\omega_3) = 0{,}6.$$

Найти вероятности следующих событий: $A =$ {первый игрок выиграл 3 партии, проиграл 3 партии и остальные свел вничью}, $B =$ {один из игроков выиграл 4 партии и проиграл 3 партии}.

18.351 (продолжение). В условиях предыдущей задачи найти вероятность события $C =$ {сыграно 2 результативные партии}.

3. Распределение Пуассона. Случайная величина X называется распределенной по *закону Пуассона с параметром* $\lambda > 0$, если ее возможные значения равны 0, 1, 2, ..., а соответствующие вероятности определяются формулой

$$\mathrm{P}\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}. \tag{10}$$

Характерной особенностью распределения Пуассона является совпадение математического ожидания и дисперсии, причем

$$m_x = D_x = \lambda.$$

Распределение Пуассона может быть получено из биномиального распределения путем предельного перехода при $n \to \infty$, $p \to 0$ при условии $np = \lambda = \mathrm{const}$ и в этом случае интерпретируется как *закон «редких» явлений.* Если n достаточно велико, а p мало, то формулу Пуассона (10) часто используют в качестве приближения вместо точных биномиальных формул для вероятностей k успехов в n испытаниях.

В таблицах П3 и П4 приведены вероятности распределения Пуассона и суммарные вероятности для распределения Пуассона соответственно.

Пример 6. На факультете насчитывается 500 студентов. Какова вероятность того, что 1 сентября является днем рождения одновременно для k студентов данного факультета? Вычислить указанную вероятность для значений $k = 0, 1, 2, 3$.

$\triangleleft$ Так как $n = 500 \gg 1$ и $p = \mathrm{P}$ {родился 1 сентября любой из студентов факультета} $= \dfrac{1}{365} \ll 1$, то можно считать, что случайное число студентов X, родившихся 1 сентября, подчиняется закону распределения Пуассона с параметром $\lambda = 1{,}36986$. Поэтому по формуле (10) $p_0 = \mathrm{P}\{X = 0\} = e^{-\lambda} \approx 0{,}2541$. Далее находим рекуррентно:

$$p_1 = \mathrm{P}\{X = 1\} = \frac{\lambda}{1!}e^{-\lambda} = \lambda p_0 \approx 0{,}3481,$$
$$p_2 = \mathrm{P}\{X = 2\} = \frac{\lambda^2}{2!}e^{-\lambda} = \frac{\lambda}{2}p_1 \approx 0{,}2385,$$
$$p_3 = \mathrm{P}\{X = 3\} = \frac{\lambda}{3}p_2 \approx 0{,}1089.$$

Значения искомых вероятностей, соответствующих биномиальному распределению $B(500,\ 1/365)$ и вычисленных с четырьмя верными знаками после запятой по рекуррентной формуле задачи 18.327, таковы: $p_0 = 0{,}2537$, $p_1 = 0{,}3485$, $p_2 = 0{,}2389$, $p_3 = 0{,}1089$. ▷

18.352 [2]). Аппаратура состоит из 1000 элементов, каждый из которых независимо от остальных выходит из строя за время T с вероятностью $p = 5 \cdot 10^{-4}$. Найти вероятности следующих событий: $A = \{$за время T откажет ровно 3 элемента$\}$, $B = \{$за время T откажет хотя бы один элемент$\}$, $C = \{$за время T откажет не более 3 элементов$\}$.

18.353. Среднее число вызовов, поступающих на АТС в минуту, ровно 120. Найти вероятности следующих событий: $A = \{$за две секунды на АТС не поступит ни одного вызова$\}$, $B = \{$за две секунды на АТС поступит менее двух вызовов$\}$.

18.354 (продолжение). В условиях предыдущей задачи найти вероятности событий: $C = \{$за одну секунду на АТС поступит ровно три вызова$\}$, $D = \{$за три секунды на АТС поступит не менее трех вызовов$\}$.

18.355. Случайная величина X — число электронов, вылетающих с нагретого катода электронной лампы в течение времени t, λ — среднее число электронов, испускаемых в единицу времени. Определить вероятности следующих событий: $A = \{$за время t_1 число испускаемых электронов будет меньше m, $m \in \mathbb{N}\}$, $B = \{$за время t_2 вылетит четное число электронов$\}$.

18.356. Корректура в 500 страниц содержит 1300 опечаток. Найти наиболее вероятное число опечаток на одной странице текста и вероятность этого числа.

18.357. Радиостанция ведет автоматическую передачу цифрового текста в течение 10 мкс. Работа ее происходит при наличии

---

[2]) Считаем, что в задачах 18.352–18.357 соответствующая случайная величина имеет распределение Пуассона.

хаотической импульсной помехи, среднее число импульсов которой в одну секунду составляет 10^4. Для срыва передачи достаточно попадание двух импульсов помехи в период работы станции. Вычислить вероятность срыва передачи.

18.358. Число элементарных частиц, регистрируемых прибором, случайно и образует пуассоновскую случайную величину со средним значением n частиц. Каждая регистрируемая частица может нести заряд с вероятностью p и быть нейтральной с вероятностью $1-p$. Определить закон распределения числа заряженных частиц, регистрируемых прибором, и найти среднее значение и дисперсию полученного распределения.

18.359. При испытании легированной стали на содержание углерода вероятность того, что в случайно взятой пробе процент углерода превысит допустимый уровень, равна $p = 0{,}01$. Считая применимым закон редких явлений, вычислить, сколько в среднем необходимо испытать образцов, чтобы с вероятностью $p = 0{,}95$ указанный эффект наблюдался по крайней мере 1 раз.

18.360 (продолжение). Ответить на вопрос предыдущей задачи, если требуется, чтобы указанный эффект наблюдался не менее двух раз.

4. Нормальный закон распределения. Случайная величина называется распределенной по *нормальному* (*гауссовскому*) *закону с параметрами* $m \in \mathbb{R}$ *и* $\sigma > 0$, если плотность распределения вероятностей имеет вид

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}, \quad -\infty < x < +\infty. \tag{11}$$

Параметры m и σ совпадают с основными характеристиками распределения:

$$m = m_X, \quad \sigma = \sigma_X = \sqrt{D_X}.$$

Для краткости говорят, что случайная величина X распределена по закону $N(m, \sigma)$, если ее плотность вероятностей записывается в виде (11).

Если X распределена по закону $N(0,1)$, то она называется *стандартизованной нормальной величиной.* Функция распределения стандартизованной гауссовской величины

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left\{-\frac{t^2}{2}\right\} dt$$

называется *функцией нормального распределения.* С ее помощью можно вычислять интервальные вероятности для нормального распределения $N(m, \sigma)$:

$$\mathrm{P}\{x_1 \leqslant X < x_2\} = \Phi\left(\frac{x_2 - m}{\sigma}\right) - \Phi\left(\frac{x_1 - m}{\sigma}\right).$$

Значения функции $\Phi(x)$ приведены в приложении (таблица П1). При решении задач на нормальное распределение часто требуется использовать табличные значения функции нормального распределения. Поскольку для этой функции справедливо соотношение

$$\Phi(-x) = 1 - \Phi(x),$$

достаточно иметь табличные значения функции $\Phi(x)$ только для положительных значений аргумента. Для вероятности попадания на симметричный относительно математического ожидания интервал справедлива формула

$$\mathrm{P}\{|X - m_x| < \varepsilon\} = 2\Phi\left(\frac{\varepsilon}{\sigma}\right) - 1.$$

Центральные моменты нормального распределения удовлетворяют рекуррентному соотношению

$$\mu_{n+2} = (n+1)\,\sigma^2 \mu_n, \quad n = 0, 1, 2, \ldots \tag{12}$$

Отсюда следует, что все центральные моменты нечетного порядка равны нулю (так как $\mu_1 = 0$).

Пример 7. Производится измерение без систематических ошибок диаметра вала. Случайные ошибки измерения X подчиняются нормальному распределению со *стандартным отклонением* 10 мм. Найти вероятность того, что измерение будет произведено с ошибкой, не превосходящей по абсолютной величине 15 мм.

◁ Так как по условию систематические ошибки отсутствуют, то $m_x = 0$ мм. «Стандартное отклонение» — это другое название для среднеквадратического отклонения, часто используемое на практике. Поэтому $\sigma = 10$ мм.

Для искомой вероятности попадания в симметричный интервал используем формулу

$$\mathrm{P}\{|X| < 15\} = 2\Phi\left(\frac{15}{\sigma}\right) - 1 = 2\Phi(1{,}5) - 1.$$

По таблице П1 находим $\Phi(1{,}5) \approx 0{,}9332$. Таким образом,

$$\mathrm{P}\{|X| < 15\} \approx 0{,}8664. \ ▷$$

18.361. Случайная величина X нормально распределена с параметрами $m = 1$, $\sigma = 2$. Выразить ее функцию распределения через функцию $\Phi(x)$.

18.362. Случайная величина X распределена по закону $N(m, \sigma)$. Пользуясь таблицей функции нормального распределения, вычислить вероятность p_k того, что отклонение величины X от ее математического ожидания не превзойдет величины $k\sigma$ (ответ получить для трех значений $k = 1, 2, 3$).

18.363. Измеряемая случайная величина X подчиняется закону распределения $N\,(10{,}5)$. Найти симметричный относительно m_X интервал, в который с вероятностью p попадет измеренное значение. Рассмотреть следующие числовые значения: а) $p = 0{,}9974$; б) $p = 0{,}9544$; в) $p = 0{,}50$.

18.364. Химический завод изготовляет серную кислоту номинальной плотности $1{,}84\,\text{г}/\text{см}^3$. В результате статистических испытаний обнаружено, что практически 99,9 % всех выпускаемых реактивов имеют плотность в интервале $(1{,}82; 1{,}86)$. Найти вероятность того, что кислота удовлетворяет стандарту, если для этого достаточно, чтобы ее плотность не отклонялась от номинала более, чем на $0{,}01\,\text{г}/\text{см}^3$.

18.365. В нормально распределенной совокупности 15 % значений x меньше 12 и 40 % значений x больше 16,2. Найти среднее значение и стандартное отклонение данного распределения.

18.366. Деталь, изготовленная автоматом, считается годной, если отклонение X контролируемого размера от номинала не превышает 10 мм. Точность изготовления деталей характеризуется стандартным отклонением σ. Считая, что для данной технологии $\sigma = 5$ и X нормально распределена, выяснить, сколько процентов годных деталей изготовляет автомат.

18.367 (продолжение). В условиях предыдущей задачи выяснить, какой должна быть точность изготовления, чтобы процент годных деталей повысился до 98?

18.368. Коробки с шоколадом упаковываются автоматически. Их средняя масса равна 1,06 кг. Известно, что 5 % коробок имеют массу, меньшую 1 кг. Каков процент коробок, масса которых превышает 940 г?

18.369*. Деталь изготавливается на станке. Ее размер X представляет собой случайную величину, распределенную по нормальному закону со средним значением 20 см и стандартным отклонением 0,2 см. Какую относительную точность изделия можно гарантировать с вероятностью 0,95?

18.370. Браковка шариков для подшипников производится следующим образом: если шарик проходит через отверстие диаметра d_2, но не проходит через отверстие диаметра $d_1 < d_2$, то шарик считается годным. Если какое-либо из этих условий нарушается, то шарик бракуется. Считается, что диаметр шарика X — случайная величина, распределенная по закону $N\left(\dfrac{d_1+d_2}{2},\ \alpha\,(d_2-d_1)\right)$, где параметр α $(0 < \alpha < 1/2)$ определяет точность изготовления шариков. Определить вероятность того, что шарик будет забракован.

18.371 (продолжение). В условиях предыдущей задачи определить, какую точность изготовления следует установить (т.е. каким следует выбрать параметр α), чтобы брак составлял не более 2% всей продукции?

18.372. Случайная величина X подчиняется закону $N(1, \sigma)$. Известно, что $\mathrm{P}\{X < 2\} = 0{,}99$. Вычислить $\mathrm{M}[X^2]$ и $\mathrm{P}\{X^2 > 2\}$.

18.373. Случайная величина X распределена по закону $N(m, \sigma)$. Вычислить $p_1 = \mathrm{P}\{X \geqslant x_{n2}\}$ и $p_2 = \mathrm{P}\{x_{n1} \leqslant X \leqslant x_{n2}\}$, где x_{n1} и x_{n2} — точки перегиба кривой плотности распределения вероятностей.

18.374** (продолжение). Пусть (a, b) — интервал, не содержащий m_X. В условиях предыдущей задачи определить, при каком σ вероятность $Q(\sigma) = \mathrm{P}\{a \leqslant X < b\}$ будет наибольшей?

18.375 (продолжение). В условиях задачи 18.373 вычислить первый абсолютный центральный момент $\mathrm{M}(|X - m_X|)$.

18.376. Случайная величина X распределена по закону $N(-1, 1)$. Вычислить a_X и e_X.

18.377* (продолжение). Для случайной величины X из предыдущей задачи найти $\mathrm{M}[X^4]$ и $\mathrm{M}[X^6]$.

§ 3. Случайные векторы

1. Законы распределения и числовые характеристики случайных векторов. Пусть для данного эксперимента определены случайные величины $X_1 = X_1(\omega)$, $X_2 = X_2(\omega)$, ..., $X_n = X_n(\omega)$, $\omega \in \Omega$. Каждому элементарному событию ω можно поставить в соответствие *n-мерный случайный вектор* (*n-мерную случайную величину*)

$$X(\omega) = (X_1, X_2, \ldots, X_n),$$

задающий отображение $\Omega \to \mathbb{R}^n$.

Функцией распределения n-мерного случайного вектора или *функцией совместного распределения случайных величин* $X_1, X_2, \ldots, X_n$ называется неслучайная функция n действительных переменных x_1, x_2, ..., x_n (функция точки $(x_1, x_2, \ldots, x_n)$ в n-мерном евклидовом пространстве $\mathbb{R}^n$), определяемая как вероятность совместного выполнения n неравенств

$$F_X(x) = F_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = \\ = \mathrm{P}\{X_1 < x_1, X_2 < x_2, \ldots, X_n < x_n\}.$$

В частном случае, для двумерного случайного вектора (X, Y), имеем по определению

$$F_{X,Y}(x, y) = \mathrm{P}\{X < x, Y < y\}.$$

Функция распределения $F_{X,\,Y}(x,\,y)$ обладает следующими свойствами:

1. $\lim\limits_{x\to-\infty} F_{X,\,Y}(x,\,y) = \lim\limits_{y\to-\infty} F_{X,\,Y}(x,\,y) = 0$.
2. $\lim\limits_{x\to+\infty} F_{X,\,Y}(x,\,y) = F_Y(y)$, $\lim\limits_{y\to+\infty} F_{X,\,Y}(x,\,y) = F_X(x)$.
3. $\lim\limits_{\substack{x\to+\infty\\ y\to+\infty}} F_{X,\,Y}(x,\,y) = 1$.
4. Функция $F_{X,\,Y}(x,\,y)$ — неубывающая функция своих аргументов.
5. Функция $F_{X,\,Y}(x,\,y)$ непрерывна слева по каждому из аргументов.

Свойство 2 обычно называют *условием согласованности*. Оно означает, что функции распределения отдельных компонент двумерного случайного вектора могут быть найдены предельным переходом из функции совместного распределения этих компонент.

Вероятность попадания случайной точки на плоскость $(X,\,Y)$ в прямоугольник со сторонами, параллельными осям координат, может быть вычислена с помощью функции распределения по формуле

$$\mathrm{P}\{x_1 \leqslant X < x_2,\ y_1 \leqslant Y < y_2\} = \\ = F_{X,\,Y}(x_1,\,y_1) + F_{X,\,Y}(x_2,\,y_2) - F_{X,\,Y}(x_1,\,y_2) - F_{X,\,Y}(x_2,\,y_1).$$

Двумерный случайный вектор $(X,\,Y)$ называется *случайным вектором дискретного типа* (сокращенно С. В. Д. Т.), если множество его возможных значений $G\,(x,\,y)$ не более чем счетно.

Перечень возможных значений пар компонент $\{(x_i,\,y_j)\,|\,(x_i,\,y_j)) \in \in G\,(x,\,y)\}$ и соответствующих каждой такой паре вероятностей

$$p_{ij} = \mathrm{P}\{X = x_i,\ Y = y_j\},$$

удовлетворяющих условию

$$\sum_i \sum_j p_{ij} = 1,$$

где суммирование распространяется на все возможные значения индексов i и j, называется законом распределения С. В. Д. Т.

Одномерные законы распределения отдельных компонент С. В. Д. Т. выражаются через вероятности совместных значений p_{ij}, по формулам

$$p_{i\cdot} = \mathrm{P}\{X = x_i\} = \sum_j p_{ij}, \quad p_{\cdot j} = \mathrm{P}\{Y = y_j\} = \sum_i p_{ij}, \tag{1}$$

где суммирование распространяется на все возможные значения индексов i или j.

Пусть $(X,\,Y)$ — двумерный случайный вектор дискретного типа, $G\,(x,\,y)$ — множество его возможных значений. *Условным законом распределения* случайной компоненты X при условии, что компонента Y приняла определенное значение y_j, называется совокупность $\{x_i\,|\,x_i \in \in G\,(x,\,y_j)\}$ возможных значений компоненты X и соответствующих этим значениям условных вероятностей $\mathrm{P}\{X = x_i/Y = y_j\}$, определяемых равенством (см. § 1, формула (6))

$$\mathrm{P}\{X = x_i/Y = y_j\} = \frac{P\{X = x_i,\ Y = y_j\}}{P\{Y = y_j\}} = \frac{p_{ij}}{p_{\cdot j}}. \tag{2}$$

Если (X, Y) — С. В. Д. Т. и G — произвольная область на плоскости, то

$$\mathrm{P}\{(X, Y) \in G\} = \sum_{i}\sum_{\substack{j \\ (x_i,\, y_j)\in G}} p_{ij}. \tag{3}$$

Двумерный случайный вектор (X, Y) называется случайным *вектором непрерывного типа* (сокращенно С. В. Н. Т.), если функция распределения $F_{X,Y}(x, y)$ непрерывна на всей плоскости и существует такая неотрицательная интегрируемая по Риману в бесконечных пределах по каждой из координат функция $f_{X,Y}(x, y)$, называемая *плотностью распределения вероятностей* случайного вектора (X, Y), что

$$F_{X,Y}(x, y) = \int\limits_{-\infty}^{x} ds \int\limits_{-\infty}^{y} f_{X,Y}(s, t)\, dt.$$

Плотность распределения вероятностей обладает следующими свойствами:

1. $f_{X,Y}(x, y) \geqslant 0,\ (x, y) \in \mathbb{R}^2$.
2. $\int\limits_{-\infty}^{+\infty} ds \int\limits_{-\infty}^{+\infty} f_{X,Y}(s, t)\, dt = 1$ (условие нормировки).
3. Если (x, y) — точка непрерывности плотности $f_{X,Y}(x, y)$, то

$$f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial x\, \partial y}.$$

4. Плотности распределения вероятностей отдельных компонент случайного вектора выражаются в виде интегралов от совместной плотности:

$$f_X(x) = \int\limits_{-\infty}^{+\infty} f_{X,Y}(x, y)\, dy, \quad f_Y(y) = \int\limits_{-\infty}^{+\infty} f_{X,Y}(x, y)\, dx. \tag{4}$$

Если (X, Y) — С. В. Н. Т., то вероятность попадания случайной точки в произвольную квадрируемую область G на плоскости определяется по формуле

$$\mathrm{P}\{(X, Y) \in G\} = \iint\limits_{G} f_{X,Y}(x, y)\, dx\, dy. \tag{5}$$

Пусть (X, Y) — двумерный С. В. Н. Т. *Условной плотностью распределения вероятностей* случайной компоненты X при условии, что компонента Y приняла определенное значение y такое, что $f_Y(y) > 0$, называется неотрицательная функция $f_X(x/y)$ действительной переменной x, определяемая при всех $x \in \mathbb{R}$ следующей *формулой умножения для плотностей:*

$$f_X(x/y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}. \tag{6}$$

Аналогично, при всех $y \in \mathbb{R}$ и всех $x \in \mathbb{R}$ таких, что $f_X(x) > 0$,

$$f_Y(y/x) = \frac{f_{X,Y}(x, y)}{f_X(x)}.$$

Случайные величины $X_1, X_2, \ldots, X_n$ называются *независимыми (в совокупности)*, если для любого набора событий $\{X_i \in B\}$, $i = 1, 2, \ldots, n$, где $B_1, B_2, \ldots, B_n$ — подмножества числовой прямой, выполняется равенство

$$\mathrm{P}\{X_1 \in B_1,\ X_2 \in B_2, \ldots, X_n \in B_n\} = \mathrm{P}\{X_1 \in B_1\}\,\mathrm{P}\{X_2 \in B_2\} \ldots \mathrm{P}\{X_n \in B_n\}.$$

Теорема. *Случайные величины $X_1, X_2, \ldots, X_n$ независимы тогда и только тогда, когда в любой точке $x \in \mathbb{R}^n$ имеет место равенство*

$$F_X(x) = F_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = F_{X_1}(x_1)\,F_{X_2}(x_2) \ldots F_{X_n}(x_n).$$

Из этой теоремы, в частности, следует, что для независимости компонент случайного вектора непрерывного типа $(X_1, X_2, \ldots, X_n)$ необходимо и достаточно, чтобы в любой точке $x \in \mathbb{R}^n$

$$f_X(x) = f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1)\,f_{X_2}(x_2) \ldots f_{X_n}(x_n).$$

Если же $(X_1, X_2, \ldots, X_n)$ — С. В. Д. Т., то соответствующее условие независимости его компонент записывается в виде

$$\mathrm{P}\{X_1 = x_1,\ X_2 = x_2, \ldots, X_n = x_n\} = \mathrm{P}\{X_1 = x_1\}\,\mathrm{P}\{X_2 = x_2\} \ldots \mathrm{P}\{X_n = x_n\}$$

для всех $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

Для двумерного случайного вектора (X, Y) вводятся следующие числовые характеристики.

Начальным моментом порядка $k + s$ случайного вектора (X, Y) называется действительное число $\alpha_{k,s}$, определяемое формулой

$$\alpha_{k,s} = \mathrm{M}[X^k Y^s] = \begin{cases} \sum\limits_i \sum\limits_j x_i^k y_j^s p_{ij}, & \text{если } (X, Y) - \text{С. В. Д. Т.} \\ \iint\limits_{-\infty}^{+\infty} x^k y^s f_{X,Y}(x, y)\,dx\,dy, & \text{если } (X, Y) - \text{С. В. Н. Т.,} \end{cases}$$

Начальный момент $\alpha_{k,s}$ существует, если ряд (соответственно интеграл) в правой части равенства абсолютно сходится. В частности, $\alpha_{k,0} = \mathrm{M}[X^k]$, $\alpha_{0,s} = \mathrm{M}[Y^s]$ — соответствующие начальные моменты отдельных компонент.

Вектор с неслучайными координатами $(m_X, m_Y) = (\alpha_{1,0}, \alpha_{0,1})$ называется *математическим ожиданием* случайного вектора (X, Y) или *центром рассеивания.*

Центральным моментом порядка $k+s$ случайного вектора (X, Y) называется действительное число $\mu_{k,s}$, определяемое формулой

$$\mu_{k,s} = \begin{cases} \sum\limits_i \sum\limits_j (x_i - m_X)^k (y_j - m_Y)^s p_{ij}, & \text{если } (X, Y) - \text{С. В. Д. Т.} \\ \iint\limits_{-\infty}^{+\infty} (x - m_X)^k (y - m_Y)^s f_{X,Y}(x, y)\, dx\, dy, & \text{если } (X, Y) - \text{С. В. Н. Т.} \end{cases}$$

Центральный момент $\mu_{k,s}$ существует, если ряд (соответственно интеграл) в правой части равенства абсолютно сходится. В частности, $\mu_{2,0} = D_X$, $\mu_{0,2} = D_Y$.

Центральный момент $\mu_{1,1}$ называется *ковариацией* и обозначается K_{XY}. Таким образом, по определению

$$K_{XY} = \mu_{1,1} = \mathrm{M}[(X - m_X)(Y - m_Y)] = \mathrm{M}[XY] - m_X m_Y.$$

Нормированная ковариация $\rho_{XY} = \dfrac{K_{XY}}{\sigma_X \sigma_Y}$ называется *коэффициентом корреляции* двух случайных компонент X и Y случайного вектора.

Коэффициент корреляции удовлетворяет условию $|\rho_{XY}| \leqslant 1$ и определяет степень линейной зависимости между X и Y. Случайные величины, для которых $\rho_{XY} = 0$, называются *некоррелированными.* Из независимости случайных величин X и Y вытекает их некоррелированность (обратное, вообще говоря, неверно).

Пусть (X, Y) — С. В. Д. Т. *Условным математическим ожиданием случайной величины* X *при условии, что* Y *приняла одно из своих возможных значений* y_j, называется действительное число $\hat{y}_j$, обозначаемое также $\mathrm{M}[X/Y = y_j]$ и определяемое формулой

$$\hat{y}_j = \mathrm{M}[X/Y = y_j] = \sum_i x_i \mathrm{P}\{X = x_i/Y = y_j\}, \tag{7}$$

где $\mathrm{P}\{X = x_i/Y = y_j\}$ — условная вероятность, определяемая формулой (2), а суммирование в правой части распространяется на все возможные значения индекса i.

Условным математическим ожиданием случайной величины X *при условии* Y называется случайная величина $\hat{Y}$, обозначаемая также $\mathrm{M}[X/Y]$, возможные значения которой $\hat{y}_j$ определяются формулой (7), а соответствующие вероятности равны

$$P\{\hat{Y} = \hat{y}_j\} = \mathrm{P}\{Y = y_j\}.$$

Имеет место следующая *формула полного математического ожидания*:

$$\mathrm{M}[X] = \mathrm{M}[\mathrm{M}[X/Y]] = \sum_j \mathrm{M}[X/Y_j = y_j]\,\mathrm{P}\{Y = y_j\}. \tag{8}$$

Пусть (X, Y) — С. В. Н. Т. *Условным математическим ожиданием случайной величины* X *при условии* $Y = y$ называется действительное число, обозначаемое $\mathrm{M}[X/Y = y]$ и определяемое формулой

$$\mathrm{M}[X/Y = y] = \int_{-\infty}^{+\infty} x f_X(x/y)\,dx,$$

где $f_X(x/y)$ — условная плотность, определяемая формулой (6). *Формула полного математического ожидания*, аналогичная (8), имеет вид

$$\mathrm{M}[X] = \mathrm{M}[\mathrm{M}[X/Y]] = \int_{-\infty}^{+\infty} f_Y(y)\,\mathrm{M}[X/Y = y]\,dy. \tag{9}$$

Приведенные выше формулы для числовых характеристик двумерного случайного вектора (X, Y) без труда обобщаются на случай n-мерного случайного вектора $(X_1, X_2, \ldots, X_n)$. Так, например, вектор с неслучайными координатами $(m_1, m_2, \ldots, m_n)$, где m_i — математическое ожидание случайной величины X_i, определяемое формулой

$$m_i = \mathrm{M}[X_i] =$$
$$= \iint_{-\infty}^{+\infty}\cdots\int x_i\, f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n)\,dx_1\,dx_2\ldots dx_n,$$

называется *центром рассеивания* случайного вектора $(X_1, X_2, \ldots, X_n)$.

Ковариационной матрицей n-мерного случайного вектора $(X_1, X_2, \ldots, X_n)$ называется симметрическая действительная матрица, элементы которой представляют собой ковариации соответствующих пар компонент:

$$K = \begin{pmatrix} K_{11} & K_{12} & \ldots & K_{1n} \\ & K_{22} & \ldots & K_{2n} \\ & & \ldots & \ldots \\ & & & K_{nn} \end{pmatrix},$$

где $K_{ij} = \mathrm{M}\,[\mathring{X}_i\mathring{X}_j] = K_{ji}$ — ковариация i-й и j-й компонент. Очевидно, $K_{ii} = \mathrm{M}\,[\mathring{X}_i^2] = D_i$ — дисперсия i-й компоненты.

Корреляционной матрицей n-мерного случайного вектора называется нормированная ковариационная матрица

$$R = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \dots & \rho_{1n} \\ & 1 & \rho_{23} & \dots & \rho_{2n} \\ & & & \dots & \dots \\ & & & & 1 \end{pmatrix},$$

где $\rho_{ij} = \dfrac{K_{ij}}{\sigma_i \sigma_j}$ — коэффициент корреляции i-й и j-й компонент.

Пример 1. Дважды бросается игральная кость. Случайные величины: X — число появлений шестерки, Y — число появлений четной цифры.

а) Описать закон распределения случайного вектора (X, Y).

б) Установить, зависимы или независимы компоненты X и Y.

в) Описать законы распределения отдельных компонент.

г) Вычислить вероятность $\mathrm{M}\{X \geqslant Y\}$.

◁ Для описания закона распределения дискретного случайного вектора (X, Y) необходимо определить множество всех возможных пар значений (x_i, y_j) и соответствующие вероятности. Результат удобно представлять в виде следующей таблицы:

x_i	y_j			$\mathrm{P}\{X = x_i\}$
	0	1	2	
0	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{9}$	$\frac{25}{36}$
1	0	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{10}{36}$
2	0	0	$\frac{1}{36}$	$\frac{1}{36}$
$\mathrm{P}\{Y = y_j\}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

В первом столбце указываются возможные значения случайной величины X, а в первой строке — возможные значения Y; в последнем столбце и в последней строке указываются безусловные вероятности возможных значений соответственно X и Y. В каждой клетке таблицы, стоящей на пересечении i-й строки и j-го столбца, указываются вероятности совместного осуществления события $\{X = x_i,\ Y = y_j\}$, т.е. $p_{ij} = \mathrm{P}\{X = x_i,\ Y = y_j\}$.

Множество Ω для данного эксперимента состоит из равновероятных исходов следующего вида:

$$\Omega = \{(i, j) \,|\, i, j = 1, \dots, 6\};$$

число элементов данного множества $N(\Omega) = 6^2 = 36$ (схема выбора двух элементов из множества $\{1, 2, \ldots, 6\}$ с возвращением и упорядочиванием). Заметим также, что данный эксперимент можно рассматривать как повторение двух независимых испытаний с одним и тем же множеством элементарных исходов в каждом опыте $\Omega_i = \{1, 2, \ldots, 6\}$ для $i = 1, 2$. Поэтому множество Ω является прямым произведением $\Omega_1 \times \Omega_2$, что соответствует приведенной выше записи.

Случайные величины X и Y определены на множестве Ω и ставят в соответствие каждому элементарному исходу одно из своих возможных значений 0, 1 или 2 (множества возможных значений для обеих компонент совпадают).

Для данного опыта проще всего сначала описать законы распределения отдельных компонент X и Y, т.е. ответить на вопрос в).

Событию $\{X = i\}$, $i = 0, 1, 2$, соответствует множество таких исходов из Ω, у которых встречается ровно i шестерок. Вероятность данного события равна вероятности получить ровно i успехов в двух опытах по схеме Бернулли с вероятностью успеха в данном опыте, равной $p = 1/6$, поэтому по формуле (7) § 2

$$\mathrm{P}\{X = i\} = P_{2,i}\left(\frac{1}{6}\right) = C_2^i \left(\frac{1}{6}\right)^i \left(\frac{5}{6}\right)^{2-i}, \quad i = 0, 1, 2.$$

Аналогично, используя соответствующую схему Бернулли для вычисления вероятностей событий $\{Y = j\}$, $j = 0, 1, 2$, получим

$$\mathrm{P}\{Y = j\} = P_{2,j}\left(\frac{1}{2}\right) = C_2^j \left(\frac{1}{2}\right)^j \left(\frac{1}{2}\right)^{2-j} = C_2^j \cdot \frac{1}{4}, \quad j = 0, 1, 2.$$

Таким образом, заполняем последний столбец и последнюю строку таблицы.

а) Описание закона распределения случайного вектора дискретного типа (X, Y) можно осуществить в такой последовательности. Заметим, прежде всего, что

$$p_{10} = p_{20} = p_{21} = 0,$$

так как данные вероятности относятся к невозможным событиям (соответствующие этим событиям множества элементарных исходов пусты).

Используя условие нормировки «по столбцу» $\sum\limits_{i=0}^{2} p_{i0} = p_{\cdot 0} = \mathrm{P}\{Y = 0\}$, находим, что $p_{00} = P\{X = 0,\ Y = 0\} = 1/4$. Аналогичное условие нормировки «по строке» дает $p_{22} = \mathrm{P}\{X = 2,\ Y = 2\} = 1/36$.

Нетрудно убедиться, что достаточно вычислить еще одну какую-либо вероятность из четырех оставшихся $p_{i,j}$, после чего все остальные вероятности в таблице восстанавливаются из условия нормировки. Вычислим, например, p_{11}. Вычисления можно провести двумя способами.

Первый способ. Непосредственно вычисляем вероятность совместного осуществления двух событий: $\{X = 1\}$ и $\{Y = 1\}$. Очевидно.

что событию $\{X = 1,\ Y = 1\}$ благоприятствуют только такие исходы $(i, j) \in \Omega$, у которых либо на первом месте стоит шестерка, а на втором месте любая нечетная цифра, либо наоборот. Поэтому $\{X = 1,$ $Y = 1\} = \{(6, 1), (6,3), (6,5), (1,6), (3,6), (5,6)\}$. Отсюда по формуле классической вероятности получаем

$$p_{11} = \frac{N(\{X = 1,\ Y = 1\})}{N(\Omega)} = \frac{1}{6}.$$

Второй способ. Используем формулу умножения вероятностей

$$\mathrm{P}\{X = 1,\ Y = 1\} = \mathrm{P}\{X = 1\}\,\mathrm{P}\{Y = 1/X = 1\}.$$

Условную вероятность находим методом вспомогательного эксперимента: $\mathrm{P}\{Y = 1/X = 1\} = \mathrm{P}\{$при одном подбрасывании выпала нечетная цифра из состава $\{1,2,3,4,5\}\} = 3/5$ согласно формуле классической вероятности. Учитывая, что $\mathrm{P}\{X = 1\} = 10/36$, получаем

$$\mathrm{P}\{X = 1,\ Y = 1\} = \frac{10}{36}\cdot\frac{3}{5} = \frac{1}{6}.$$

Оставшиеся клетки таблицы заполняем с помощью условия нормировки, например, в такой последовательности:

$$p_{01} = \frac{1}{2} - \frac{1}{6} = \frac{1}{3},$$

$$p_{12} = \frac{10}{36} - \frac{1}{6} = \frac{1}{9}.$$

Наконец,

$$p_{02} = \frac{25}{36} - \frac{1}{3} - \frac{1}{4} = \frac{1}{9}.$$

б) Компоненты X и Y зависимы, так как, например,

$$p_{11} = \mathrm{P}\{X = 1,\ Y = 1\} = \frac{1}{6} \neq \mathrm{P}\{X = 1\}\cdot\mathrm{P}\{Y = 1\} = \frac{10}{36}\cdot\frac{1}{2} = \frac{5}{36}.$$

в) Искомая вероятность определяется как вероятность попадания в область на плоскости, определяемую соответствующим неравенством,

$$\mathrm{P}\{X \geqslant Y\} = \mathrm{P}\{(X, Y) \in G\},$$

где $G = \{(x, y) \mid (x, y) \in \mathbb{R}^2,\ x \geqslant y\}$. Так как (X, Y) — С. В. Д. Т., то по формуле (3) получаем

$$\mathrm{P}\{X \geqslant Y\} = \sum_{\substack{i=0 \\ (x_i \geqslant y_j)}}^{2}\sum_{j=0}^{2} p_{ij} = p_{00} + p_{10} + p_{11} + p_{20} + p_{21} + p_{22} =$$

$$= \frac{1}{4} + \frac{1}{6} + \frac{1}{36} = \frac{4}{9}. \ \triangleright$$

Пример 2 (продолжение). Описать функцию распределения вектора из примера 1.

◁ Согласно определению функции распределения

$$F_{X,Y}(X, Y) = \mathrm{P}\{X < x,\ Y < y\} = \mathrm{P}\{(X, Y) \in G_{x,y}\},$$

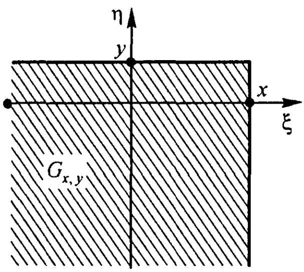


Рис. 8

где

$$G_{x,y} = \{(\xi, \eta) | (\xi, \eta) \in \mathbb{R}^2,\ \xi < x,\ \eta < y\}.$$

Область $G_{x,y}$ заштрихована на рис. 8 и представляет собой бесконечный прямоугольник с вершинами в точках $(-\infty, -\infty)$, $(-\infty, y)$, (x, y) и $(x, -\infty)$.

При каждом фиксированном значении точки с координатами (x, y) значение $F_{X,Y}(x_i, y_j)$ равно сумме вероятностей тех возможных значений вектора (x_i, y_j), которые попадают внутрь указанного прямоугольника. Результат удобно представить в виде следующей таблицы:

x	y			
	$y \leqslant 0$	$0 < y \leqslant 1$	$1 < y \leqslant 2$	$2 < y$
$x \leqslant 0$	0	0	0	0
$0 < x \leqslant 1$	0	1/4	7/12	25/36
$1 < x \leqslant 2$	0	1/4	3/4	31/36
$2 < x$	0	1/4	3/4	1

Пример 3 (продолжение). Вычислить основные характеристики m_X, m_Y, D_X, D_Y, K_{XY}, ρ_{XY} случайного вектора из примера 1.

◁ По формулам для начальных моментов первого порядка имеем

$$m_X = \alpha_{1,0} = \sum_{i=0}^{2}\sum_{j=0}^{2} i p_{ij} = \sum_{i=0}^{2} i \sum_{j=0}^{2} p_{ij} = \sum_{i=0}^{2} i p_{i\cdot} = \frac{1}{3},$$

$$m_Y = \alpha_{0,1} = \sum_{i=0}^{2}\sum_{j=0}^{2} j p_{ij} = \sum_{j=0}^{2} j \sum_{i=0}^{2} p_{ij} = \sum_{j=0}^{2} j p_{\cdot j} = 1.$$

Дисперсии удобно вычислять через второй начальный момент:

$$D_X = \mu_{2,0} = \alpha_{2,0} - m_X^2 = \sum_{i=0}^{2} i^2 p_{i\cdot} - m_X^2 = \frac{5}{18},$$

$$D_Y = \mu_{0,2} = \alpha_{0,2} - m_Y^2 = \sum_{j=0}^{2} j^2 p_{\cdot j} - m_Y^2 = \frac{1}{2}.$$

Для вычисления ковариации используем аналогичную формулу через начальный момент $\alpha_{1,1}$:

$$K_{XY} = \alpha_{1,1} - m_X m_Y = \sum_{i=0}^{2}\sum_{j=0}^{2} ijp_{ij} - m_X m_Y = \frac{1}{6}.$$

Коэффициент корреляции определяется как нормированная ковариация, поэтому

$$\rho_{XY} = \frac{K_{XY}}{\sigma_X \sigma_Y} = \frac{K_{XY}}{\sqrt{D_X}\,\sqrt{D_Y}} = \frac{1}{\sqrt{5}}. \ \triangleright$$

Пример 4 (продолжение). Описать условный закон распределения случайной величины X из примера 1 при условии $Y = 2$ и вычислить условное математическое ожидание $\mathrm{M}[X/Y = 2]$.

$\triangleleft$ Условный закон распределения случайной компоненты X при условии, что компонента Y приняла значение, равное 2, находим, используя формулу (2):

$$\mathrm{P}\{X = i/Y = 2\} = \frac{\mathrm{P}\{X = i,\ Y = 2\}}{\mathrm{P}\{Y = 2\}} = \frac{p_{i2}}{p_{\cdot 2}} = 4p_{i2}, \quad i = 0,\,1,\,2.$$

Результат оформим в виде таблицы

x_i	0	1	2	$\sum\limits_{i=0}^{2} \mathrm{P}\{X = x_i/Y = 2\}$
$\mathrm{P}\{X = x_i/Y = 2\}$	4/9	4/9	1/9	1

При этом условное математическое ожидание $\mathrm{M}[X/Y = 2] = 2/3$. $\triangleright$

В задачах 18.378–18.403 рассматриваются случайные векторы дискретного типа.

18.378. Закон распределения случайного вектора (X, Y) дискретного типа определяется следующей таблицей:

x_i	y_j		
	-1	0	1
1	0,15	0,3	0,35
2	0,05	0,05	0,1

а) Найти безусловные законы распределения отдельных компонент X и Y.

б) Установить, зависимы или нет компоненты X и Y?

в) Вычислить вероятности $\mathrm{P}\{X = 2, Y \geqslant 0\}$ и $\mathrm{P}\{X > Y\}$.

18.379 (продолжение). Построить функцию распределения $F_{X,Y}(x, y)$, оформив результат в виде таблицы, и найти m_X и m_Y.

18.380 (продолжение). Вычислить ковариационную матрицу случайного вектора из задачи 18.378.

18.381 (продолжение). Для случайного вектора из задачи 18.378 описать условный закон распределения случайной величины Y при условии $X = 1$ и найти условное математическое ожидание $\mathrm{M}[Y/X = 1]$.

18.382. В условиях примера 1 вычислить вероятности p_{12} и p_{01}, не используя условие нормировки. Чему равна вероятность $\mathrm{P}\{X + Y > 1\}$?

18.383. Один раз подбрасывается игральная кость. Случайные величины: X — индикатор четного числа выпавших очков ($X = 1$, если выпало четное число очков, и $X = 0$ в противном случае), Y — индикатор числа очков, кратного трем ($Y = 1$, если выпало число очков, кратное трем, и $Y = 0$ в противном случае). Описать закон распределения случайного вектора (X, Y) и безусловные законы распределения компонент.

18.384 (продолжение). Вычислить момент $\mu_{2,2}$ для распределения случайного вектора предыдущей задачи.

18.385. Иван и Петр наудачу извлекают по одному шару из урны, содержащей 6 белых и 4 черных шара. Иван извлекает шар первым. Случайные величины: X — число белых шаров у Ивана, Y — число белых шаров у Петра. Описать закон распределения случайного вектора (X, Y) и безусловные законы распределения компонент, если выбор шаров производится без возвращения.

18.386 (продолжение). Решить задачу 18.385 при условии, что выбор шаров осуществляется с возвращением. В каком случае вероятность события $\{X > Y\}$ больше: в опыте из задачи 18.385 или в данном опыте?

18.387. Случайная величина X принимает значения 0, 1 или 2 с вероятностями соответственно 0,2; 0,7 и 0,1, а не зависящая от нее случайная величина Y — значения -1, 0, 1 соответственно с вероятностями 0,3; 0,5; 0,2. Описать закон распределения случайного вектора (X, Y) и вычислить функцию распределения в точках (1,5; -0,5) и (0,5; 4).

18.388. По цели производится два независимых выстрела. Вероятность попадания в цель при первом выстреле равна p_1, при втором p_2. Случайные величины: X — число попаданий при первом выстреле, Y — число попаданий при втором выстреле. Найти функцию распределения $F_{X,Y}(x, y)$.

18.389. Из урны, содержащей 6 белых и 4 черных шара, наудачу извлекают 2 шара без возвращения. Случайные величины: X — число белых шаров в выборке, Y — число черных шаров в выборке. Описать закон распределения случайного вектора (X, Y) и вычислить ρ_{XY}.

18.390. Производится два выстрела по мишени в неизменных условиях. Вероятность попадания в мишень при одном выстреле равна p. Случайные величины: X — число выстрелов до первого попадания (включительно), Y — число промахов. Описать закон распределения случайного вектора (X, Y). Найти центр рассеивания данного распределения и значения σ_X^2 и σ_Y^2.

18.391 (продолжение). Вычислить вероятность события $\{X \geqslant Y+1\}$.

18.392 (продолжение). Вычислить коэффициент корреляции для компонент случайного вектора из задачи 18.390.

18.393. Функция распределения случайного вектора (X, Y) дискретного типа имеет вид, определяемый таблицей

x	y			
	$y \leqslant 1$	$1 < y \leqslant 3$	$3 < y \leqslant 5$	$5 < y$
$x \leqslant 2$	0	0,25	0,25	0,25
$2 < x \leqslant 4$	0	0,25	0,3	0,4
$4 < x$	0	0,4	0,75	1

Описать закон распределения случайного вектора (X, Y) и найти центр рассеивания.

18.394. Бросаются две одинаковые игральные кости. Случайные величины: X — индикатор четности суммы выпавших очков (т.е. $X = 1$, если эта сумма четна, и $X = 0$ в противном случае), Y — индикатор четности произведения выпавших очков (т.е. $Y = 1$, если это произведение четно, и $Y = 0$ в противном случае). Описать закон распределения случайного вектора (X, Y).

18.395 (продолжение). Построить функцию распределения случайного вектора из предыдущей задачи.

18.396 (продолжение). Найти ковариационную матрицу случайного вектора из задачи 18.394.

18.397. Число X выбирается случайным образом из множества целых чисел $\{1, 2, 3\}$. Затем из того же множества выбирается наудачу число Y, большее первого или равное ему. Описать закон распределения случайного вектора (X, Y) и определить, зависимы или независимы случайные компоненты X и Y.

18.398 (продолжение). В условиях предыдущей задачи найти коэффициент корреляции ρ_{XY}.

18.399 (продолжение). В условиях задачи 18.397 описать условный закон распределения компоненты X при условии $Y = 3$ и вычислить $\mathrm{M}[X/Y = 3]$.

18.400*. В продукции завода брак вследствие дефекта α составляет 3%, а вследствие дефекта β — 4,5%. Годная продукция составляет 95%. Определить, какой процент всей продукции обладает дефектами обоих типов.

18.401 (продолжение). В условиях предыдущей задачи найти коэффициент корреляции дефектов α и β.

18.402*. В продукции завода брак вследствие дефекта α составляет 6%, причем среди забракованной по признаку α продукции в 4% случаев встречается дефект β, а в продукции, свободной от дефекта α, дефект β встречается в 1% случаев. Найти вероятность встретить дефект β во всей продукции.

18.403 (продолжение). В условиях предыдущей задачи вычислить коэффициент корреляции дефектов α и β.

Пример 5. Плотность распределения вероятностей случайного вектора (X, Y) непрерывного типа задана следующим образом:

$$f_{X,Y}(x, y) = \begin{cases} 0, & x^2 + y^2 \geqslant a^2, \\ c\sqrt{a^2 - (x^2 + y^2)}, & x^2 + y^2 < a^2. \end{cases}$$

Определить константу c и вычислить $\mathrm{P}\{(X, Y) \in G\}$, где $G = \{(x, y) \mid x > 0,\ y > 0\}$ — первый квадрант плоскости Oxy.

◁ Константу c находим из условия нормировки, которое в данном случае записывается в виде

$$c \iint\limits_{x^2+y^2\leqslant a^2} \sqrt{a^2 - (x^2 + y^2)}\, dx\, dy = 1.$$

Переходя к полярной системе координат, находим

$$c\int_0^{2\pi} d\varphi \int_0^a r\sqrt{a^2 - r^2}\, dr = \frac{c \cdot 2\pi a^3}{3} = 1,$$

откуда следует, что $c = \dfrac{3}{2\pi a^3}$.

Вероятность попадания в область G для случайного вектора (X, Y) непрерывного типа определяется, согласно формуле (5), как двойной интеграл от плотности, взятый по области G; поэтому

$$\mathrm{P}\{(X, Y) \in G\} = \frac{3}{2\pi a^3} \int_0^{\pi/2} d\varphi \int_0^a r\sqrt{a^2 - r^2}\, dr = \frac{1}{4}. \ \triangleright$$

Пример 6 (продолжение). Вычислить центр рассеивания и ковариационную матрицу случайного вектора из примера 5.

◁ В силу осевой симметрии функции плотности очевидно, что $(m_X, m_Y) = (0, 0)$. Отсюда следует, что для данного распределения $D_X = \alpha_{2,0}$, $D_Y = \alpha_{0,2}$, и $K_{XY} = \alpha_{1,1}$. Записывая соответствующие моменты в полярной системе координат, получаем

$$D_X = \frac{3}{2\pi a^3}\int_0^{2\pi}\cos^2\varphi\,d\varphi\int_0^a r^3\sqrt{a^2-r^2}\,dr =$$

$$= \frac{3}{4\pi a^3}\int_0^{2\pi}d\varphi\int_0^a r^3\sqrt{a^2-r^2}\,dr + \frac{3}{4\pi a^3}\int_0^{2\pi}\cos 2\varphi\,d\varphi\int_0^a r^3\sqrt{a^2-r^2}\,dr =$$

$$= \frac{3}{4a^3}\int_0^a (a^2-r^2)^{3/2}\,d(a^2-r^2) - \frac{3}{4a}\int_0^a (a^2-r^2)^{1/2}\,d(a^2-r^2) =$$

$$= \frac{3}{4a^3}\left(\frac{2}{5}(a^2-r^2)^{5/2} - \frac{2}{3}a^2(a^2-r^2)^{3/2}\right)\Bigg|_0^a = \frac{3a^2}{2}\left(\frac{1}{3}-\frac{1}{5}\right) = \frac{a^2}{5}.$$

Так как выражение для $\alpha_{0,2}$ сводится к тому же интегралу, то $D_Y = a^2/5$.

Ковариация находится следующим образом:

$$K_{XY} = \frac{3}{2\pi a^3}\int_0^{2\pi}\sin\varphi\cos\varphi\,d\varphi\int_0^a r^3\sqrt{a^2-r^2}\,dr = 0.$$

Таким образом, ковариационная матрица имеет вид

$$K = \begin{pmatrix} a^2/5 & 0 \\ 0 & a^2/5 \end{pmatrix}. \ \triangleright$$

В задачах 18.404–18.422 рассматриваются случайные векторы непрерывного типа.

18.404. Плотность распределения вероятностей случайного вектора (X, Y) имеет следующий вид:

$$f_{X,Y}(x, y) = \begin{cases} c(x+y) & \text{при } 0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 1, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Определить константу c и вычислить $\mathrm{P}\{X + Y < 1\}$.

18.405 (продолжение). В условиях предыдущей задачи определить безусловную плотность распределения компоненты X и установить, зависимы компоненты X и Y или нет.

18.406 (продолжение). Для случайного вектора из задачи 18.404 вычислить центр рассеивания и функцию распределения $F_X(x)$.

18.407. Плотность распределения вероятностей двумерного случайного вектора (X, Y) имеет следующий вид:

$$f_{X,Y}(x, y) = \begin{cases} c\,(xy + y^2) & \text{при } 0 \leqslant x \leqslant 2,\ 0 \leqslant y \leqslant 2, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Вычислить значение постоянной c и вероятность $\mathrm{P}\{X + Y < 2\}$.

18.408 (продолжение). Найти центр рассеивания случайного вектора из предыдущей задачи.

18.409. Функция распределения случайного вектора (X, Y) непрерывного типа задана в виде

$$F_{X,Y}(x, y) = \begin{cases} 0, & \text{если } x < 0 \text{ или } y < 0, \\ \dfrac{1}{2}(\sin x + \sin y - \sin(x + y)), & \text{если } 0 \leqslant x \leqslant \pi/2, \\ & \phantom{\text{если }} 0 \leqslant y \leqslant \pi/2, \\ 1, & \text{если } x > \pi/2 \text{ и } y > \pi/2. \end{cases}$$

Вычислить вероятности $p_1 = \mathrm{P}\{(X,Y) \in G_1\}$ и $p_2 = \mathrm{P}\{(X,Y) \in G_2\}$, где области G_1 и G_2 изображены на рис. 9 и 10 соответственно.

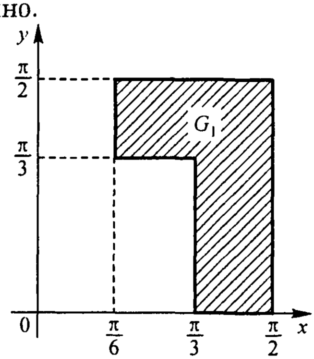


Рис. 9

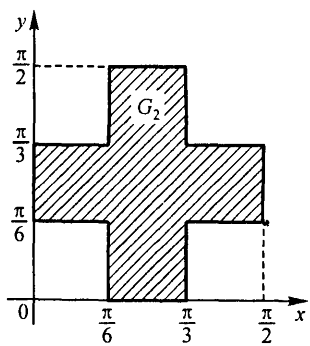


Рис. 10

18.410 (продолжение). Вычислить плотность распределения вероятностей случайного вектора из предыдущей задачи и найти его центр рассеивания.

Говорят, что случайный вектор (X, Y) непрерывного типа *распределен равномерно в области* $G \subset \mathbb{R}^2$, если G — квадрируемая область и плотность распределения вероятностей такого вектора имеет вид

$$f_{X,Y}(x, y) = \begin{cases} 0, & \text{если} \quad (x, y) \notin G, \\ \dfrac{1}{S(G)}, & \text{если} \quad (x, y) \in G, \end{cases}$$

где $S(G)$ — площадь области G.

18.411. Случайный вектор (X, Y) распределен равномерно внутри прямоугольника $G = \{(x, y) \mid -1 \leqslant x \leqslant 2,\ 1 \leqslant y \leqslant 2\}$. Получить выражения для плотности распределения $f_{X,Y}(x, y)$, функции распределения $F_{X,Y}(x, y)$ и установить, зависимы или нет компоненты X и Y.

18.412 (продолжение). В условиях предыдущей задачи вычислить центр рассеивания (m_X, m_Y) и дисперсии D_X и D_Y.

18.413 (продолжение). В условиях задачи 18.411 вычислить вероятности $p_i = \mathrm{P}\{(X, Y) \in G_i\}$, $i = 1, 2$, где $G_1 = \{(x, y) \mid 0{,}5 \leqslant x \leqslant 1{,}5,\ 1{,}5 \leqslant y \leqslant 2{,}5\}$, $G_2 = \{(x, y) \mid x^2 + (y - 1{,}5)^2 \leqslant 0{,}25\}$.

18.414. Случайные величины X и Y независимы и распределены по законам $R(-1, 1)$и $R(0, 2)$ соответственно. Найти выражение для плотности $f_{X,Y}(x, y)$, вычислить вероятность события $A = \{(x, y) \in G\}$, где область G — треугольник с вершинами в точках $(-1, 0)$, $(0, 1)$, $(1, 0)$.

18.415. Случайный вектор (X, Y) распределен равномерно в треугольнике с вершинами в точках $(-1, 0)$, $(1, 2)$ и $(1, 0)$. Вычислить центр рассеивания данного распределения.

18.416. Случайный вектор (X, Y) равномерно распределен в квадрате со стороной a и диагоналями, совпадающими с осями координат. Установить, зависимы или независимы его компоненты.

18.417 (продолжение). Вычислить ковариационную матрицу случайного вектора из предыдущей задачи и выяснить, коррелированы или некоррелированы его компоненты.

18.418 (продолжение). Для случайного вектора из задачи 18.416 вычислить вероятности $p_1 = \mathrm{P}\{XY > 0\}$ и $p_2 = \mathrm{P}\left\{X^2 + Y^2 < \dfrac{a^2}{4}\right\}$

18.419. Точка M наудачу ставится в круг $G = \{(x, y) \mid x^2 + y^2 < a^2\}$. Исследовать вопрос о коррелированности или некоррелированности, а также зависимости или независимости случайных координат X и Y точки M.

18.420. Функция совместного распределения двух случайных величин X и Y имеет следующий вид:

$$F_{X,Y}(x, y) = \begin{cases} 0, & \text{если} \quad \min(x, y) < 0, \\ \min(x, y), & \text{если} \quad 0 \leqslant \min(x, y) < 1, \\ 1, & \text{если} \quad \min(x, y) \geqslant 1. \end{cases}$$

Найти одномерные законы распределения компонент и решить вопрос об их зависимости или независимости.

18.421. Случайные величины X и Y независимы и распределены следующим образом; X — по закону $\mathrm{Ex}(2)$, Y — по закону $R(-2, 2)$. Найти вероятности следующих событий: $A = \{(X, Y) \in D_1\}$, $B = \{(X, Y) \in D_2\}$, где $D_1 = \{(x, y) \mid 0 \leqslant x \leqslant 2,\ 0 \leqslant y \leqslant 1\}$, $D_2 = \{(x, y) \mid y > x - 2\}$.

18.422. Случайные величины X и Y независимы и распределены каждая по показательному закону с параметрами соответственно λ_1 и λ_2. Найти $\mathrm{P}\{X \geqslant \lambda_1^{-1},\ Y \geqslant \lambda_2^{-1}\}$.

2. Нормальный закон на плоскости. Говорят, что двумерный случайный вектор (X, Y) *распределен по нормальному* (*гауссовскому*) *закону*, если совместная плотность распределения вероятностей случайных компонент имеет вид

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \exp\left\{-\frac{1}{2(1-\rho_{XY}^2)}\left[\frac{(x-m_X)^2}{\sigma_X^2} - \frac{2\rho_{XY}(x-m_X)(y-m_Y)}{\sigma_X\sigma_Y} + \frac{(y-m_Y)^2}{\sigma_Y^2}\right]\right\}, \quad (10)$$

где (m_X, m_Y) — центр рассеивания, σ_X, σ_Y — среднеквадратичные отклонения случайных компонент вектора (X, Y), ρ_{XY} — коэффициент корреляции.

Для нормального закона справедливо следующее правило:

Если компоненты двумерного нормального вектора некоррелированы ($\rho_{XY} = 0$), *то они и независимы*, так как в этом случае

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y} \exp\left\{-\left[\frac{(x-m_X)^2}{2\sigma_X^2} + \frac{(y-m_Y)^2}{2\sigma_Y^2}\right]\right\} = f_X(x)\, f_Y(y). \quad (11)$$

При этом оси координат Ox и Oy называются *главными осями рассеивания*. Если к тому же $\sigma_X = \sigma_Y = \sigma$, то рассеивание называется *круговым*. В общем случае ($\rho_{XY} \neq 0$, $\sigma_X \neq \sigma_Y$) *эллипсом рассеивания* называется эллипс, в каждой точке которого плотность имеет одно и то же постоянное значение. Эллипс, в точках которого плотность постоянна и равна

$$f_{X,Y}(x, y) = \mathrm{const} = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \exp\left\{-\frac{\lambda^2}{2(1-\rho_{XY}^2)}\right\},$$

описывается уравнением

$$\frac{(x-m_X)^2}{\sigma_X^2}-\frac{2\rho_{XY}(x-m_X)(y-m_Y)}{\sigma_X\sigma_Y}+\frac{(y-m_Y)^2}{\sigma_Y^2}=\lambda^2.$$

С помощью преобразования поворота системы координат плотность нормального распределения всегда может быть приведена к *каноническому виду* (11) с *главным эллипсом рассеивания*, описываемым уравнением

$$\frac{(x-m_X)^2}{\sigma_X^2}+\frac{(y-m_Y)^2}{\sigma_Y^2}=\lambda^2.$$

Пример 7. Случайный вектор (X, Y) подчиняется каноническому нормальному распределению с параметрами $m_X=m_Y=0$, σ_X, σ_Y. Вычислить вероятность попадания случайной точки (X, Y) в область D_λ, ограниченную главным эллипсом рассеивания с полуосями

$$a=\lambda\sigma_X, \quad b=\lambda\sigma_Y.$$

◁ Уравнение эллипса

$$\frac{x^2}{(\lambda\sigma_X)^2}+\frac{y^2}{(\lambda\sigma_Y)^2}=1,$$

$$\mathrm{P}\{(X, Y)\in D_\lambda\}=\iint\limits_{D_\lambda} f_{X,Y}(x, y)\,dx\,dy.$$

Для вычисления интеграла перейдем к полярной системе координат:

$$x=\sigma_X r\cos\varphi, \quad y=\sigma_Y r\sin\varphi.$$

Якобиан этого преобразования $I=\sigma_X\sigma_Y r$. При этом уравнение эллипса преобразуется в уравнение окружности радиуса λ. Поэтому

$$\mathrm{P}\{(X, Y)\in D_\lambda\}=\frac{1}{2\pi}\int\limits_0^{2\pi}d\varphi\int\limits_0^{\lambda}re^{-r^2/2}\,dr=1-e^{-\lambda^2/2}. \ \triangleright$$

Пример 8. Двумерный случайный вектор распределен по нормальному закону с совместной плотностью вероятностей, определяемой формулой (10). Найти безусловную плотность вероятностей компоненты X, условную плотность $f_Y(y/x)$ и условное математическое ожидание $\mathrm{M}[Y/X=x]$.

◁ По формуле (4)

$$f_X(x)=\int\limits_{-\infty}^{+\infty}f_{X,Y}(x, y)\,dy,$$

где $f_{X,Y}(x, y)$ определяется формулой (10). Сделаем замену переменных: $\dfrac{x-m_X}{\sqrt{2}\,\sigma_X}=u$, $\dfrac{y-m_Y}{\sqrt{2}\,\sigma_Y}=v$, и обозначим для краткости $\rho=\rho_{XY}$. Тогда

$$f_X(x)=\frac{1}{\sqrt{2}\pi\sigma_X\sqrt{1-\rho^2}}\exp\left(-\frac{u^2}{1-\rho^2}\right)\times \\ \times\int\limits_{-\infty}^{+\infty}\exp\left(-\frac{1}{1-\rho^2}v^2+\frac{2\rho uv}{1-\rho^2}\right)dv.$$

Дополняя до полного квадрата в показателе экспоненты под знаком интеграла, получим интеграл Пуассона:

$$f_X(x)=\frac{1}{\sqrt{2}\pi\sigma_X\sqrt{1-\rho^2}}\exp\left(-\frac{u^2-\rho^2u^2}{1-\rho^2}\right)\int\limits_{-\infty}^{+\infty}\exp\left\{-\frac{(v-u\rho)^2}{1-\rho^2}\right\}dv= \\ =\frac{1}{\sqrt{2\pi}\,\sigma_X}\exp(-u^2)=\frac{1}{\sigma_X\sqrt{2\pi}}\exp\left\{-\frac{(x-m_X)^2}{2\sigma_X^2}\right\}.$$

Условную вероятность находим по формуле (6):

$$f_Y(y/x)=\frac{f_{X,Y}(x,y)}{f_X(x)}=\frac{1}{\sigma_Y\sqrt{2\pi}\sqrt{1-\rho_{XY}^2}}\exp\left\{-\frac{1}{2(1-\rho_{XY}^2)\sigma_Y^2}\times\right. \\ \left.\times\left[y-m_y-\rho_{XY}\frac{\sigma_Y}{\sigma_X}(x-m_X)\right]^2\right\}.$$

Это значит, что $f_Y(y/x)$ представляет собой гауссовскую плотность с параметрами, имеющими смысл условного математического ожидания и условной дисперсии:

$$\begin{aligned} \mathrm{M}[Y/X=x]&=m_Y+\rho_{XY}\frac{\sigma_Y}{\sigma_X}(x-m_X),\\ \mathrm{D}[Y/X=x]&=\sigma_Y^2(1-\rho_{XY}^2). \end{aligned}\tag{12}$$

Уравнение (12), определяющее условное математическое ожидание как функцию x, называется уравнением (линейной) регрессии Y на X. ▷

18.423. Случайная точка на плоскости (X, Y) распределена по каноническому нормальному закону с центром рассеивания $(m_X, m_Y)=(0, 1)$ и среднеквадратичными отклонениями $\sigma_X = 1$, $\sigma_Y=2$. Вычислить вероятность попадания случайной точки в прямоугольник с вершинами $(-1, 1)$, $(2, 1)$, $(2, 3)$ и $(-1, 3)$.

18.424. Случайная точка на плоскости (X, Y) распределена по круговому нормальному закону ($\rho_{XY}=0$, $\sigma_X=\sigma_Y=\sigma=1$)

с центром рассеивания в начале координат. Вычислить вероятности следующих событий: $A = \{Y > X\}$, $B = \{|Y| > X\}$, $C = \{Y < 3X\}$.

18.425 (продолжение). В условиях предыдущей задачи вычислить вероятность событий $D = \{|X| < 1\}$ и $E = \{X < 1,\ Y < 2\}$.

18.426. Координаты точек на плоскости независимы и распределены по законам $N(a,\ \sigma)$ и $N(b,\ \sigma)$.

Найти радиус круга с центром в точке $(a,\ b)$, вероятность попадания в который равна 0,997.

18.427. Случайная точка $(X,\ Y)$ распределена по круговому нормальному закону с параметрами $m_X = m_Y = 0$, $\sigma_X = \sigma_Y = 2$. Найти вероятности попадания случайной точки в области G_i, $i = 1,\ 2$, где $G_1 = \left\{(r,\ \varphi)\,|\,0 < r \leqslant 2,\ 0 < \varphi \leqslant \dfrac{\pi}{6}\right\}$ (сектор), $(r,\ \varphi)$ — полярные координаты точки, $G_2\{(x,\ y)\,|\,4 \leqslant x^2 + y^2 \leqslant 9\}$ (кольцо).

18.428 (продолжение). В условиях предыдущей задачи найти вероятности попадания случайной точки в область $G_3 = \{(x,\ y)\,\big|\,|x| + |y| \leqslant 3\}$ (квадрат).

18.429*. Случайная точка на плоскости $(X,\ Y)$ распределена по круговому нормальному закону с параметрами $m_X = m_Y = 0$, $\sigma_X = \sigma_Y = 1$. Найти вероятности попадания случайной точки в области G_i, $i = 1,\ 2$, где G_1 — треугольник с вершинами (0, 0), (1, 0), (1, 1), G_2 — треугольник с вершинами (0, 0), (1, 1), (2, 0).

18.430 (продолжение). В условиях предыдущей задачи найти вероятности попадания случайной точки в области G_3 — трапеция с вершинами (1, 0), (2, 0), (2, 2), (1, 1), G_4 — треугольник с вершинами (2, 0), (2, 2), (1, 1).

18.431. Заданы следующие характеристики двумерного нормального вектора: $m_X = -2$, $m_Y = 3$ и ковариационная матрица

$$K = \begin{pmatrix} 16 & 12 \\ 12 & 25 \end{pmatrix}.$$

Записать выражение для плотности распределения вероятностей $f_{X,Y}(x,\ y)$ и вычислить вероятность попадания в главный эллипс рассеивания с полуосями $a = 2\sigma_X$, $b = 2\sigma_Y$.

18.432 (продолжение). В условиях предыдущей задачи написать уравнение регрессии Y на X.

18.433. Случайный вектор $(X,\ Y)$ подчиняется нормальному распределению с параметрами $m_X = -1$, $m_Y = 1$, $\sigma_X = 1$, $\sigma_Y = 2$, $\rho_{XY} = 0$. Написать уравнение главного эллипса, ограничивающего область G, вероятность попадания случайной точки $(X,\ Y)$ в которую равна 0,9.

18.434. Закон распределения двумерного случайного вектора описывается плотностью распределения вероятностей следующего вида:

$$f_{X,Y}(x, y) = \frac{1}{\pi} e^{-(x^2+2xy+5y^2)/2}.$$

Записать выражение для безусловной плотности $f_X(x)$ и указать значения основных параметров совместного распределения.

18.435. Производится стрельба по точечной (малоразмерной) цели, зона поражения которой представляет собой круг радиуса r с центром в начале координат. Рассеивание точки попадания снаряда — нормальное круговое с параметрами $m_X = m_Y = 0$, $\sigma_X = \sigma_Y = 2r$. Сколько выстрелов надо произвести, чтобы поразить цель с вероятностью, не меньшей 0,95?

§ 4. Функции случайных величин

1. Числовые характеристики функций случайных величин. Если X — дискретная или непрерывная случайная величина с известным законом распределения и $Y = \varphi(X)$, где φ — неслучайная функция, то математическое ожидание и дисперсия случайной величины Y в случае, если они существуют, могут быть найдены по формулам

$$m_Y = \mathrm{M}[Y] = \begin{cases} \sum\limits_k \varphi(x_k)\,\mathrm{P}\{X = x_k\}, & \text{если} \quad \text{X – С.В.Д.Т.}, \\ \int\limits_{-\infty}^{+\infty} \varphi(x)\, f_X(x)\, dx, & \text{если} \quad \text{X – С.В.Н.Т.}; \end{cases} \tag{1}$$

$$D_Y = \mathrm{M}[\mathring{Y}^2] = \begin{cases} \sum\limits_k (\varphi(x_k) - m_Y)^2\,\mathrm{P}\{X = x_k\}, & \text{если} \quad \text{X – С.В.Д.Т.}, \\ \int\limits_{-\infty}^{+\infty} (\varphi(x) - m_Y)^2 f_X(x)\, dx, & \text{если} \quad \text{X – С.В.Н.Т.} \end{cases}$$

Аналогичные формулы имеют место и для всех прочих начальных и центральных моментов распределения случайной величины Y, которая является неслучайной функцией X. Таким образом, для вычисления числовых характеристик неслучайной функции случайной величины не надо знать закона распределения зависящей от X случайной величины Y, а достаточно знать закон распределения случайного аргумента X. Сформулированное правило естественно обобщается на функции от большего числа случайных переменных. Например, если

$Z = \varphi(X, Y)$, то

$$m_z = \mathrm{M}[Z] = \begin{cases} \sum\limits_i \sum\limits_j \varphi(x_i, y_j)\, \mathrm{P}\{X = x_i,\ Y = y_j\}, \\ \qquad\qquad\qquad\qquad \text{если } (\mathrm{X}, \mathrm{Y}) - \text{С. В. Д. Т.}, \\ \iint\limits_{-\infty}^{+\infty} \varphi(x, y) f_{XY}(x, y)\, dx\, dy, \\ \qquad\qquad\qquad\qquad \text{если } (\mathrm{X}, \mathrm{Y}) - \text{С. В. Н. Т.} \end{cases}$$

Если существуют соответствующие моменты, то справедливы следующие свойства математического ожидания и дисперсии:

1. *Для любых случайных величин* X_k $(k = 1, 2, \ldots, n)$

$$\mathrm{M}\left[\sum_{k=1}^{n} a_k X_k + b\right] = \sum_{k=1}^{n} a_k \mathrm{M}[X_k] + b \quad (\textit{свойство линейности}).$$

2. *Если* $X \geqslant 0$, *то* $\mathrm{M}[X] \geqslant 0$.
3. *Если* $X \geqslant 0$ *и* $\mathrm{M}[X] = 0$, *то* $\mathrm{P}\{X = 0\} = 1$.
4. *Для любых случайных величин* X_k $(k = 1, 2, \ldots, n)$

$$D\left[\sum_{k=1}^{n} a_k X_k + b\right] = \sum_{k=1}^{n} a_k^2 D[X_k] + 2 \sum_{\substack{i,j=1 \\ i<j}}^{n} a_i a_j K_{ij},$$

где $K_{ij} = M[\mathring{X}_i \mathring{Y}_j]$. В частности,

$$\mathrm{D}[aX + bY + c] = a^2 \mathrm{D}[X] + b^2 \mathrm{D}[Y] + 2abK_{XY}.$$

5. $\mathrm{M}[XY] = \mathrm{M}[X]\mathrm{M}[Y] + K_{XY}$.
6. $\mathrm{M}^2[XY] \leqslant \mathrm{M}[X^2]\mathrm{M}[Y^2]$ (*неравенство Коши–Буняковского*).
7. *Если* X *и* Y *независимы, то*

$$\mathrm{D}[XY] = \mathrm{D}[X]\mathrm{D}[Y] + m_X^2 \mathrm{D}[Y] + m_Y^2 \mathrm{D}[X].$$

Свойство 1 может быть записано в более общей форме в матричных обозначениях:

1*. $\mathrm{M}[A\boldsymbol{X}B + C] = A\mathrm{M}[\boldsymbol{X}]B + C$,

где $\boldsymbol{X}$ — случайный n-мерный вектор-столбец, $\mathrm{M}[\boldsymbol{X}]$ — неслучайный n-мерный вектор-столбец, компоненты которого равны математическим ожиданиям случайных компонент вектора $\boldsymbol{X}$, A, B и C — постоянные матрицы порядков соответственно $m \times n$, $n \times p$ и $m \times p$. (Свойство 1 является частным случаем свойства 1* при $m = p = 1$.)

Пример 1. На вход измерительного прибора поступает случайный вектор (X, Y) со следующими характеристиками: $m_X = -1$, $m_Y = 1$, $\sigma_X = 2$, $\sigma_Y = 3$, $\rho_{XY} = 0{,}5$. На выходе прибора измеряется величина $Z = (X - Y)^2$. Определить математическое ожидание случайной величины Z.

◁ Воспользуемся свойствами 1 и 5 математического ожидания:

$$\mathrm{M}[Z] = \mathrm{M}[X^2 + Y^2 - 2XY] = \mathrm{M}[X^2] + \mathrm{M}[Y^2] - 2\mathrm{M}[XY] =$$
$$= D_X + m_X^2 + D_Y + m_Y^2 - 2(m_X m_Y + K_{XY}) = 11. \ ▷$$

Пример 2. На окружность радиуса r наудачу ставятся две точки, которые затем соединяются между собой и с центром окружности. Найти математическое ожидание площади полученного треугольника.

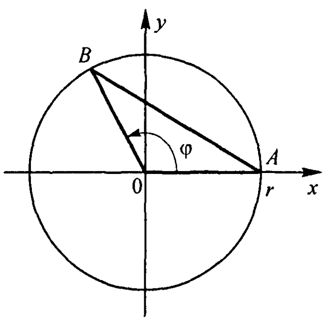


Рис. 11

◁ Так как в данном опыте важно лишь взаимное расположение точек на окружности, то можно считать, что первая точка имеет фиксированные координаты $(r, 0)$. Тогда положение второй точки, случайно поставленной на окружности, полностью определяется случайным углом Φ между положительным направлением оси Ox и радиус-вектором, проведенным во вторую точку, как показано на рис. 11. Поскольку все значения угла Φ равновозможны в пределах от 0 до 2π, то можно считать, что случайная величина Φ распределена по закону $R(0, 2\pi)$. Поэтому

$$f_\Phi(\varphi) = \begin{cases} 0, & \text{если} \quad \varphi \notin (0, 2\pi), \\ \dfrac{1}{2\pi}, & \text{если} \quad \varphi \in (0, 2\pi). \end{cases}$$

При фиксированных точках A и B площадь треугольника OAB записывается в виде

$$S = \frac{r^2}{2}|\sin\Phi|.$$

По формуле (1) для С. В. Н. Т. имеем

$$\mathrm{M}[S] = \frac{1}{2}r^2 \int_0^{2\pi} |\sin\varphi| \frac{1}{2\pi}\, d\varphi = \frac{r^2}{2\pi}\int_0^{\pi} \sin\varphi\, d\varphi = \frac{r^2}{\pi}. \ ▷$$

18.436. Один раз брошены две игральные кости. Случайная величина S — сумма выпавших очков. Вычислить среднее значение и дисперсию случайной величины S.

18.437. Случайные величины X и Y независимы и имеют следующие характеристики: $m_X = 1$, $m_Y = 2$, $\sigma_X = 1$, $\sigma_Y = 2$. Вычислить математические ожидания случайных величин $U = X^2 + 2Y^2 - XY - 4X + Y + 4$, $V = (X + Y - 1)^2$.

18.438. Случайная точка (X, Y) характеризуется центром рассеивания $(-1, 1)$ и ковариационной матрицей $\begin{pmatrix} 3 & -2 \\ -2 & 4 \end{pmatrix}$. Найти математическое ожидание и дисперсию случайной величины $Z = 2X - 4Y + 3$.

18.439. В прямоугольник с вершинами $(0, 0)$, $(2, 0)$, $(2, 1)$ и $(0, 1)$ наудачу ставится точка. Обозначим (X, Y) случайные координаты этой точки. Вычислить $\mathrm{M}[X \pm Y]$, $\mathrm{M}[X^2 + Y^2]$ и $\mathrm{M}[XY]$.

18.440 (продолжение). В условиях предыдущей задачи вычислить $\mathrm{D}[X \pm Y]$ и $\mathrm{D}[XY]$.

18.441. Случайные величины X и Y независимы и имеют одинаковое распределение с математическим ожиданием m и дисперсией σ^2. Найти коэффициент корреляции случайных величин $U = \alpha X + \beta Y$ и $V = \alpha X - \beta Y$.

18.442 (продолжение). Показать, что случайные величины $Z = X + Y$ и $W = X - Y$, где X и Y описаны в условии предыдущей задачи, некоррелированы.

18.443. Случайная величина X дискретного типа распределена по закону, определяемому таблицей

x_i	-1	0	1
p_i	$1/6$	$1/3$	$1/2$

Найти коэффициент корреляции между X и X^2.

18.444. Случайные величины X и Y независимы и распределены: X по закону $R(0, 2)$, Y — по закону $N(1, 2)$. Вычислить $\mathrm{D}[X - Y]$ и $\mathrm{M}[XY^2 + X^2Y]$.

18.445. Случайная величина X распределена по закону $N(-1, 2)$, а не зависимая от нее случайная величина Y распределена по закону $R(-1, 3)$. Вычислить $\mathrm{M}[Z]$ и $\mathrm{D}[Z]$, где $Z = X + Y - XY$.

18.446*. Случайная величина X распределена по закону $N(1, 1)$. Вычислить ковариационную матрицу случайного вектора $Y = (X, X^2, X^3)$.

18.447*. Случайная величина Φ — угол поворота азимутального лимба прибора — может принимать конечный набор значений $\varphi_k = k \cdot \dfrac{2\pi}{n}$ $(k = 0, 1, 2, \ldots, n)$ с равными вероятностями. На выходе прибора измеряется случайная величина $Y = \sin \Phi$. Найти $\mathrm{M}[Y]$ и $\mathrm{D}[Y]$.

18.448. Случайная величина Φ принимает дискретные значения $\varphi_k = \dfrac{k\pi}{2}$ $(k = 0, 1, \ldots)$ с вероятностями, убывающими в геометрической прогрессии. Найти $\mathrm{M}[Y]$ и $\mathrm{D}[Y]$, если $Y = \sin \Phi$.

18.449 (продолжение). В условиях предыдущей задачи найти $\mathrm{M}[Z]$ и $\mathrm{D}[Z]$, если $Z = \cos \Phi$.

18.450. На окружность радиуса r наудачу ставятся две точки. Найти математическое ожидание и дисперсию случайной длины L хорды, соединяющей эти точки.

18.451. На отрезок AB длины l наудачу ставится точка M и проводится окружность радиуса AM. Найти $\mathrm{M}[L]$ и $\mathrm{D}[L]$, где L — длина окружности.

18.452 (продолжение). В условиях предыдущей задачи найти $\mathrm{M}[S]$ и $\mathrm{D}[S]$, где S — площадь круга.

18.453. На отрезок $[0, l]$ наудачу ставятся две точки A и B. Найти $\mathrm{M}[S]$, где S — площадь квадрата со стороной $R = |\overline{AB}|$.

18.454 (продолжение). В условиях предыдущей задачи найти $\mathrm{M}[R]$ и $\mathrm{D}[R]$.

18.455. На экране индикатора радиолокационной станции кругового обзора отраженный импульс от цели представляется в виде светящейся точки с координатами (x, y). При поиске очередной цели светящаяся точка появляется случайным образом в любом месте круга радиуса a. Найти математическое ожидание и дисперсию расстояния R от точки до центра экрана.

18.456. На смежные стороны прямоугольника со сторонами a и b $(a > b)$ наудачу и независимо ставится по одной точке. Найти математическое ожидание и дисперсию квадрата расстояния R^2 между ними.

18.457. Шарики для подшипников изготовляются из стали с плотностью $\rho = 7{,}8\,\text{г/см}^3$. Они считаются годными, если проходят через отверстие диаметром $d_2 = 10{,}1$ мм и не проходят через отверстие диаметром $d_1 = 9{,}9$ мм. Технология изготовления шариков такова, что диаметр шарика можно считать нормально распределенной случайной величиной с параметрами $m_X = \dfrac{d_1 + d_2}{2}$, $\sigma_X = \dfrac{d_2 - d_1}{5}$. Найти математическое ожидание массы шарика Z с тремя верными знаками после запятой.

18.458 (продолжение). В условиях предыдущей задачи допустим, что диаметр шариков распределен равномерно в пределах поля допуска (по закону $R(d_1, d_2)$) с теми же значениями d_1, d_2 и ρ. Найти математическое ожидание массы шарика Z с тремя верными знаками после запятой.

18.459. Проводятся последовательные испытания по схеме Бернулли с вероятностью успеха в одном испытании, равной p. Случайная величина Z — относительная частота успехов в n испытаниях. Вычислить $\mathrm{M}[Z]$ и $\mathrm{D}[Z]$, выразив Z через сумму индикаторов успеха в одном испытании и воспользовавшись свойствами операций математического ожидания и дисперсии.

18.460*. Автоматическая линия производит детали, удовлетворяющие стандарту с вероятностью q и не удовлетворяющие стандарту с вероятностью $p = 1 - q$. Перестройка линии производится после получения k нестандартных деталей. Случайная величина Z — число деталей, сошедших с автоматической линии между двумя перестройками. Вычислить m_Z и D_Z.

18.461*. Проводятся последовательные независимые испытания по схеме Бернулли с вероятностью успеха в одном испытании, равной p. Назовем «серией успехов» такую конечную цепочку последовательных испытаний, в которой каждое испытание закончилось успехом, а испытание, предшествующее цепочке, и испытание, следующее за ней, — неудачей. Пусть Z_n — число серий успехов в n испытаниях. Найти $\mathrm{M}[Z_n]$.

18.462*. Случайная величина Z представима в виде

$$Z = \sum_{k=1}^{Y} X_k,$$

где X_k $(k = 1, 2, \ldots)$ — попарно независимые, одинаково распределенные случайные величины с математическим ожиданием m и дисперсией σ^2, а Y — С. В. Д. Т., множество значений которой есть $\{1, 2, \ldots, N\}$ (N конечно или бесконечно) с известными вероятностями их реализаций $p_k = \mathrm{P}\{Y = k\}$, $k = 1, 2, \ldots$ Показать, что справедливы следующие формулы для моментов $\mathrm{M}[Z]$ и $\mathrm{M}[Z^2]$:

$$\mathrm{M}[Z] = m\sum_{n=1}^{N} np_n, \quad \mathrm{M}[Z^2] = \sum_{n=1}^{N} p_n(n\sigma^2 + n^2m^2).$$

18.463*. Бросается игральная кость. Пусть выпало k очков. После этого та же кость подбрасывается k раз. Случайная величина Z — сумма выпавших при этом очков. Найти $\mathrm{M}[Z]$ и $\mathrm{D}[Z]$.

18.464*. Пусть $Z = \sum\limits_{k=1}^{Y} X_k$, где X_k, $k = 1, 2, \ldots$, подчиняются закону $N(m, \sigma)$ и независимы в совокупности, Y распределена по геометрическому закону с параметром p. Вычислить m_Z и D_Z.

18.465. Пусть для случайной величины X существует начальный момент 4-го порядка, т.е. $\mathrm{M}[X^4]$. Используя неравенство Коши–Буняковского, доказать, что тогда существуют и начальные моменты 1-го, 2-го и 3-го порядков.

18.466. Показать, что коэффициенты асимметрии и эксцесса не зависят от начала отсчета и масштаба при линейном преобразовании, т.е. если $Y = \lambda(X - a)$, то $a_Y = a_X$, $e_Y = e_X$.

18.467*. Пусть $X_1, X_2, \ldots, X_n$ — стандартизованные случайные величины, одинаково коррелированные между собой (т.е. $\rho_{ij} = \rho$ для всех $i \neq j$). Показать, что $\rho \geqslant -\dfrac{1}{n-1}$.

18.468*. Пусть X и Y — две стандартизованные случайные величины с коэффициентом корреляции $\rho_{XY} = \rho$. Доказать, что

$$1 \leqslant \mathrm{M}[\max(X^2, Y^2)] \leqslant 1 + \sqrt{1-\rho^2}.$$

18.469*. Опыт состоит в том, что точка X_1 наудачу выбирается из отрезка $(0, 1)$, затем точка X_2 наудачу выбирается из отрезка $(X_1, 1)$, ..., точка X_n наудачу выбирается из отрезка $(X_{n-1}, 1)$. Найти $\mathrm{M}[X_n]$.

18.470. Два равносильных шахматиста договорились сыграть между собой матч на следующих условиях. Общее число партий не ограничивается, за каждую выигранную партию победитель получает одно очко, за ничьи очки не присуждаются. Выигравшим матч считается тот, кто первый наберет 6 очков. Определить среднее число партий, сыгранных в данном матче, если вероятность выиграть очередную партию любому из игроков равна p $(p < 1/2)$, а вероятность ничейного исхода равна $1 - 2p$.

$\lhd$ Обозначим X число партий, сыгранных в матче, X_k $(k = 1, 2, \ldots, 11)$ — число партий, сыгранных от $(k-1)$-й победы кого-либо из участников до k-й победы кого-либо из участников (последнее — включительно), Z_m $(m = 1, 2, \ldots, 11)$ — число партий, сыгранных до достижения m-й результативной партии включительно. Очевидно, $Z_m = \sum\limits_{k=1}^{m} X_k$. Так как вероятность результативной партии равна $2p$, то X_k (при любом k) подчиняется геометрическому распределению (см. задачу 18.330) с параметром $2p$, поэтому $\mathrm{M}[X_k] = \dfrac{1}{2p}$. Согласно свойству 1) математического ожидания $\mathrm{M}[Z_m] = \sum\limits_{k=1}^{m} \mathrm{M}[X_k] = \dfrac{m}{2p}$. Обозначим Y число результативных партий, сыгранных в матче. Множество возможных значений случайной величины Y: $\{6, 7, \ldots, 11\}$. Очевидно, событие $\{Y = 6 + k\}$ $(k = 0, 1, \ldots, 5)$ означает, что один из участников выиграл матч, набрав 6 очков при общем числе результативных партий $6 + k$. Таким образом, $\mathrm{P}\{Y = 6 + k\} = \mathrm{P}\{$матч выиграл первый участник при общем числе результативных партий $6 + k\} + \mathrm{P}\{$матч выиграл второй участник при общем числе результативных партий $6+k\} = 2\mathrm{P}\{$составить наудачу $(6 + k)$-буквенное слово из алфавита с двумя буквами $\{\Pi, B\}$ (Π — выиграл первый, B — выиграл второй), причем слово должно содержать 6 букв Π и оканчиваться буквой $\Pi\} = \dfrac{C_{5+k}^{k}}{2^{5+k}}$ $(k = 0, 1, \ldots, 5)$. Найдем условное математическое ожидание

$$\mathrm{M}[X/Y = 6 + k] = \mathrm{M}[Z_{6+k}] = \frac{6 + k}{2p}.$$

По формуле (8) § 3 полного математического ожидания

$$\mathrm{M}[X] = \mathrm{M}[\mathrm{M}[X/Y]] = \\ = \sum_{k=0}^{5} \mathrm{P}\{Y = 6 + k\}\,\mathrm{M}[X/Y = 6 + k] = \sum_{k=0}^{5} \frac{C_{5+k}^{k}}{2^{k+5}}\,\frac{6 + k}{2p} = \frac{4{,}6465}{p}.$$

Поскольку возможные значения X — натуральные числа, то в качестве $\mathrm{M}[X]$ нужно взять ближайшее натуральное число, не меньшее чем $\dfrac{4{,}6465}{p}$. Например, при $p = \dfrac{1}{4}$ имеем $\mathrm{M}[X] = 19$. ▷

18.471*. В n почтовых ящиков, установленных в данном районе города, случайно и независимо опускают по одному письму в течение длительного времени.

а) Найти математическое ожидание $\mathrm{M}[X_n]$ общего числа писем, опущенных до момента, пока не останется пустых ящиков.

б) Получить числовые значения при $n = 2;\ 5;\ 10;\ 100$.

2. Характеристические функции случайных величин. Если $Z = X + iY$ — комплекснозначная случайная величина, где X и Y — действительные случайные величины, то $\mathrm{M}[Z] = \mathrm{M}[X] + i\,\mathrm{M}[Y]$.

Характеристической функцией $E_X(t)$ случайной величины X называется комплекснозначная неслучайная функция действительного аргумента t, определяемая равенством

$$E_X(t) = \mathrm{M}\left[e^{itX}\right] = \begin{cases} \sum\limits_k e^{itx_k}\, \mathrm{P}\{X = x_k\}, & \text{если} \quad X - \text{С. В. Д. Т.}, \\ \int\limits_{-\infty}^{+\infty} e^{itx} f_X(x)\, dx, & \text{если} \quad X - \text{С. В. Н. Т.} \end{cases}$$

Свойства характеристической функции:

1. $E_X(0) = 1$, $|E_X(t)| \leqslant 1$.

2. Если $E_X(t)$ — характеристическая функция случайной величины X и $Y = aX + b$, то
$$E_Y(t) = e^{itb} E_X(at).$$

3. Если существует m-й абсолютный момент $\mathrm{M}\left[|X|^m\right]$, то существуют производные характеристической функции $E_X(t)$ до m-го порядка включительно, причем
$$\left.\frac{d^k E_X(t)}{dt^k}\right|_{t=0} = i^k \alpha_k, \quad \text{где} \quad \alpha_k = \mathrm{M}\left[X^k\right], \quad k = 1, 2, \ldots, m.$$

4. Если $Y = \sum\limits_{k=1}^{n} X_k$, причем $\{X_k\}$ $(k = 1, 2, \ldots, n)$ независимы в совокупности, то $E_Y(t) = E_{X_1}(t)\, E_{X_2}(t) \ldots E_{X_n}(t)$.

5. $\overline{E_X(t)} = E_X(-t) = E_{-X}(t)$, где черта означает операцию комплексного сопряжения. В частности, отсюда следует, что если $E_X(t)$ — действительная функция, то она обязательно четная.

6. По характеристической функции $E_X(t)$ однозначно восстанавливается функция распределения $F_X(x)$. Если же X — С. В. Н. Т. (т.е. для нее существует плотность) и функция $E_X(t)$ абсолютно интегрируема, то
$$f_X(x) = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} e^{-itx} E_X(t)\, dt.$$

Характеристической функцией случайного вектора $X_1, X_2, \ldots, X_n$ называется комплекснозначная неслучайная функция n действительных переменных $t_1, t_2, \ldots, t_n$, определяемая равенством
$$E_{X_1, X_2, \ldots, X_n}(t_1, t_2, \ldots, t_n) = \mathrm{M}\left[\exp\left\{i \sum_{k=1}^{n} t_k X_k\right\}\right].$$

Пример 3. Случайная величина X подчиняется закону распределения $N(0, 1)$. Найти ее характеристическую функцию.

◁ По определению характеристической функции

$$E_X(t) = \mathrm{M}\left[\exp\{itX\}\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\{itx\} \exp\left\{-\frac{x^2}{2}\right\} dx =$$

$$= \frac{\exp\left\{\dfrac{t^2}{2}\right\}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left\{-\frac{(x-it)^2}{2}\right\} dx = \frac{\exp\left\{-\dfrac{t^2}{2}\right\}}{\sqrt{2\pi}} \int_{-\infty-it}^{+\infty-it} \exp\left\{-\frac{z^2}{2}\right\} dz =$$

$$= \frac{\exp\left\{-\dfrac{t^2}{2}\right\}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left\{-\frac{x^2}{2}\right\} dx = \exp\left\{-\frac{t^2}{2}\right\}.$$

Переход от интегрирования по контуру $L = (-\infty - it,\ +\infty - it)$ к интегрированию по вещественной оси оправдывается аналитичностью подынтегральной функции в части нижней полуплоскости, ограниченной действительной осью и прямой L, и возможностью деформировать контур интегрирования в области аналитичности согласно теореме Коши. ▷

Пример 4. Указать, какие из нижеприведенных функций действительной переменной t не являются характеристическими функциями и почему.

$$E_1(t) = \frac{1}{1+t}, \quad E_2(t) = \frac{1}{1+t^2}, \quad E_3(t) = \sin bt,$$

$$E_4(t) = \cos bt, \quad E_5(t) = 1 - it.$$

◁ Не являются характеристическими функциями $E_1(t)$, $E_3(t)$ (не выполняется свойство 5) и $E_5(t)$ (не выполняется свойство 1). ▷

Пример 5. Найти характеристическую функцию случайной величины X, распределенной по закону Пуассона с параметром λ, и с ее помощью вычислить m_X и D_X.

◁ По определению характеристической функции

$$E_X(t) = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}.$$

По свойству 3 имеем

$$i\alpha_1 = E_X'(t)\big|_{t=0} = e^{\lambda(e^{it}-1)} \lambda i e^{it}\big|_{t=0} = \lambda i,$$

следовательно,

$$m_X = \alpha_1 = \lambda.$$

Далее, по тому же свойству

$$i^2\alpha_2 = E_X''(t)\big|_{t=0} = i\lambda\,(1+\lambda i e^{it})\,e^{it+\lambda\,(e^{it}-1)}\big|_{t=0} = i^2\lambda\,(1+\lambda),$$

т.е. $\alpha_2 = \lambda\,(1+\lambda)$.

Таким образом,

$$D_X = \alpha_2 - m_X^2 = \lambda + \lambda^2 - \lambda^2 = \lambda. \ \triangleright$$

18.472. Случайная величина X дискретного типа может принимать только два возможных значения: -1 или 1, с равными вероятностями. Вычислить характеристическую функцию данного распределения.

18.473. Случайная величина X дискретного типа распределена по закону, определяемому таблицей

x_k	-2	0	2
p_k	$1/4$	$1/2$	$1/4$

Найти характеристическую функцию и с ее помощью вычислить дисперсию σ_X^2.

18.474. Проводятся последовательные независимые испытания с двумя исходами. Случайные величины: I_k — индикатор успеха в k-м испытании, X — число успехов в n испытаниях. Построить характеристическую функцию для I_k и, используя ее свойства, найти характеристическую функцию случайной величины X, если вероятность успеха от испытания к испытанию не меняется и равна p.

18.475 (продолжение). Найти характеристическую функцию случайной величины X из предыдущей задачи, если вероятность успеха в k-м испытании равна p_k.

18.476. Случайная величина X распределена по биномиальному закону с параметрами n и p. Найти ее характеристическую функцию.

18.477 (продолжение). В условиях предыдущей задачи найти m_X и D_X, используя характеристическую функцию.

18.478. Случайная величина X распределена по геометрическому закону с параметром $p > 0$ (см. задачу 18.330). Найти характеристическую функцию и с ее помощью вычислить математическое ожидание m_X.

18.479. Случайные величины X и Y независимы и одинаково распределены с характеристической функцией $E\,(t)$. Пусть $Z = X - Y$. Найти $E_Z\,(t)$.

18.480. Пусть для случайной величины X непрерывного типа существует $\mathrm{M}[|X|]$. Показать, что

$$\left|\frac{dE_X(t)}{dt}\right| \leqslant \mathrm{M}[|X|].$$

18.481. Пусть X— случайная величина непрерывного типа с вещественной характеристической функцией. Показать, что

$$E_X(t) = \int\limits_{-\infty}^{+\infty} \cos tx f_X(x)\,dx.$$

18.482 (продолжение). В условиях предыдущей задачи показать, что $m_X = 0$.

18.483* (продолжение). В условиях задачи 18.481 известно, что дисперсия случайной величины существует и равна σ^2. Показать, что $E_X(t) \geqslant 1 - \dfrac{t^2\sigma^2}{2}$.

18.484. Случайная величина X подчиняется показательному распределению с параметром λ. Найти ее характеристическую функцию.

18.485 (продолжение). Используя найденную в предыдущей задаче характеристическую функцию показательного распределения, вычислить основные его характеристики: m_X, D_X и a_X.

18.486. Случайная величина X подчиняется равномерному закону распределения на отрезке $[a, b]$ (закону $R(a, b)$). Найти характеристическую функцию.

18.487*. Случайная величина X подчиняется *закону Коши с параметрами* $c \in \mathbb{R}$ *и* $a > 0$ с плотностью распределения вероятностей

$$f_X(x) = \frac{a}{\pi}\,\frac{1}{(x-c)^2 + a^2}.$$

Найти ее характеристическую функцию.

Семейство законов распределения, описываемых функциями распределения $F\left(\dfrac{x-a}{b}\right)$, где $F(x)$ — фиксированная функция распределения, $a \in \mathbb{R}$, $b > 0$, называется *видом распределения.* При этом параметр a называется *параметром сдвига*, b — *масштабным множителем.*

Из этого определения вытекают два простых следствия:

Следствие 1. *Семейство законов распределения, описываемых плотностями* $\dfrac{1}{b} f\left(\dfrac{x-a}{b}\right)$, *где* $f(x)$ — *фиксированная плотность распределения вероятностей,* $a \in \mathbb{R}$, $b > 0$, *является видом распределения.*

Следствие 2. *Семейство законов распределения, описываемых характеристическими функциями* $e^{ita}E(bt)$, *где* $a \in \mathbb{R}$, $b > 0$, $E(t)$ — *фиксированная характеристическая функция, является видом распределения.*

Пример 6. Характеристическая функция случайной величины X имеет вид $E_X(t) = \dfrac{e^{it} - 1}{it}$. Какому закону распределения она соответствует?

◁ Характеристическая функция случайной величины X, распределенной по закону $R(a, b)$, имеет вид (см. задачу 18.486)

$$E_X(t) = \frac{e^{itb} - e^{ita}}{it(b-a)}.$$

Сравнивая полученную характеристическую функцию с заданной, находим $b = 1$, $a = 0$. По следствию 2 заключаем отсюда, что заданная характеристическая функция соответствует закону распределения $R(0, 1)$. ▷

18.488. Пусть X — случайная величина непрерывного типа с функцией распределения $F_X(x)$. Найти характеристическую функцию случайной величины $Y = F_X(X)$ и указать, какому распределению она соответствует.

18.489. Пусть X — случайная величина непрерывного типа с функцией распределения $F_X(x)$. Найти характеристическую функцию случайной величины $Y = -\ln F_X(X)$ и указать, какому распределению она соответствует.

18.490. Случайная величина X распределена по закону $N(0, 1)$. Найти характеристическую функцию случайной величины $Y = \sigma X + m$ и установить вид закона распределения.

18.491. Случайная величина $Z = \dfrac{X_1 + X_2 + \ldots + X_n}{n}$, где $\{X_k\}$, $k = 1, 2, \ldots, n$, независимы в совокупности и одинаково распределены по закону Коши с параметрами $c = 0$ и $a = 1$. Используя аппарат характеристических функций, установить вид закона распределения случайной величины Z.

18.492. Дана характеристическая функция непрерывной случайной величины X:

$$E_X(t) = \frac{1}{1+t^2}.$$

Найти выражение для плотности распределения вероятностей и установить вид закона распределения.

18.493. Пусть X подчиняется закону $N(0, 1)$. Найти характеристическую функцию случайной величины $Y = X^2$ и показать, что она распределена по закону с одной степенью свободы.

18.494*. Найти характеристическую функцию случайной величины $Z_n = X_1^2 + X_2^2 + \ldots + X_n^2$, где все X_k подчиняются закону $\mathcal{N}(0, 1)$ и независимы в совокупности, и с ее помощью установить, что Z_n подчиняется закону $\chi^2(n)$.

18.495 (продолжение). Используя найденную в предыдущей задаче характеристическую функцию, вычислить основные характеристики распределения $\chi^2(n)$ — математическое ожидание и дисперсию.

18.496. Случайная величина $Y = \dfrac{\sigma}{\sqrt{2}} X + m$, где X подчиняется закону распределения Лапласа с параметрами $m_X = 0$ и $\sigma_X = \sqrt{2}$. Найти характеристическую функцию $E_Y(t)$ и с ее помощью вычислить m_Y и D_Y.

18.497. Случайная величина X непрерывного типа имеет характеристическую функцию вида

$$E_X(t) = \begin{cases} 1 - |t|, & |t| \leqslant 1, \\ 0, & |t| > 1. \end{cases}$$

Найти ее плотность распределения вероятностей.

3. Законы распределения функций случайной величины. Если X — С.В.Д.Т. и $Y = \varphi(X)$, где φ — неслучайная функция, то Y — также С.В.Д.Т., причем ее возможные значения $y_k = \varphi(x_k)$. Если при этом все y_k различны (например, функция $\varphi(x)$ строго монотонна), то $\mathrm{P}\{Y = y_k\} = \mathrm{P}\{X = x_k\}$. Если же среди y_k имеются одинаковые значения, то

$$\mathrm{P}\{Y = y_k\} = \sum_{i;\,\varphi(x_i) = y_k} \mathrm{P}\{X = x_i\}$$

(т.е. необходимо сложить вероятности тех значений x_i, для которых $\varphi(x_i) = y_k$).

Если X — С.В.Н.Т. и $Y = \varphi(X)$, причем $\varphi(x)$ — монотонно возрастающая непрерывно дифференцируемая функция, то Y — также С.В.Н.Т., причем

$$F_Y(y) = \mathrm{P}\{Y < y\} = \int_{-\infty}^{\varphi^{-1}(y)} f_X(x)\,dx,$$

$$f_Y(y) = F_Y'(y) = f_X(\varphi^{-1}(y)) \left| \frac{d\varphi^{-1}(y)}{dy} \right|, \tag{2}$$

где $\varphi^{-1}(y)$ — обратная функция к $\varphi(x)$. Если же $\varphi(x)$ — немонотонная функция, то

$$F_Y(y) = \sum_i \int_{\Delta_i(y)} f_X(x)\,dx, \tag{3}$$

где $\Delta_i(y)$ означает i-й интервал на оси Ox, на котором $\varphi(x) < y$. Плотность $f_Y(y)$ получается дифференцированием $F_Y(y)$ по y.

Пример 7. Случайная величина Φ принимает с равной вероятностью любое из конечного числа значений $\varphi_k = k\pi/n$, $k = 0, 1, \ldots, n$. Найти закон распределения случайной величины $Y = \sin\Phi$.

◁ В силу равновероятности возможных значений Φ имеем $\mathrm{P}\{\Phi = \varphi_k\} = \dfrac{1}{n+1}$. По условию $y_k = \sin\dfrac{k\pi}{n}$. Так как значению $y_0 = 0$ соответствуют два возможных значения Φ: $\varphi_0 = 0$ и $\varphi_n = \pi$, то $\mathrm{P}\{Y = y_0\} = \mathrm{P}\{\Phi = 0\} + \mathrm{P}\{\Phi = \pi\} = \dfrac{2}{n+1}$. Аналогично, $\mathrm{P}\left\{Y = y_1 = \sin\dfrac{\pi}{n}\right\} = \mathrm{P}\left\{\Phi = \dfrac{\pi}{n}\right\} + \mathrm{P}\left\{\Phi = \dfrac{n-1}{n}\pi\right\} = \dfrac{2}{n+1}$ и т.д. Таким образом, при нечетном n закон распределения случайной величины Y описывается следующим образом:

$$\mathrm{P}\left\{Y = \sin\frac{k\pi}{n}\right\} = \frac{2}{n+1}, \quad k = 0, 1, \ldots, \frac{n-1}{2}.$$

При четном n

$$\mathrm{P}\left\{Y = \sin\frac{k\pi}{n}\right\} = \begin{cases} \dfrac{2}{n+1}, & \text{если} \quad k = 0, 1, \ldots, \dfrac{n-2}{2}, \\ \dfrac{1}{n+1}, & \text{если} \quad k = \dfrac{n}{2}. \end{cases} \ ▷$$

Пример 8. Случайная величина X подчиняется закону распределения $N(m, \sigma)$. Найти плотность распределения вероятностей случайной величины $Y = (X - m)^3$.

◁ Плотность распределения вероятностей случайной величины X по условию нормальна, т.е.

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}.$$

Функция $y = \varphi(x) = (x - m)^3$ монотонно возрастающая всюду, поэтому справедлива формула (2) для $f_Y(y)$. Обратная функция $\varphi^{-1}(y)$ и ее производная имеют вид

$$\varphi^{-1}(y) = m + \sqrt[3]{y}, \qquad \frac{d\varphi^{-1}(y)}{dy} = \frac{1}{3\sqrt[3]{y^2}}.$$

Подставляя это в формулу для определения $f_Y(y)$, получим

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \frac{1}{3\sqrt[3]{y^2}} \exp\left\{-\frac{\sqrt[3]{y^2}}{2\sigma^2}\right\}, \quad -\infty < y < +\infty. \ ▷$$

Пример 9 (продолжение). Для случайной величины X из предыдущего примера найти плотность распределения случайной величины $Z = (X - m)^2$.

◁ Функция $z = \varphi(x) = (x-m)^2$ имеет два интервала монотонности: $(-\infty, m)$ и (m, ∞), в каждом из которых определена обратная функция $x = m - \sqrt{z}$ при $x < m$ $(z > 0)$ и $x = m + \sqrt{z}$ при $x > m$ $(z > 0)$. Поэтому удобнее сначала найти функцию распределения (см. формулу (3))

$$F_Z(z) = \mathrm{P}\{Z < z\} = \begin{cases} 0, & z \leqslant 0, \\ \displaystyle\int\limits_{x_1(z)}^{m} f_X(x)\,dx + \int\limits_{m}^{x_2(z)} f_X q(x)\,dx, & z > 0, \end{cases}$$

где $x_1(z)$ и $x_2(z)$ — точки пересечения прямой $z = \mathrm{const}$ с ветвями параболы $z = (x-m)^2$, т.е. $x_1(z) = m - \sqrt{z}$, $x_2(z) = m + \sqrt{z}$, следовательно, при $z > 0$

$$F_Z(z) = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{m-\sqrt{z}}^{m} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} dx + \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{m}^{m+\sqrt{z}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} dx.$$

Применяя формулу дифференцирования определенного интеграла по параметру (формулу Лейбница), получим выражение для плотности распределения вероятностей при $z > 0$:

$$f_Z(z) = \frac{dF_{Z(z)}}{dz} = -\frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{z}{2\sigma^2}\right\} \frac{d}{dz}(m - \sqrt{z}) +$$
$$+ \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{z}{2\sigma^2}\right\} \frac{d}{dz}(m + \sqrt{z}).$$

Окончательно получим

$$f_Z(z) = \begin{cases} 0, & \text{если} \quad z \leqslant 0, \\ \dfrac{1}{\sigma\sqrt{2\pi z}} \exp\left\{-\dfrac{z}{2\sigma^2}\right\}, & \text{если} \quad z > 0. \end{cases}$$ ▷

18.498. Случайная величина X принимает значения $x_k = k\pi/8$, $k = 0, 1, \ldots, 8$, с вероятностями $\mathrm{P}\{X = x_k\} = (k+1)/45$, $k = 0, 1, \ldots, 8$. Описать закон распределения случайной величины $Y = \cos 2X$.

18.499. Шесть раз бросается правильная монета. Случайная величина X — модуль разности числа появлений герба и числа появлений цифры в данном эксперименте. Описать закон распределения.

18.500. Один раз брошены две одинаковые игральные кости. Случайная величина X — сумма очков на верхних гранях игральных костей. Описать закон распределения.

18.501*. Доказать, что две непрерывные случайные величины X и Y, связанные между собой линейной зависимостью $Y = aX + b$, подчиняются закону распределения одного и того же вида. В частности, указать закон распределения случайной величины $Y = (X - m)/\sigma$ и определить m_Y и σ_Y, если X подчиняется закону $N(m, \sigma)$.

18.502. Известна функция распределения $F_X(x)$ случайной величины X непрерывного типа. Найти функции распределения случайных величин $Y = 9X^2 - 4$, $Z = |X - 1|$, $V = e^{-2X}$, выразив их через функцию распределения случайной величины X.

18.503 (продолжение). Используя результат предыдущей задачи, найти плотность распределения вероятностей случайной величины $Z = |X - 1|$, если X подчиняется закону $N(1, \sigma)$.

18.504* (продолжение). Пусть $W = F_X(X)$. Вычислить плотность распределения вероятностей случайной величины W и установить вид закона распределения.

18.505. На плоскости Oxy через точку $(a, 0)$ $(a > 0)$ наудачу проводится прямая линия. Найти плотность распределения вероятностей ординаты Y точки пересечения прямой с осью Oy.

18.506. Через точку, наудачу выбранную на окружности радиуса 1 с центром в начале координат, проводится касательная к окружности. Найти плотность распределения вероятностей длины отрезка касательной, заключенного между осями координат.

18.507. Случайная величина X распределена равномерно на интервале $\left(-\frac{l}{2}, \frac{l}{2}\right)$. Найти плотности распределения вероятностей функций $Y_1 = 3X^3$, $Y_2 = a\sin\left(\frac{2\pi}{l}X\right)$, $a > 0$.

18.508. Случайная величина X распределена по показательному закону с параметром λ. Найти плотность распределения вероятностей случайных величин

$$Y = \sqrt{X}, \quad Z = X^2, \quad U = 1 - e^{-\lambda X}.$$

18.509. Случайная величина X подчиняется закону распределения Коши с параметрами $c = 0$, $a = 1$ (см. задачу 18.487). Найти плотности распределения вероятностей следующих случайных величин:

$$Y_1 = 3X - 2, \quad Y_2 = \frac{1}{3X - 2}.$$

18.510 (продолжение). В условиях предыдущей задачи найти плотности распределения вероятностей случайных величин

$$Y_3 = \frac{1}{\pi} \operatorname{arctg} X, \quad Y_4 = 4X^2.$$

18.511. Указать закон распределения случайной величины $Y = \ln\left(\dfrac{X}{x_0}\right)$, если случайная величина X подчиняется закону распределения Парето (см. задачу 18.305) с параметрами $a > 0$ и $x_0 > 0$.

Случайная величина T имеет *распределение Стьюдента с $n \in \mathbb{N}$ степенями свободы* (T распределена по закону $\mathrm{St}\,(n)$), если она непрерывного типа и ее плотность распределения вероятностей имеет вид

$$f_T(t) = \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\dfrac{n}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-(n+1)/2}, \quad -\infty < t < +\infty.$$

Распределение Стьюдента играет важную роль в математической статистике в связи со следующим утверждением, сформулированным в задаче 18.512.

18.512*. Пусть случайная величина X распределена по закону $N\,(0,\ 1)$, а независимая от нее случайная величина Y распределена по закону $\chi^2(n)$. Тогда случайная величина

$$T = X\sqrt{\frac{n}{Y}}$$

распределена по закону $\mathrm{St}\,(n)$. Доказать это.

18.513. Случайная величина T подчиняется закону распределения $\mathrm{St}(n)$. Определить наибольшее натуральное значение m, для которого существует момент $\mathrm{M}\left[|T|^m\right]$.

Если $(X,\ Y)$ — случайный вектор с заданным законом распределения и $Z = \varphi\,(X,\ Y)$, где $\varphi\,(x,\ y)$ — произвольная неслучайная функция, то

$$F_X(z) = \mathrm{P}\,\{Z < z\} = \\ = \begin{cases} \sum\limits_{\substack{i \quad j \\ \varphi\,(x_i,\, y_j) < z}} \sum p_{ij}, & \text{если} \quad (X,\ Y) \ - \ \text{С.В.Д.Т.}, \\ \iint\limits_{\varphi\,(x,\, y) < z} f_{X,Y}\,(x,\ y)\,dx\,dy, & \text{если} \quad (X,\ Y) \ - \ \text{С.В.Н.Т.} \end{cases} \tag{4}$$

Найдя функцию распределения $F_Z(z)$, далее по известным правилам можно найти закон распределения случайной величины Z. В частности, плотность вероятности $f_Z(z)$ для случая непрерывного вектора (X, Y) находится дифференцированием $F_Z(z)$ по z, если в точке z функция $F_Z(z)$ дифференцируема.

Пример 10. Случайный вектор (X, Y) дискретного типа распределен по закону, определяемому таблицей

x_i	y_j		
	0	1	2
−1	0,05	0,06	0,05
0	0,05	0,30	0,15
1	0,09	0,15	0,10

Описать закон распределения случайной величины $Z = |Y| - |X|$.

◁ Для каждой пары возможных значений (x_i, y_j) вычислим соответствующее значение $z(x_i, y_j) = |y_j| - |x_i|$ и результат оформим в виде таблицы.

(x_i, y_j)	(−1,0)	(−1,1)	(−1,2)	(0,0)	(0,1)	(0,2)	(1,0)	(1,1)	(1,2)
$z(x_i, y_j)$	−1	0	1	0	1	2	−1	0	1
$\mathrm{P}\{X = x_i,\ Y = y_j\}$	0,05	0,06	0,05	0,05	0,3	0,15	0,09	0,15	0,1

Из анализа таблицы заключаем, что множеством возможных значений случайной величины Z является множество $\{-1, 0, 1, 2, 3\}$. Вероятности реализации соответствующих значений получаем по правилу сложения вероятностей. Например,

$$\mathrm{P}\{Z=-1\}=\mathrm{P}\{X=-1,\ Y=0\}+\mathrm{P}\{X=1,\ Y=0\}=0{,}05+0{,}09=0{,}18.$$

Окончательный результат оформляем в виде таблицы распределения

z_k	−1	0	1	2
p_k	0,14	0,26	0,45	0,15

▷

Пример 11. Случайный вектор (X, Y) распределен равномерно в круге радиуса a с центром в начале координат. Найти плотность распределения вероятностей случайной величины $Z = Y/X$.

◁ Согласно условию плотность распределения вероятностей случайного вектора (X, Y) имеет вид

$$f_{X,Y}(x, y) = \begin{cases} \dfrac{1}{\pi a^2}, & \text{если} \quad (x, y) \in D_a, \\ 0, & \text{если} \quad (x, y) \notin D_a, \end{cases}$$

где $D_a = \{(x, y) \mid x^2 + y^2 \leqslant a^2\}$.

Найдем сначала функцию распределения случайной величины Z. По определению, учитывая формулу (4), находим

$$F_Z(z) = \mathrm{P}\{Z < z\} = \mathrm{P}\left\{\frac{Y}{X} < z\right\} = \iint\limits_{G_z} f_{X,Y}(x, y)\,dx\,dy, \qquad (5)$$

где $G_z = \{(x, y) \mid y/x < z\}$ — область на плоскости, зависящая от значений действительной переменной z. Для фиксированного значения z эта область показана на рис. 12 и 13 штриховкой.

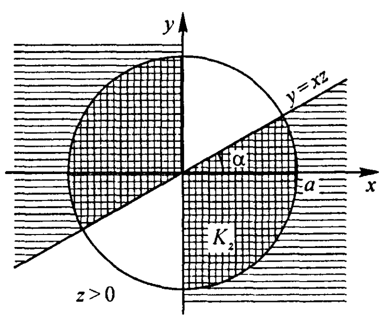


Рис. 12

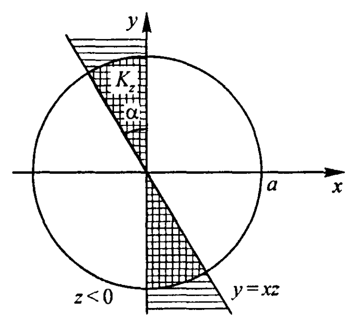


Рис. 13

Учитывая, что $f_{X,Y}(x, y)$ отлична от нуля только в круге D_a, из (5) находим

$$F_Z(z) = \frac{1}{\pi a^2} \iint\limits_{G_z \cap D_a} dx\,dy = \frac{2}{\pi a^2} \iint\limits_{K_z} dx\,dy = \frac{2}{\pi a^2} S(K_z),$$

где K_z — один из секторов, составляющий область $G_z \cap D_a$, выделенную на рис. 12 и 13 двойной штриховкой.

Так как в случае $z > 0$ площадь сектора

$$S(K_z) = \frac{1}{2}\left(\frac{\pi}{2} + \alpha\right)a^2 = \frac{1}{2}\left(\frac{\pi}{2} + \operatorname{arctg} z\right)a^2,$$

то получаем

$$F_z(z) = \frac{1}{\pi}\left(\operatorname{arctg} z + \frac{\pi}{2}\right) \quad \text{при} \quad z > 0.$$

Аналогично, при $z < 0$, как видно из рис. 13,

$$F_z(z) = \frac{1}{\pi}\left(\frac{\pi}{2} + \operatorname{arctg} z\right).$$

Дифференцируя функцию распределения по z, получаем независимо от знака z

$$f_z(z) = \frac{dF_z(z)}{dz} = \frac{1}{\pi}\cdot\frac{1}{1+z^2},$$

что соответствует закону распределения Коши. ▷

18.514. Случайный вектор (X, Y) дискретного типа распределен по закону, определяемому таблицей

x_i	y_j			
	-1	0	1	2
-1	0,05	0,03	0,15	0,05
1	0,1	0,05	0,25	0,05

Описать законы распределения случайных величин $U = |Y - X|$ и $V = Y^2 - X^2$.

18.515. Случайные величины X и Y независимы и подчиняются одному и тому же индикаторному распределению $B(1, p)$. Описать законы распределения случайных величин $Z = X + Y$ и $V = XY$.

18.516. Вычислить функцию распределения случайной величины $Z = XY$, если случайный вектор (X, Y) распределен по закону, определяемому таблицей

x_i	y_j		
	-1	0	1
0	0,1	0,2	0,1
1	0,2	0,3	0,1

18.517. Случайные величины X и Y независимы и одинаково распределены по закону $R(0, 1)$. Найти плотность распределения случайной величины $Z = Y/X^2$.

18.518. Случайный вектор (X, Y) распределен по закону, определяемому плотностью распределения вероятностей

$$f_{X,Y}(x, y) = \begin{cases} (x+y) & \text{при } 0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 1, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Найти плотности распределения вероятностей функций $Z = X + Y$, $U = XY$.

18.519 (продолжение). В условиях предыдущей задачи найти плотность распределения вероятностей $g_{U,V}(u, v)$, где $U = X^2$, $V = Y^2$.

18.520. В круг радиуса r наудачу ставится точка. Описать закон распределения расстояния от этой точки до центра круга.

18.521. Случайные величины X и Y независимы и одинаково распределены по закону $N(0, \sigma)$. Установить, по какому закону распределена случайная величина $R = \sqrt{X^2 + Y^2}$.

18.522. Случайные величины X и Y являются стандартизованными и независимыми нормальными величинами. По какому закону распределена случайная величина $Z = X/Y$?

18.523. Случайная точка (X, Y) распределена по нормальному закону с параметрами $m_X = m_Y = 0$, $\sigma_X > 0$, $\sigma_Y > 0$, $\rho_{XY} = 0$. Написать плотность совместного распределения вероятностей полярных координат точки (R, Φ).

18.524*. Пусть X и Y — две независимые случайные величины непрерывного типа. Доказать, что случайные величины X^n и Y^n $(n \in \mathbb{N})$ также независимы.

18.525*. Случайные величины X и Y независимы и одинаково распределены с функцией распределения $F(x)$. Найти функции распределения случайных величин $U = \min\{X, Y\}$ и $V = \max\{X, Y\}$.

18.526 (продолжение). В условиях предыдущей задачи найти совместную функцию распределения $F_{U,V}(u, v)$.

18.527*. Случайные величины X и Y независимы и распределены каждая по закону $R(a, b)$. Найти плотность $f_{U,V}(u, v)$ совместного распределения вероятностей случайных величин $U = \min\{X, Y\}$, $V = \max\{X, Y\}$.

4. Задача композиции. В одном из важных частных случаев функциональной зависимости $Z = \varphi(X, Y) = X + Y$ возникает задача определения закона распределения суммы компонент случайного вектора по известному закону совместного распределения его компонент. Если, например, (X, Y) — С. В. Н. Т. с известной плотностью совместного распределения компонент $f_{X,Y}(x, y)$ и $Z = X + Y$, то

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{X,Y}(x,\; z-x)\,dx = \int_{-\infty}^{+\infty} f_{X,Y}(z-y,\; y)\,dy. \tag{6}$$

Если (X, Y) — С. В. Д. Т., то закон распределения С. В. Д. Т. $Z = X + Y$ записывается в виде

$$\mathrm{P}\{Z = z_k\} = \sum_{\substack{i \quad j \\ x_i + y_j = z_k}} \sum \mathrm{P}\{X = x_i,\; Y = y_j\},$$

где суммирование распространяется на все значения индексов i и j, для которых выполняется условие $x_i + y_j = z_k$.

В частности, если (X, Y) — С. В. Д. Т. с независимыми компонентами, то

$$\mathrm{P}\{Z = z_k\} = \sum_i \mathrm{P}\{X = x_i\}\,\mathrm{P}\{Y = z_k - x_i\}. \tag{7}$$

Если (X, Y) — С. В. Н. Т. с независимыми компонентами, то формула (6) приводится к свертке двух плотностей:

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)\,f_Y(z-x)\,dx = \int_{-\infty}^{+\infty} f_X(z-y)\,f_Y(y)\,dy. \tag{8}$$

Задача определения закона распределения суммы независимых случайных величин носит название *задачи композиции*. Описанные выше формулы (7) и (8) дают непосредственное решение задачи композиции. Формулу (8) удобно применять в тех случаях, когда плотности распределения вероятностей компонент описываютя одной формулой на всей оси (что, например, справедливо для нормального закона, закона Коши и т.д.). Другой подход к решению задачи композиции основан на применении свойств 4 и 6 характеристической функции. Так как $E_Z(t) = E_X(t)\,E_Y(t)$, то, найдя $E_Z(t)$, можно по характеристической функции восстановить закон распределения случайной величины Z (согласно свойству 6).

Закон распределения W определенного вида называется *композиционно устойчивым*, если из того, что две независимые случайные величины X и Y подчиняются закону распределения данного вида, следует, что их сумма $X+Y$ подчиняются закону распределения W того же вида (см. определение вида распределения на с. 117).

18.528. X и Y — независимые случайные величины, распределенные по одному и тому же закону, определяемому таблицей

x_i	0	1	2
p_i	1/2	3/8	1/8

Описать закон распределения суммы $Z = X + Y$.

18.529. X и Y — две независимые случайные величины, подчиняющиеся одному и тому же закону геометрического распределения (см. задачу 18.330) с параметром p. Найти закон распределения их суммы $Z = X + Y$.

18.530. Доказать композиционную устойчивость закона Пуассона и найти m_Z и D_Z, где $Z = X + Y$, X и Y — независимые пуассоновские случайные величины с параметрами соответственно λ_1 и λ_2.

18.531. Доказать композиционную устойчивость закона $B(n, p)$ при фиксированном p.

18.532. Доказать композиционную устойчивость нормального закона $N(m, \sigma)$.

18.533*. Решить задачу композиции двух равномерных распределений на отрезке $[-1, 1]$. Найти плотность распределения вероятностей суммы и указать, какому закону она соответствует.

18.534. Известно, что X — случайная величина непрерывного типа с функцией распределения $F_X(x)$, Y — независимая от X случайная величина, подчиняющаяся закону распределения $R(a, b)$. Найти плотность распределения вероятностей случайной величины $Z = X + Y$.

18.535*. Измеряется некоторая физическая величина X, равномерно распределенная на отрезке $[-3, 3]$. Процесс измерения проводится в условиях воздействия аддитивной независимой от X помехи Y, распределенной по нормальному закону с параметрами $m_Y = 0$, $\sigma_Y = 2$. Написать плотность распределения вероятностей фактически измеряемой величины $Z = X + Y$.

18.536. Решить задачу композиции двух показательных распределений с параметрами, равными соответственно λ_1 и λ_2. Найти плотность распределения вероятностей суммы $Z = X+Y$ непосредственно, вычислив сначала функцию распределения, а затем плотность.

18.537 (продолжение). В условиях предыдущей задачи найти плотность распределения вероятностей Z, используя аппарат характеристических функций.

18.538* (продолжение). В условиях задачи 18.536 найти плотность распределения вероятностей Z в частном случае $\lambda_1 = \lambda_2 = \lambda$ (композиция двух одинаковых показательных распределений).

Случайная величина X непрерывного типа распределена по *закону Эрланга n-го порядка* ($n \in \mathbb{N}$) *с параметром* $\lambda > 0$, если ее плотность распределения вероятностей описывается формулой

$$f_X(x) = \begin{cases} 0, & \text{если} \quad x \leqslant 0, \\ \dfrac{\lambda(\lambda x)^n}{n!}\, e^{-\lambda x}, & \text{если} \quad x > 0. \end{cases}$$

Из решения задачи 18.538 вытекает, что композиция двух одинаковых показательных распределений с параметром λ есть распределение Эрланга первого порядка.

18.539*. Показать, что композиция n одинаковых показательных распределений с параметром λ есть распределение Эрланга $(n-1)$-го порядка с параметром λ.

18.540*. Доказать композиционную устойчивость распределения $\chi^2(n)$. В частности, показать, что сумма двух независимых распределений χ^2 соответственно с n_1 и n_2 степенями свободы есть снова распределение χ^2 с $n_1 + n_2$ степенями свободы.

18.541*. Случайная величина Z представима в виде $Z = \sum\limits_{k=1}^{Y} X_k$, где X_k — попарно независимые случайные величины, одинаково распределенные по показательному закону с параметром λ, а Y подчиняется геометрическому закону распределения с параметром p. Указать закон распределения случайной величины Z.

§ 5. Закон больших чисел и предельные теоремы теории вероятностей

1. Закон больших чисел. Следующие утверждения и теоремы составляют содержание группы законов, объединенных общим названием *закон больших чисел.*

Если случайная величина X имеет конечный первый абсолютный момент $\mathrm{M}[|X|]$, то $\forall \varepsilon > 0$

$$\mathrm{P}\{|X| \geqslant \varepsilon\} \leqslant \frac{\mathrm{M}[|X|]}{\varepsilon}.$$

В частности, если $X \geqslant 0$ и существует m_x, то

$$\mathrm{P}\{X \geqslant \varepsilon\} \leqslant \frac{m_x}{\varepsilon}$$

(*первое неравенство Чебышева*).

Если существует $\mathrm{M}[X^2]$, то при любом $\varepsilon > 0$ справедливы *второе неравенство Чебышева в нецентрированной форме*:

$$\mathrm{P}\{X \geqslant \varepsilon\} \leqslant \frac{\mathrm{M}[X^2]}{\varepsilon^2}$$

и второе неравенство Чебышева в центрированной форме:

$$\mathrm{P}\{|X - m_X| \geqslant \varepsilon\} \leqslant \frac{D_X}{\varepsilon^2}.$$

Последовательность случайных величин $X_1, X_2, \ldots, X_n, \ldots$ называется *сходящейся по вероятности при $n \to \infty$ к случайной величине* X (краткое обозначение: $X_n \xrightarrow{p} X$ при $n \to \infty$), если $\forall \varepsilon > 0$

$$\lim_{n \to \infty} \mathrm{P}\{|X_n - X| \geqslant \varepsilon\} = 0.$$

Теорема Чебышева (*закон больших чисел*). *Если случайные величины в последовательности $X_1, X_2, \ldots, X_n, \ldots$ попарно независимы, а их дисперсии удовлетворяют условию*

$$\lim_{n \to \infty} \frac{1}{n^2} \sum_{k=1}^{n} \mathrm{D}[X_k] = 0, \tag{1}$$

то $\forall \varepsilon > 0$

$$\lim_{n \to \infty} \mathrm{P}\left\{\left|\frac{1}{n} \sum_{k=1}^{n} X_k - \frac{1}{n} \sum_{k=1}^{n} \mathrm{M}[X_k]\right| \geqslant \varepsilon\right\} = 0. \tag{2}$$

Другими словами: при выполнении сформулированных условий последовательность средних арифметических n случайных величин сходится по вероятности к среднему арифметическому их математических ожиданий. В частности, если дисперсии попарно независимых случайных величин X_k, $k = 1, 2, \ldots$ равномерно ограничены (т.е. $\mathrm{D}[X_k] \leqslant \sigma^2$ для $k = 1, 2, \ldots$), то выполняется (1), а следовательно, и (2).

Теорема Маркова (*закон больших чисел в общей формулировке*). *Если дисперсии произвольных случайных величин в последовательности $X_1, X_2 \ldots, Xn, \ldots$ удовлетворяют условию*

$$\lim_{n \to \infty} \frac{1}{n^2} \mathrm{D}\left[\sum_{k=1}^{n} X_k\right] = 0, \tag{3}$$

то имеет место утверждение (2).

Пример 1. Дана последовательность случайных величин $X_1, X_2, \ldots$, удовлетворяющих условиям:

1) X_k имеют конечные вторые начальные моменты, причем

$$\mathrm{M}\,[X_k^2] < \alpha^2 \quad \text{для всех} \quad k = 1,\, 2,\, \ldots$$

2) Каждая из случайных величин в последовательности зависит лишь от случайных величин с соседними номерами.

а) Доказать, что к этой последовательности применима теорема Маркова.

б) С помощью неравенств Чебышева оценить сверху вероятность

$$\mathrm{P}\left\{\left|\frac{X_k + X_{k+1}}{2} - \frac{m_k + m_{k+1}}{2}\right| \geqslant \sigma_k + \sigma_{k+1}\right\}.$$

◁ а) Покажем прежде всего, что из конечности второго начального момента следует конечность математического ожидания. Пусть для определенности X — С. В. Н. Т. Тогда согласно неравенству Коши–Буняковского (свойство 6) математического ожидания (см. с. 107) имеем

$$m_X = \int\limits_{-\infty}^{+\infty} x f_X(x)\,dx = \int\limits_{-\infty}^{+\infty} x\sqrt{f_X(x)}\,\sqrt{f_X(x)}\,dx \leqslant$$

$$\leqslant \left(\int\limits_{-\infty}^{+\infty} x^2 f_X(x)\,dx \int\limits_{-\infty}^{+\infty} f_X(x)\,dx\right)^{1/2} = \sqrt{\mathrm{M}\,[X^2]} \leqslant \sqrt{\alpha^2}.$$

Таким образом, $|\mathrm{M}\,[X_k]\,| \leqslant \alpha$, $k = 1,\, 2,\, \ldots$ Отсюда следует, что последовательность дисперсий $\mathrm{D}\,[X_1]$, $\mathrm{D}\,[X_2]$, ... равномерно ограничена, так как

$$\mathrm{D}\,[X_k] = \mathrm{M}\,[X_k^2] - \mathrm{M}^2[X_k] \leqslant \mathrm{M}\,[X_k^2] + \mathrm{M}^2[X_k] = 2\alpha^2.$$

Для доказательства утверждения, что к указанной последовательности $X_1, X_2, \ldots$ применима теорема Маркова, достаточно проверить выполнимость условия (3). Согласно свойству 4 дисперсии (с. 107) имеем

$$\frac{1}{n^2}\,\mathrm{D}\left[\sum_{k=1}^{n} X_k\right] = \frac{1}{n^2}\left[\sum_{k=1}^{n}\sigma_k^2 + 2\sum_{i=1}^{n-1}\rho_{i,\,i+1}\sigma_i\sigma_{i+1}\right] \leqslant$$

$$\leqslant \frac{2}{n^2}\left(n\alpha^2 + 2\alpha^2\sum_{i=1}^{n-1}|\rho_{i,\,i+1}|\right) \leqslant \frac{2\alpha^2}{n^2}\,(3n-2) \to 0 \quad \text{при} \quad n \to \infty,$$

и, следовательно, теорема Маркова применима.

б) Так как

$$\mathrm{M}\left[\frac{X_k + X_{k+1}}{2}\right] = \frac{m_k + m_{k+1}}{2}, \quad \mathrm{D}\left[\frac{X_k + X_{k+1}}{2}\right] = \frac{\mathrm{D}\,[X_k + X_{k+1}]}{4},$$

то согласно второму неравенству Чебышева в центрированной форме имеем

$$\mathrm{P}\left\{\left|\frac{X_k + X_{k+1}}{2} - \frac{m_k + m_{k+1}}{2}\right| \geqslant \sigma_k + \sigma_{k+1}\right\} \leqslant \frac{\mathrm{D}\left[X_k + X_{k+1}\right]}{4\left(\sigma_k + \sigma_{k+1}\right)^2} =$$

$$= \frac{\mathrm{D}\left[X_k\right] + \mathrm{D}\left[X_{k+1}\right] + 2\rho_{k,\,k+1}\sigma_k\sigma_{k+1}}{4\left(\sigma_k + \sigma_{k+1}\right)^2} \leqslant \frac{\left(\sigma_k + \sigma_{k+1}\right)^2}{4\left(\sigma_k + \sigma_{k+1}\right)^2} = \frac{1}{4}. \ \triangleright$$

18.542. Случайная величина X имеет характеристики $m_X = 1$, $\sigma_X = 0{,}2$. Оценить снизу вероятности событий $A = \{0,5 \leqslant X < 1{,}5\}$, $B = \{0{,}75 \leqslant X < 1{,}35\}$, $C = \{X < 2\}$.

18.543. Измеряется скорость ветра в данном пункте Земли. Случайная величина X — проекция вектора скорости ветра на фиксированное направление. Оценить вероятность события $A = \{X \geqslant 80\,\text{км/ч}\}$, если путем многолетних измерений установлено, что $\mathrm{M}[|X|] = 16\,\text{км/ч}$.

18.544 (продолжение). В условиях предыдущей задачи оценить вероятность события A, если в результате проведения дополнительных измерений установлено, что $\sigma_X = 4\,\text{км/ч}$.

18.545 (продолжение). Оценить вероятность события A, если к данным задач 18.543 и 18.544 добавить условие, что закон распределения случайной величины X симметричен относительно математического ожидания m_X.

18.546. Число X солнечных дней в году для данной местности является случайной величиной со средним значением 100 дней и среднеквадратичным отклонением 20 дней. Оценить сверху вероятности событий $A = \{X \geqslant 150\}$, $B = \{X \geqslant 200\}$.

18.547. С помощью неравенств Чебышева оценить вероятности $p_k = \mathrm{P}\{|X - m_X| \leqslant k\sigma_X\}$ для $k = 1, 2, 3$, если X подчиняется закону $N(m_X, \sigma_X)$. Сравнить с точными значениями этих вероятностей, полученными в задаче 18.362.

18.548. Пусть $\varphi(x)$ — монотонно возрастающая и положительная функция, причем существует $\mathrm{M}[\varphi(X)]$, где X — некоторая неотрицательная случайная величина. Используя первое неравенство Чебышева, показать, что для всех $\varepsilon > 0$

$$\mathrm{P}\{X \geqslant \varepsilon\} \leqslant \frac{\mathrm{M}[\varphi(X)]}{\varphi(\varepsilon)}.$$

18.549*. Показать, что если для некоторой случайной величины X существует $\mathrm{M}\left[e^{\beta X}\right]$, где $\beta > 0$, то

$$\mathrm{P}\{X \geqslant \varepsilon\} \leqslant e^{-\beta\varepsilon}\,\mathrm{M}\left[e^{\beta X}\right].$$

В задачах 18.550–18.552 заданы законы распределения попарно независимых случайных величин, образующих случайную последовательность $\{X_n\}$, $n = 1, 2, \ldots$ Выяснить, применим ли к этим последовательностям закон больших чисел.

18.550.

x_{ni}	$-\sqrt{n}$	0	$\sqrt{n}$
$\mathrm{P}\{X_n = x_{ni}\}$	$\frac{1}{n}$	$1 - \frac{2}{n}$	$\frac{1}{n}$

18.551.

x_{ni}	$-na$	0	na
$\mathrm{P}\{X_n = x_{ni}\}$	$\frac{1}{2n^2}$	$1 - \frac{1}{n^2}$	$\frac{1}{2n^2}$

18.552.

x_{ni}	$-na$	0	na
$\mathrm{P}\{X_n = x_{ni}\}$	$\frac{1}{2^n}$	$1 - \frac{1}{2^{n-1}}$	$\frac{1}{2^n}$

18.552.

x_{ni}	$-\sqrt{\ln n}$	$\sqrt{\ln n}$
$\mathrm{P}\{X_n = x_{ni}\}$	$1/2$	$1/2$

18.554. Для некоторого автопарка среднее число автобусов, отправляемых в ремонт после месяца эксплуатации на городских линиях, равно 5. Оценить вероятность события $A = \{$по истечении месяца в данном автопарке будет отправлено в ремонт меньше 15 автобусов$\}$, если информация о дисперсии отсутствует.

18.555 (продолжение). Оценить вероятность события A из предыдущей задачи, если дисперсия равна 4.

2. Предельные теоремы теории вероятностей.

Теорема Бернулли. *Относительная частота успехов в n независимых испытаниях по схеме Бернулли сходится по вероятности при $n \to \infty$ к вероятности успеха в одном испытании.*

Центральная предельная теорема (*в упрощенной формулировке Ляпунова*). *Если случайные величины в последовательности $\{X_n\}$ $(n = 1, 2, \ldots)$ независимы, одинаково распределены и имеют конечные математическое ожидание m_X и дисперсию σ_X^2, то для любого действительного x*

$$F_{Y_n}(x) \underset{n\to\infty}{\longrightarrow} \Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} \exp\left\{-\frac{t^2}{2}\right\} dt,$$

где $Y_n = \dfrac{\dfrac{1}{n}\sum\limits_{k=1}^{n} X_k - m_x}{\dfrac{\sigma_x}{\sqrt{n}}}$ *— стандартизованное среднее арифметическое n случайных величин в последовательности.*

Следствиями центральной предельной теоремы являются следующие две предельные теоремы, относящиеся к схеме Бернулли:

Пусть X_n — число успехов в n независимых испытаниях по схеме Бернулли. Тогда при достаточно больших значениях npq

$$\mathrm{P}\{m_1 \leqslant X_n < m_2\} = \Phi\left(\frac{m_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{m_1 - np}{\sqrt{npq}}\right) + O\left(\frac{1}{\sqrt{npq}}\right)$$

(*интегральная теорема Муавра–Лапласа*) *и, кроме того,*

$$\mathrm{P}\{X_n = m\} = \frac{1}{\sqrt{2\pi npq}} \exp\left\{-\frac{x_m^2}{2}\right\} + O\left(\frac{1}{\sqrt{npq}}\right),$$

где $x_m = \dfrac{m - np}{\sqrt{npq}}$ (*локальная теорема Муавра–Лапласа*).

Пример 2. В одном из экспериментов Пирсона по моделированию на вычислительной машине опытов с подбрасыванием правильной монеты из общего числа 24 000 «подбрасываний» герб выпал 12012 раз.

а) Какова априорная вероятность получить данный результат?

б) Сколь вероятно при повторении эксперимента получить такое же или еще большее отклонение относительной частоты успехов от вероятности успеха в одном опыте?

$\lhd$ Пусть X_n — число выпадений герба при n подбрасываниях правильной монеты. По условию задачи $n = 24\,000$, $p = q = 1/2$. Так как $npq = 6000 \gg 1$, то применимы обе теоремы Муавра–Лапласа.

а) По локальной теореме Муавра–Лапласа

$$\mathrm{P}\{X_n = 12012\} \approx \frac{1}{\sqrt{2\pi \cdot 6000}} \exp\left\{-\frac{1}{2}\left(\frac{12}{\sqrt{6000}}\right)^2\right\} \approx 5{,}088 \cdot 10^{-3}.$$

б) Случайная величина $\dfrac{1}{n} X_n$ имеет смысл относительной частоты успехов в n опытах, причем $\mathrm{D}\left[\dfrac{1}{n} X_n\right] = \dfrac{pq}{n}$. Так как в опыте Пирсона было получено отклонение относительной частоты успехов от вероятности успеха в одном опыте, равное

$$\frac{12012}{24\,000} - \frac{1}{2} = 0{,}0005,$$

то согласно интегральной теореме Муавра–Лапласа

$$\mathrm{P}\left\{\left|\frac{1}{n}Xn-\frac{1}{2}\right|\geqslant 0{,}0005\right\}=1-\mathrm{P}\left\{\left|\frac{1}{n}X_n-\frac{1}{2}\right|<0{,}0005\right\}\approx$$

$$\approx 2-2\Phi\left(\frac{0{,}0005}{\sqrt{\dfrac{pq}{n}}}\right)=2-2\Phi(0{,}1549)\approx 0{,}8668.\ \triangleright$$

18.556. Проводятся последовательные испытания по схеме Бернулли. Вероятность осуществления события A в одном испытании $p=0{,}6$. Считая применимыми предельные теоремы Муавра–Лапласа, вычислить вероятности следующих событий:

$B=\{$событие A произойдет в большинстве из 60 испытаний$\}$, $C=\{$число успешных осуществлений события A в 60 испытаниях будет заключено между 30 и 42$\}$.

18.557 (продолжение). В условиях предыдущей задачи найти вероятность события $D=\{$событие A осуществится 36 раз в 60 испытаниях$\}$.

18.558. Радиотелеграфная станция передает цифровой текст. В силу наличия помех каждая цифра независимо от других может быть неправильно принята с вероятностью 0,01. Найти вероятности следующих событий: $A=\{$в принятом тексте, содержащем 1100 цифр, будет меньше 20 ошибок$\}$, $B=\{$будет сделано ровно 7 ошибок$\}$.

18.559. Вероятность рождения мальчика $p=0{,}512$. Считая применимыми локальную и интегральную теоремы Муавра–Лапласа, вычислить вероятность события $A=\{$среди 100 новорожденных будет 51 мальчик$\}$, $B=\{$среди 100 новорожденных будет больше мальчиков, чем девочек$\}$.

18.560 (продолжение). В условиях предыдущей задачи найти вероятность события $C=\{$разница между количеством мальчиков и количеством девочек из 100 новорожденных не превысит 10$\}$.

18.561. Складывается 10^3 чисел, каждое из которых округлено с точностью до 10^{-3}. Предполагая, что ошибки от округления независимы и равномерно распределены в интервале $(-0{,}5\cdot 10^{-3},\ 0{,}5\cdot 10^{-3})$, найти интервал, симметричный относительно математического ожидания, в котором с вероятностью 0,998 заключена суммарная ошибка.

18.562*. Случайная величина X — результат измерения некоторой физической величины, закон распределения которой неизвестен. Определить, какую максимально возможную относительную точность измерения можно гарантировать с вероятностью, не меньшей 0,95, при следующих данных: а) известно, что $m_X = 0{,}1$, $\sigma_X = 0{,}02$ и проводится одно измерение; б) проводится 5 измерений и в качестве результата X берется среднее арифметическое измеренных значений.

18.563 (продолжение). Решить предыдущую задачу, если берется среднее арифметическое 100 измерений и считается допустимым воспользоваться предельным законом распределения суммы согласно центральной предельной теореме.

18.564. Отдел технического контроля проверяет качество наудачу отобранных 900 деталей. Вероятность p того, что деталь стандартна, равна 0,9. Случайная величина X — число стандартных деталей в партии. Найти наименьший интервал, симметричный относительно m_X, в котором с вероятностью, не меньшей 0,9544, будет заключено число стандартных деталей.

18.565. В страховой компании застраховано 10 000 автомобилей. Вероятность поломки любого автомобиля в результате аварии равна 0,006. Каждый владелец застрахованного автомобиля платит в год 12 руб. страховых и в случае поломки автомобиля в результате аварии получает от компании 1000 руб. Найти вероятность события $A = \{$по истечении года работы страховая компания потерпит убыток$\}$.

18.566 (продолжение). В условиях предыдущей задачи найти вероятность события $B_m = \{$страховая компания получит прибыль не менее m руб.$\}$, если $m = 40\,000$, $60\,000$, $80\,000$.

18.567. 500 раз подбрасывается игральная кость. Какова вероятность того, что относительная частота выпадения шестерки окажется в интервале $\left(\dfrac{1}{6} - 0{,}05;\ \dfrac{1}{6} + 0{,}05\right)$?

18.568. В опыте Бюффона монета была подброшена 4040 раз, причем герб выпал 2048 раз. С какой вероятностью можно при повторении опыта получить такое же или еще большее отклонение относительной частоты успехов от вероятности успеха в одном опыте?

18.569. Сколько раз нужно подбросить монету, чтобы с вероятностью, не меньшей 0,975, утверждать, что относительная частота выпадения герба попадет в интервал (0,4; 0,6)? Получить оценку указанного числа, используя второе неравенство Чебышева.

18.570 (продолжение). Получить оценку указанного в предыдущей задаче числа подбрасываний монеты, считая применимой интегральную теорему Муавра–Лапласа.

18.570 (1). В жюри, состоящем из нечетного числа судей, каждый судья независимо от остальных принимает правильное решение с вероятностью $p = 0{,}7$. Каково должно быть наименьшее число членов жюри, при котором решение, принимаемое большинством голосов, будет правильным с гарантированной вероятностью 0,99?

18.571*. В урне содержатся белые и черные шары в отношении 3 : 2. Производятся последовательные опыты по извлечению одного шара с возвращением, причем каждый раз фиксируется цвет вынутого шара. Каково минимальное число извлечений, при котором с вероятностью, не меньшей 0,9948, можно ожидать, что отклонение относительной частоты появления белого шара от вероятности его появления в одном опыте не превысит величины $\varepsilon = 0{,}05$?

18.572. Случайная величина X распределена по закону Пуассона с параметром λ. Показать, что предельной формой закона распределения стандартизованной случайной величины $Z = \dfrac{X-\lambda}{\sqrt{\lambda}}$ при $\lambda \to \infty$ является нормальный закон $N(0,\ 1)$.

◁ Так как X — пуассоновская величина с параметром λ, то ее характеристическая функция имеет вид (см. пример 5 § 4) $E_X(t) = \exp\{\lambda(e^{it}-1)\}$. Величина Z линейно выражается через X: $Z = \dfrac{1}{\sqrt{\lambda}}X - \sqrt{\lambda}$, поэтому по свойству 2 характеристической функции

$$E_Z(t) = e^{-it\sqrt{\lambda}}\, E_X\left(\frac{t}{\sqrt{\lambda}}\right) = e^{-it\sqrt{\lambda}} e^{\lambda}\left(e^{it/\sqrt{\lambda}}-1\right).$$

Разложим выражение в скобках в ряд Тейлора по степеням $\dfrac{t}{\sqrt{\lambda}}$ с точностью до членов второго порядка малости:

$$e^{it/\sqrt{\lambda}} - 1 \approx i\,\frac{t}{\sqrt{\lambda}} - \frac{t^2}{2\lambda} + o\left(\frac{t^2}{\lambda}\right).$$

Подставляя это разложение в выражение для $E_Z(t)$, получим

$$E_Z(t) = e^{-it\lambda}e^{\lambda[it/\sqrt{\lambda} - t^2/(2\lambda) + o(t^2/\lambda)]} = e^{-t^2/2}e^{\lambda\, o(t^2/\lambda)} \underset{\lambda\to\infty}{\longrightarrow} e^{-t^2/2},$$

что соответствует виду характеристической функции стандартизованного нормального распределения (см. пример 3 § 4). Так как характеристическая функция однозначно определяет закон распределения (свойство 6), отсюда следует утверждение, сформулированное в задаче. ▷

Из решения задачи 18.572 вытекает, что при достаточно больших значениях параметра λ можно приближенно аппроксимировать пуассоновское распределение нормальным.

18.573. Оценить вероятность события A из задачи 18.554, если предполагается, что число автобусов, отправляемых в ремонт после месяца эксплуатации, подчиняется закону распределения Пуассона с параметром $\lambda = 5$.

18.574. Среднее число вызовов на АТС за 1 минуту равно $\lambda = 20 = m_X$. Найти вероятности следующих событий: $A = \{X \geqslant 20\}$, $B = \{10 \leqslant X < 30\}$.

18.575. Известно, что в среднем 5 % студентов носят очки. Какова вероятность, что из 200 студентов, сидящих в аудитории, окажется не менее 10 % носящих очки?

18.576. Рассматривается среднее арифметическое Y 100 независимых пуассоновских случайных величин X_n с параметрами $\lambda_n = n$. Найти $\mathrm{P}\{|Y - m_Y| < \lambda_1\}$.

18.577. Пусть X_n подчиняется закону распределения $\chi^2(n)$. Показать, что случайная величина

$$Y_n = \frac{X_n - n}{\sqrt{2n}}$$

асимптотически распределена по закону $N(0, 1)$, т.е.

$$F_{Y_n}(x) \to \frac{1}{\sqrt{2n}} \int\limits_{-\infty}^{x} e^{-t^2/2}\, dt \quad \text{при} \quad n \to \infty.$$

$\triangleleft$ Так как X_n распределено по закону $\chi^2(n)$, то согласно «происхождению» распределения χ^2 (см. задачу 18.494) существуют такие попарно независимые нормальные стандартизованные случайные величины Z_k $(k = 1, 2, \ldots)$, для которых $X_n = \sum\limits_{k=1}^{n} Z_k^2$. Поскольку Z_k попарно независимы, то и Z_k^2 попарно независимы $(k = 1, 2, \ldots)$ (см. задачу 18.524). Кроме того, для дисперсии любой из величин Z_k^2 имеем

$$\mathrm{D}[Z_k^2] = \mathrm{M}[Z_k^4] - \mathrm{M}^2[Z_k^2] = \mu_4 - \mu_2^2 = 3\sigma_Z^4 - \sigma_Z^4 = 2\sigma_Z^4 = 2,$$

т.е. дисперсии случайных величин Z_k $(k = 1, 2, \ldots)$ конечны. Следовательно, для случайных величин X_n $(n = 1, 2, \ldots)$ выполнены все условия центральной предельной теоремы Ляпунова, поэтому стандартизованная случайная величина Y_n асимптотически распределена по закону $N(0, 1)$. $\triangleright$

18.578. Случайная величина X распределена по закону $\chi^2(200)$. Используя асимптотическую нормальность X, оценить вероятность события $A = \{X \geqslant 250\}$.

18.579.** Случайная величина X распределена по закону $\mathrm{St}(n)$. Доказать, что при $n \to \infty$ плотность распределения вероятностей случайной величины X сходится при любом значении аргумента x к плотности распределения вероятностей закона $N(0, 1)$ при том же значении x.

18.580. Случайная величина T распределена по закону $\mathrm{St}\,(60)$. Найти приближенно квантиль $t_{0,9}(60)$, используя нормальное приближение, и сравнить с точным значением, полученным с помощью таблицы П6.

3. Метод статистических испытаний. Предельные теоремы теории вероятностей и их следствия лежат в основе вычислительного метода, известного как *метод статистических испытаний* (*метод Монте-Карло*), который часто применяется для приближенного вычисления определенных интегралов, не выражающихся через элементарные функции, для решения систем линейных алгебраических уравнений высокого порядка, при решении краевых задач для уравнений в частных производных и во многих других задачах численного анализа, где обычные численные методы (такие, как метод сеток) приводят к чрезмерно большому объему вычислений.

Идею метода статистических испытаний проиллюстрируем на примере приближенного вычисления определенного интеграла от функции $\varphi\,(x)$. Пусть

$$I = \int\limits_a^b \varphi\,(x)\,dx = \frac{1}{b-a}\int\limits_a^b (b-a)\,\varphi\,(x)\,dx.$$

Значение интеграла I можно рассматривать как математическое ожидание функции случайной величины $Y = (b-a)\,\varphi\,(X)$, где X распределена по закону $R\,(a,\,b)$. Пусть X_k $(k = 1,\,2,\,\ldots)$ — независимые случайные числа, распределенные по закону $R\,(a,\,b)$ (существуют стандартные программы на ЭВМ, вырабатывающие так называемые «псевдослучайные» числа, которые обычно используют в качестве X_k). Случайную величину

$$I_n = \frac{1}{n}\sum_{k=1}^{n}(b-a)\,\varphi\,(X_k) \tag{4}$$

можно рассматривать как приближенное значение интеграла I. Если дисперсия $\mathrm{D}\,[\varphi\,(X)]$ существует, то из закона больших чисел следует, что $I_n \xrightarrow{p} I$ при $n \to \infty$ (следовательно, вероятности больших отклонений I_n от $\mathrm{M}\,[I_n] = I$ малы при больших n), а из центральной предельной теоремы (в силу независимости $\varphi\,(X_k)$ $(k = 1,\,2,\,\ldots)$ и ограниченности дисперсий $\mathrm{D}\,[\varphi\,(X_k)]$) вытекает, что случайная величина

$$U = \frac{I_n - I}{\sqrt{\mathrm{D}\,(I_n)}} = \frac{I_n - I}{\sigma_Y}\sqrt{n}$$

асимптотически (при больших n) распределена по закону $N\,(0,\,1)$.

Пусть $\delta = \left|\dfrac{I_n - I}{I}\right|$ — относительная погрешность вычисления интеграла I. Так как распределение случайной величины U известно (стандартизованное нормальное), то можно найти такое минимальное значение относительной погрешности, которое гарантируется с заданной вероятностью p. Обозначим это значение $\delta_{\min}$. По определению,

$$P\{\delta < \delta_{\min}\} = p \quad \text{или} \quad P\left\{|U| < \delta_{\min}\frac{|I|}{\sigma_Y}\sqrt{n}\right\} = p.$$

Отсюда следует, что

$$\delta_{\min} = \frac{1}{\sqrt{n}} \cdot \frac{\sigma_Y}{|I|}\, t_{(1+p)/2}\,,$$

где $t_{(1+p)/2}$ — квантиль порядка $\dfrac{1+p}{2}$ для нормального распределения $N(0, 1)$. Так как σ_Y конечно, то $\sigma_{\min}$ может быть сделано сколь угодно малым при увеличении n.

Пример 3. Какой объем n выборки необходимо взять в методе статистических испытаний для того, чтобы при вычислении значения гамма-функции Эйлера $\Gamma(5/4)$ можно было с вероятностью, не меньшей 0,95, считать, что относительная ошибка вычисления будет меньше одного процента?

$\lhd$ Воспользуемся выражением для гамма-функции Эйлера в виде определенного интеграла:

$$\Gamma(z) = \int\limits_0^\infty t^{z-1}e^{-t}\,dt.$$

Путем замены переменных $e^{-t} = x$ данный интеграл приводится к конечному промежутку:

$$\Gamma(z) = \int\limits_0^1 \ln^{z-1}\left(\frac{1}{x}\right)\,dx.$$

Обозначим через I искомое значение гамма-функции:

$$I = \Gamma\left(\frac{5}{4}\right) = \int\limits_0^1 \ln^{1/4}\left(\frac{1}{x}\right)\,dx.$$

Очевидно, I представимо в виде $I = M[Y]$, где $Y = \varphi(X) = \ln^{1/4}\left(\dfrac{1}{X}\right)$, и X равномерно распределена на отрезке $[0, 1]$. Вычислим величину

$\sigma_Y/|I|$. Используя известную формулу, выражающую дисперсию через второй начальный момент, получим

$$D_\varphi = \mathrm{D}\left[\ln^{1/4}\left(\frac{1}{X}\right)\right] = \mathrm{M}\,[\varphi^2(X)] - \mathrm{M}^2[\varphi\,(X)] = \int\limits_0^1 \ln^{1/2}\left(\frac{1}{x}\right)\,dx - I^2.$$

Отсюда следует, что

$$\frac{\sigma_Y}{|I|} = \sqrt{\frac{\mathrm{M}\,[\varphi^2(X)]}{I^2} - 1} = \sqrt{\frac{1}{I^2}\int\limits_0^1 \ln^{1/2}\left(\frac{1}{x}\right)\,dx - 1}\,.$$

Используя известное свойство гамма-функции: $\Gamma\,(z+1) = z\Gamma\,(z)$ и учитывая частное значение $\Gamma\,(1/2) = \sqrt{\pi}$ (интеграл Пуассона), получим значение интеграла в правой части:

$$\int\limits_0^1 \ln^{1/2}\left(\frac{1}{x}\right)\,dx = \Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\,\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}\,.$$

Чтобы найти точное значение величины $\sigma_Y/|I|$, необходимо знать точное значение интеграла I, а оно только еще вычисляется. В этом состоит известная трудность при оценивании погрешности в методе статистических испытаний. Обычно вместо точного значения величины I используют ту или иную ее приближенную оценку снизу (а для σ_Y по тем же причинам часто используют оценку сверху), при этом получают несколько завышенное значение величины $\sigma_Y/|I|$, что приводит к более «осторожным» оценкам для относительной ошибки. Используя для данного примера таблицы гамма-функции [3]), найдем, что с точностью до $5\cdot 10^{-5}$ искомое значение $\Gamma\,(5/4) = 0{,}9065$, и, таким образом, $\sigma_Y/|I| = 0{,}280$. Далее, по условию задачи $\mathrm{P}\,\{\delta < 0{,}01\} \geqslant 0{,}95$, или $\mathrm{P}\left\{|T| < 0{,}01\sqrt{n}\,\dfrac{|I|}{\sigma_Y}\right\} \geqslant 0{,}95$. Отсюда следует, что

$$\sqrt{n} \geqslant \widehat{t}_{0{,}975}\cdot\frac{0{,}280}{0{,}01} = 54{,}88;$$

следовательно, наименьший объем выборки, обеспечивающий относительную погрешность вычисления $\Gamma\,(5/4)$ не более 1 % с вероятностью не менее 0,95, $n_{\min} = [54{,}88^2 + 0{,}5] = 3012$, где $[x]$ означает целую часть числа x. ▷

[3]) См., например: *Янке Е., Эмде Ф., Леш Ф.* Специальные функции.— М.: Наука, 1977. С. 55.

18.581. Вычисление интеграла $\int\limits_0^{\pi/2} \cos x\,dx$ производится методом Монте-Карло на основании 1000 испытаний. Какую максимальную относительную погрешность вычисления можно гарантировать с надежностью 97,22 %?

18.582*. Сколько требуется провести статистических испытаний при вычислении значения функции нормального распределения $\Phi(1)$, чтобы с вероятностью, не меньшей 0,95, гарантировать величину относительной погрешности в пределах 2 %?

18.583. Вычисление интеграла $I = \int\limits_0^1 x^2\,dx$ производится методом Монте-Карло на основании 10^4 независимых испытаний. Вычислить вероятность того, что относительная погрешность вычисления не превзойдет 1 %.

18.584*. Вычисление числа π производится методом статистических испытаний, состоящих в n независимых случайных бросаниях иглы длины l на плоскость, разграфленную параллельными прямыми, отстоящими друг от друга на расстояние $2l$ (см. задачу Бюффона 18.159). Измеряется число пересечений иглой любой из параллельных прямых. Найти наименьшее число испытаний, которые требуется провести, чтобы с вероятностью 0,9996 относительная погрешность определения числа π была не более 3 %.

§ 6. Случайные функции (корреляционная теория)

1. Законы распределения и осредненные характеристики случайных функций. Пусть G_t — некоторое множество действительных чисел. Если каждому значению $t \in G_t$ поставлена в соответствие случайная величина $X(t)$, то говорят, что на множестве G_t задана *случайная функция* $X(t)$. Множество G_t при этом называется областью определения *случайной функции*. Название *случайный процесс* относится к классу случайных функций, у которых параметр t играет роль времени.

Случайная величина $X(t_0)$, соответствующая значению случайной функции при фиксированном значении аргумента $t = t_0 \in G_t$, называется *сечением*. Каждое испытание дает конкретную функцию $x(t)$, которая называется *реализацией* (*траекторией*) случайной функции.

n-мерным законом распределения случайной функции $X(t)$, *зависящим от* n *действительных параметров* $t_1, t_2, \ldots, t_n$, называется закон совместного распределения n сечений

$$(X(t_1),\ X(t_2), \ldots, X(t_n)).$$

В *корреляционной теории* описание характерных свойств случайных функций строится на основе не более чем двумерных законов распределения.

Одномерная функция распределения $F_1(x/t)$ значений случайной функции $X(t)$ при фиксированном t представляет собой функцию распределения сечения $X(t)$:

$$F_1(x/t) = \mathrm{P}\{X(t) < x\}.$$

Соответствующая одномерная плотность существует, если сечение $X(t)$ — С. В. Н. Т., причем в точках дифференцируемости функции $F_1(x/t)$ справедливо равенство

$$f_1(x/t) = \frac{\partial F_1(x/t)}{\partial x}.$$

Если сечение $X(t)$ — С. В. Д. Т., то одномерный закон распределения описывается перечнем вероятностей

$$\mathrm{P}\{X(t) = x_k(t)\} = p_k(t) = F_1(x_{k+1}/t) - F_1(x_k/t),$$
$$\sum_k p_k(t) = 1, \quad t \in G_t.$$

Двумерной функцией распределения $F_2(x, y/t_1, t_2)$ называется функция совместного распределения двух сечений случайной функции $(X(t_1), X(t_2))$, $t_1, t_2 \in G_t$:

$$F_2(x, y/t_1, t_2) = \mathrm{P}\{X(t_1) < x,\ X(t_2) < y\}.$$

Соответствующая двумерная плотность существует, если случайный вектор $(X(t_1), X(t_2))$ — С. В. Н. Т., причем если в точке (x, y) функция $F_2(x, y/t_1, t_2)$ дважды дифференцируема, то

$$f_2(x, y/t_1, t_2) = \frac{\partial^2 F_2(x, y/t_1, t_2)}{\partial x\, \partial y}.$$

Зная двумерную плотность, можно по общему правилу (см. формулу (4 § 3) вычислить одномерную плотность сечения $X(t)$:

$$f_1(x/t) = \int_{-\infty}^{+\infty} f_2(x, y/t_1, t_2)\, dy. \tag{1}$$

Если случайный вектор $(X(t_1), X(t_2))$ — С. В. Д. Т., то двумерный закон его распределения описывается перечнем вероятностей

$$\mathrm{P}\{X(t_1) = x_i(t_1),\ X(t_2) = y_j(t_2)\} = p_{ij}(t_1, t_2),$$
$$\sum_{i,j} p_{ij}(t_1, t_2) = 1, \quad t_1, t_2 \in G_t.$$

Основными характеристиками случайных функций являются *математическое ожидание, дисперсия* и *автоковариационная функция.*

Математическим ожиданием и *дисперсией* случайной функции $X(t)$ называются такие неслучайные функции $m_X(t)$ и $D_X(t) = \sigma_X^2(t)$, которые для каждого фиксированного значения t равны математическому ожиданию и дисперсии соответствующего сечения. Таким образом, если, например, $X(t)$ — С. В. Н. Т. при $t \in G_t$ и написанные ниже интегралы абсолютно сходятся, то

$$m_X(t) = \mathrm{M}[X(t)] = \int\limits_{-\infty}^{+\infty} x f_1(x/t)\,dx,$$

$$D_X(t) = \sigma_X^2 = \mathrm{D}[X(t)] = \mathrm{M}[\overset{\circ}{X}{}^2(t)] = \int\limits_{-\infty}^{+\infty} [x - m_X(t)]^2 f_1(x/t)\,dx.$$

Автоковариационной функцией называется такая неслучайная функция $K_X(t_1, t_2)$ двух действительных аргументов t_1 и t_2, которая для каждой пары фиксированных значений t_1 и t_2 равна ковариации соответствующих сечений

$$K_X(t_1, t_2) = \mathrm{M}[(X(t_1) - m_X(t_1))(X(t_2) - m_X(t_2))] =$$
$$= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} (x_1 - m_X(t_1))(x_2 - m_X(t_2))\, f_2(x_1, x_2/t_1, t_2)\,dx_1\,dx_2.$$

Нормированная автоковариационная функция

$$\rho_X(t_1, t_2) = \frac{K_X(t_1, t_2)}{\sigma_X(t_1)\,\sigma_X(t_2)}$$

называется *автокорреляционной функцией.* Основные свойства автоковариационной (автокорреляционной функции):

1. $K_X(t_1, t_2) = K_X(t_2, t_1)$ $(\rho_X(t_1, t_2) = \rho_X(t_2, t_1))$ (свойство симметрии).

2. $|K_X(t_1, t_2)| \leqslant \sigma_X(t_1)\,\sigma_X(t_2)$ $(|\rho_X(t_1, t_2)| \leqslant 1)$.

3. Функция $K_X(t_1, t_2)$ неотрицательно определенная.

Это значит, что для любого натурального n, любых вещественных значений $x_1, x_2, \ldots, x_n$ и любого набора значений аргументов $t_1, t_2, \ldots, t_n$ из области определения функции $K_X(t', t'')$ имеет место неравенство

$$\sum_{i=1}^{n}\sum_{j=1}^{n} K_X(t_i, t_j)\,x_i x_j \geqslant 0.$$

Взаимной ковариационной функцией связи двух действительных случайных функций $X(t)$ и $Y(t)$ называется неслучайная функция

$$K_{XY}(t_1, t_2) = \mathrm{M}[\overset{\circ}{X}(t_1)\,\overset{\circ}{Y}(t_2)],$$

где $\overset{\circ}{X}(t_1)$ и $\overset{\circ}{Y}(t_2)$ — центрированные сечения случайных функций.

Взаимная корреляционная функция связи (*нормированная автоковариационная функция связи*) определяется равенством

$$\rho_{XY}(t_1, t_2) = \frac{K_{XY}(t_1, t_2)}{\sigma_X(t_1)\,\sigma_Y(t_2)}.$$

В дальнейшем для краткости часто будем опускать приставку «авто» в названии «автоковариационная функция» в тех случаях, когда речь будет идти о характеристиках самого процесса $X(t)$.

Пример 1. Задана двумерная плотность случайного процесса $X(t)$ в виде

$$f_2(x, y/t_1, t_2) = \frac{1}{2\pi}\exp\left\{-\frac{(x+\sin t_1)^2 + (y+\sin t_2)^2}{2}\right\}.$$

Вычислить основные характеристики процесса $m_X(t)$, $D_X(t)$ и $K_X(t_1, t_2)$.

◁ Найдем сначала одномерную плотность процесса. Из формулы (1) получаем

$$f_1(x/t) = \int_{-\infty}^{+\infty} f_2(x, y/t, t_2)\,dy =$$

$$= \frac{1}{2\pi}\exp\left\{-\frac{(x+\sin t)^2}{2}\right\}\int_{-\infty}^{+\infty}\exp\left\{-\frac{(y+\sin t_2)^2}{2}\right\}dy.$$

С помощью замены $u = y + \sin t_2$ приходим к интегралу Пуассона

$$f_1(x/t) = \frac{1}{2\pi}\exp\left\{-\frac{(x+\sin t_1)^2}{2}\right\}\int_{-\infty}^{+\infty}\exp\left\{-\frac{u^2}{2}\right\}du =$$

$$= \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(x+\sin t_1)^2}{2}\right\}.$$

Отсюда, в частности, следует, что сечения $X(t_1)$ и $X(t_2)$ при $t_1 \neq t_2$ независимы, так как

$$f_2(x, y/t_1, t_2) = f_1(x/t_1) f_2(y/t_2).$$

Для математического ожидания процесса по определению имеем

$$m_X(t) = \int_{-\infty}^{+\infty} x f_1(x/t)\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} x\exp\left\{-\frac{(x+\sin t)^2}{2}\right\}dx.$$

Используя в интеграле замену переменных $u = x + \sin t$, получаем

$$m_X(t) = \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{+\infty} u \exp\left\{-\frac{u^2}{2}\right\} du - \sin t \int_{-\infty}^{+\infty} \exp\left\{-\frac{u^2}{2}\right\} du \right) = -\sin t.$$

По определению ковариационной функции

$$K_X(t_1, t_2) = \mathrm{M}[\overset{\circ}{X}(t_1)\overset{\circ}{X}(t_2)] = \mathrm{M}[X(t_1)X(t_2)] - m_X(t_1)m_X(t_2).$$

Так как при $t_1 \neq t_2$ сечения процесса независимы, то

$$\mathrm{M}[X(t_1)X(t_2)] = \mathrm{M}[X(t_1)]\,\mathrm{M}[X(t_2)] = \sin t_1 \sin t_2,$$

откуда следует, что $K_X(t_1, t_2) = 0$ при $t_1 \neq t_2$. Если же $t_1 = t_2 = t$, то по определению

$$K_X(t, t) = D_X(t) = \int_{-\infty}^{+\infty} x^2 f_1(x/t)\,dx - m_X^2(t).$$

Подставляя сюда выражение для плотности $f_1(x/t)$ и проводя замену переменных, получаем

$$D_X(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 \exp\left\{-\frac{(x+\sin t_1)^2}{2}\right\} dx - \sin^2 t =$$

$$= \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{+\infty} u^2 \exp\left\{-\frac{u^2}{2}\right\} du - 2\sin t \int_{-\infty}^{+\infty} u \exp\left\{-\frac{u^2}{2}\right\} du + \right.$$

$$\left. + \sin^2 t \int_{-\infty}^{+\infty} \exp\left\{-\frac{u^2}{2}\right\} du \right) - \sin^2 t = 1. \ \triangleright$$

18.585. Двумерный закон распределения случайной функции $X(t)$ описывается плотностью

$$f_2(x, y/t_1, t_2) = \frac{1}{2\pi\sqrt{(1+t_1^2)(1+t_2^2)}} \exp\left\{-\frac{(x-vt_1)^2}{2(1+t_1^2)} - \frac{(y-vt_2)^2}{2(1+t_2^2)}\right\},$$

где $v > 0$. Найти основные характеристики: $m_X(t)$, $D_X(t)$ и $K_X(t_1, t_2)$.

18.586. Случайная величина является частным случаем такой случайной функции, у которой отсутствует зависимость от t. Пусть $X(t) \equiv X$ для всех $t \in \mathbb{R}$, причем X — С. В. Н. Т., подчиняющаяся показательному распределению с параметром $\lambda = 2$. Найти $m_X(t)$, $D_X(t)$ и $F_2(x, y/t_1, t_2)$.

18.587. Случайный процесс $X(t)$ имеет вид

$$X(t) = Vt^2 \quad (t > 0),$$

где V — случайная величина, равномерно распределенная на $[0, 3]$. Найти одномерную функцию распределения и одномерную плотность этого процесса.

18.588. Случайная функция $X(t)$ задана в виде $X(t) = Vt + b$, где V — С. В. Н. Т., подчиняющаяся закону $N(m, \sigma)$, а b — неслучайная константа. Найти одномерную плотность $f_1(x/t)$ и основные характеристики процесса: $m_X(t)$, $\sigma_X(t)$, $K_X(t_1, t_2)$.

18.589. Случайная функция $X(t)$ задана в виде $X(t) = U + Vt$, где U и V — независимые случайные величины, подчиняющиеся одному и тому же закону распределения $N(m, \sigma)$. Используя свойства математического ожидания и дисперсии, вычислить $m_X(t)$, $D_X(t)$ и $K_X(t_1, t_2)$.

18.590 (продолжение). В условиях предыдущей задачи записать одномерную плотность $f_1(x/t)$.

18.591. Заданы плотности $f_U(u)$ и $f_V(v)$ независимых случайных величин U и V. Записать одномерную плотность $f_1(x/t)$ процесса $X(t) = U + Vt$ при $t > 0$.

18.591 (1). Случайная функция $X(t)$ задана в виде $X(t) = W\exp\{-Vt\}$, где случайный вектор (W, V) подчиняется круговому нормальному распределению с центром в точке $(0, 0)$ и ковариационной матрицей $K = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$. Найти $m_X(t)$, $D_X(t)$ и $K_X(t_1, t_2)$.

18.592. Заданы ковариационные функции $K_X(t_1, t_2)$ и $K_Y(t_1, t_2)$ и математические ожидания $m_X(t)$ и $m_Y(t)$ двух независимых случайных процессов $X(t)$ и $Y(t)$. Найти ковариационную функцию процесса $Z(t) = X(t)\,Y(t)$.

18.593. Показать, что если две случайные функции $X(t)$ и $Y(t)$ независимы при любом фиксированном t и имеют нулевые математические ожидания, то автоковариационная функция их произведения равна произведению автоковариационных функций отдельных сомножителей.

18.594. Доказать следующее свойство автоковариационной функции: если $Y(t) = \psi(t) X(t) + \varphi(t)$, где $\varphi(t)$ и $\psi(t)$ — неслучайные функции, то

$$K_Y(t_1, t_2) = \psi(t_1)\psi(t_2) K_X(t_1, t_2).$$

(Отсюда, в частности, следует, что прибавление к случайному процессу неслучайной функции не изменяет автоковариационной функции.)

18.595. Дана ковариационная функция случайного процесса $X(t)$:

$$K_X(t_1, t_2) = \frac{1}{1 + (t_1 - t_2)^2}.$$

Найти ковариационную функцию и дисперсию процесса

$$Y(t) = e^{-t^2} X(t) + \sin 2t.$$

18.596. Случайный процесс $Z(t)$ задан в виде

$$Z(t) = X(t) + tY(t) + t^2,$$

где $X(t)$ и $Y(t)$ — некоррелированные случайные процессы с характеристиками

$$m_X(t) = 4, \ K_X(t_1, t_2) = 9e^{-2|t_2 - t_1|},$$

$$m_Y(t) = 1, \ K_Y(t_1, t_2) = 4e^{-2|t_2 - t_1|}.$$

Найти $m_Z(t)$ и $D_Z(t)$.

18.597 (продолжение). В условиях предыдущей задачи процессы $X(t)$ и $Y(t)$ коррелированы, причем

$$K_{XY}(t_1, t_2) = 6e^{-2|t_2 - t_1|}.$$

Найти корреляционную функцию $\rho_Z(t_1, t_2)$.

18.598. Заданы случайные функции

$$X(t) = -U \sin t + V \cos t,$$

$$Y(t) = U \cos t + V \sin t,$$

где U и V — некоррелированные стандартизованные случайные величины. Найти автокорреляционные функции $\rho_X(t_1, t_2)$ и $\rho_Y(t_1, t_2)$ процессов $X(t)$ и $Y(t)$, а также корреляционную функцию связи $\rho_{XY}(t_1, t_2)$.

18.599. Случайный процесс $X(t)$ задан в виде $X(t) = \varphi(t, Y)$, где $\varphi(t, y)$ — произвольная неслучайная функция двух действительных аргументов, t — действительный параметр (время), а Y — С. В. Н. Т. с известной плотностью распределения $f_Y(y)$. Записать выражения для основных характеристик процесса: $m_X(t)$, $D_X(t)$, $K_X(t_1, t_2)$.

18.600. Реализации случайного процесса $X(t)$ формируются следующим образом. В начальный момент времени значение функции с равной вероятностью равно либо $+1$, либо -1. Смена значений функции может происходить лишь в фиксированные моменты времени $t_k = k$ ($k = 1, 2, \ldots$), в каждый из которых независимо от предыдущих значений происходит очередной розыгрыш одного из двух равновероятных значений: $+1$ или -1, которое и сохраняется до следующего момента времени $t_{k+1} = k + 1$. Одна из возможных реализаций процесса изображена на рис. 14. Определить основные характеристики: $m_X(t)$, $D_X(t)$ и $K_X(t, t + \tau)$.

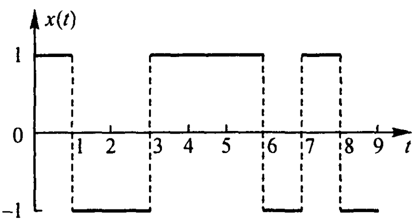


Рис. 14

18.601. Случайный процесс $X(t)$ представляет собой *случайную ступеньку* $X(t) = A\eta(t - T)$, где

$$\eta(u) = \begin{cases} 1, & u > 0, \\ 0, & u \leqslant 0, \end{cases}$$

— единичная функция Хевисайда, A — случайная амплитуда с характеристиками $m_A > 0$, σ_A^2; T — случайное, независимое от A время начала ступеньки, распределенное по закону с плотностью $f_T(t)$. Найти $m_X(t)$ и $K_X(t, t + \tau)$ при $\tau > 0$.

Случайный процесс $X(t)$ называется *нормальным* (или *гауссовским*) процессом, если одномерные и двумерные законы распределения любых его сечений нормальны.

18.602. Случайное гармоническое колебание задано в виде $X(t) = A\cos\omega t + B\sin\omega t$, где ω — неслучайная частота, а случайные амплитуды A и B независимы и подчиняются каждая закону распределения $N(0, \sigma)$. Найти одномерную и двумерную плотности процесса.

18.603. Случайный процесс $Z(t)$ задан в виде $Z(t) = aX(t) +$

ками

$$m_X(t) = t, \quad m_Y(t) = 1 + t^2,$$

$$K_X(t_1,\, t_2) = \frac{4}{1 + 2\,(t_1 - t_2)^2}, \quad K_Y(t_1,\, t_2) = 9e^{-2\,(t_2 - t_1)^2},$$

а ковариационная функция связи процессов $X\,(t)$ и $Y\,(t)$ имеет вид $K_{XY}(t_1,\, t_2) = 4\cos\omega\,(t_2 - t_1)$. Написать одномерную плотность процесса $Z\,(t)$.

18.604 (продолжение). В условиях предыдущей задачи найти автокорреляционную функцию $\rho_Z(t_1,\, t_2)$.

18.605*. Угол крена корабля $X\,(t)$ представляет собой нормальный случайный процесс с характеристиками $m_X = 0$, $K_X(t, t + \tau) = \sigma^2 \rho_X(\tau)$. Известно, что в момент времени t_1 угол крена корабля составлял $X\,(t_1) = \alpha$ градусов. Какова вероятность того, что в момент $t_2 = t_1 + \tau$ угол крена будет больше, чем β градусов?

Процесс $X\,(t)$, определенный при $t \in G_t = [a,\, b]$, называется *процессом с независимыми приращениями*, если для любых $t_0,\, t_1,\, \ldots,\, t_n$ таких, что $a \leqslant t_0 \leqslant t_1 \leqslant \ldots \leqslant t_n \leqslant b$, случайные величины $X\,(t_0)$, $X\,(t_1) - X\,(t_0),\, \ldots,\, X\,(t_n) - X\,(t_{n-1})$ независимы.

Случайный процесс с независимыми приращениями называется *однородным*, если закон распределения случайной величины $X\,(t) - X\,(t_0)$ не зависит от t_0, а определяется лишь длиной интервала $(t_0,\, t)$.

Пример 2 (*пуассоновский процесс*). Случайный процесс $N\,(t)$ удовлетворяет следующим условиям:

1) $N\,(t)$ определен при всех $t \geqslant 0$, причем $P\,\{N\,(0) = 0\} = 1$.

2) $N\,(t)$ — однородный по времени.

3) $N\,(t)$ — процесс с независимыми приращениями.

4) В случайный момент времени происходит приращение значения функции $N\,(t)$ на единицу, причем для любого момента времени $t \geqslant 0$

$$P\,\{N\,(t + \Delta t) - N\,(t) = 1\} = \lambda\Delta t + o\,(\Delta t),$$

$$P\,\{N\,(t + \Delta t) - N\,(t) = 0\} = 1 - \lambda\Delta t + o\,(\Delta t),$$

$$P\,\{N\,(t + \Delta t) - N\,(t) \geqslant 2\} = o\,(\Delta t),$$

где λ — постоянное для данного процесса число.

На рис. 15 показана одна из реализаций пуассоновского процесса. Найти одномерный закон распределения случайной функции $N\,(t)$.

$\lhd$ Обозначим $\Delta N\,(t) = N\,(t + \Delta t) - N\,(t)$ случайное приращение процесса за время Δt и положим $p_n(t) = \mathrm{P}\{N\,(t) = n\}$, $n = 0,\, 1,\, 2,\, \ldots$ Задача состоит в отыскании значений вероятностей $p_n(t)$ для $n = 1,\, 2,\, \ldots$ и всех $t > 0$.

Для данного процесса событие $\{N(t) = n\}$ означает «число единичных приращений за время t равно n». Выразим сложное событие $\{N(t+\Delta t) = n\}$ в алгебре событий следующим образом:

$$\{N(t+\Delta t) = n\} = \sum_{k=0}^{n} \{N(t) = n-k\}\{\Delta N(t) = k\}.$$

По условию 3) $N(t)$ — процесс с независимыми приращениями. Полагая $t_0 = 0$, $t_1 = t$, $t_2 = t + \Delta t$ $(\Delta t > 0)$, получаем по определению,

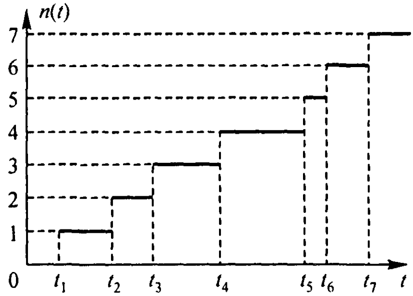


Рис. 15

что случайные величины $N(t_1) - N(t_0) = N(t)$ и $N(t_2) - N(t_1) = \Delta N(t)$ независимы. Поэтому пары событий $\{N(t) = n-k\}$ и $\{\Delta N(t) = k\}$ независимы при $\forall k = 0, 1, \ldots$ Применяя формулы сложения и умножения вероятностей и используя условия 4), получим

$$p_n(t+\Delta t) = \sum_{k=0}^{n} p_{n-k}(t)\mathrm{P}\{\Delta N(t) = k\} = p_n(t)\ (1 - \lambda\Delta t + o(\Delta t)) +$$

$$+p_{n-1}(t)\ (\lambda\Delta t + o(\Delta t)) + \sum_{k=2}^{n} p_{n-k}(t)\mathrm{P}\,\{\Delta N(t) = k\}.$$

Учитывая, что $0 \leqslant p_{n-k}(t) \leqslant 1$, $k = 0, 1, \ldots, n$, оценим последнюю сумму:

$$0 \leqslant \sum_{k=2}^{n} p_{n-k}(t)\mathrm{P}\,\{\Delta N(t) = k\} \leqslant$$

$$\leqslant \sum_{k=2}^{n} \mathrm{P}\,\{\Delta N(t) = k\} \leqslant \sum_{k=2}^{\infty} \mathrm{P}\,\{\Delta N(t) = k\} = \mathrm{P}\,\{\Delta N(t) \geqslant 2\} = o(\Delta t)$$

в силу последнего из условий 4).

После элементарных преобразований окончательно получаем

$$p_n(t+\Delta t) - p_n(t) = -\lambda\Delta t\,\big(p_n(t) - p_{n-1}(t)\big) + o(\Delta t).$$

Деля обе части на Δt и переходя к пределу при $\Delta t \to 0$, получим дифференциальное уравнение для искомых вероятностей

$$\frac{dp_n(t)}{dt} = -\lambda p_n(t) + \lambda p_{n-1}(t), \qquad n = 1, 2, \ldots, \tag{2}$$

с начальными условиями

$$p_n(0) = 0, \quad n = 1, 2, \ldots,$$
$$p_0(0) = 1.$$

Уравнение (2) решаем рекуррентно, считая, что на n-м шаге значение $p_{n-1}(t)$ уже известно. При $n = 0$ решением задачи

$$p_0'(t) = -\lambda p_0(t),$$
$$p_0(0) = 1$$

является функция $p_0(t) = e^{-\lambda t}$. При $n \neq 0$ согласно интегралу Дюамеля (см. ч. 3, гл. 14, § 1, п. 1)

$$p_n(t) = \lambda \int_0^t e^{-\lambda (t-\tau)} p_{n-1}(\tau)\, d\tau,$$

откуда по индукции получаем

$$p_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad t \geqslant 0, \quad n = 0, 1, \ldots$$

Параметр λ трактуется как среднее число единичных приращений в единицу времени. ▷

К пуассоновским процессам относятся процесс радиоактивного распада ($N(t)$ — число атомов, распавшихся за время t, λt — среднее число атомов, распадающихся за время t), поток заявок на АТС ($N(t)$ — число вызовов на АТС за время t, λ — число вызовов в единицу времени), сбои радиоэлектронной аппаратуры ($N(t)$ — число элементов аппаратуры, вышедших из строя за время t, λ — среднее число отказов в единицу времени) и т.д.

18.606. Пусть $X(t)$ процесс с независимыми приращениями, удовлетворяющий условию $X(t) = 0$ при $t \leqslant 0$. Доказать, что дисперсия $D_X(t)$ является неубывающей функцией t.

18.607. Процесс $N(t)$ представляет собой простейший пуассоновский поток отказов радиотехнической системы с интенсивностью 0,002 отказа в час. Найти вероятность того, что за 100 часов наступит не менее 3 отказов.

18.608 (продолжение). В условиях предыдущей задачи найти вероятность того, что за 200 часов работы системы поступит четное число отказов.

18.609. АТС обслуживает 6000 абонентов, каждый из которых в среднем занимает линию связи в течение одной минуты в час. Какое минимальное число каналов n надо иметь на АТС, чтобы вероятность того, что число поступивших в течение одной минуты вызовов превысит число каналов, была не более 0,003?

18.610. Найти математическое ожидание и дисперсию процесса Пуассона $N(t)$, определенного в примере 2.

18.611. $X(t)$ — пуассоновский процесс с параметром λ. Описать условный закон распределения

$$\mathrm{P}\{X(t_2) = m/X(t_1) = n\}, \quad m, n \in \mathbb{N}; \quad t_2 > t_1.$$

18.612*. Вычислить автоковариационную функцию процесса Пуассона $X(t)$ с параметром λ.

18.612 (1). Пусть $X(t)$ — пуассоновский процесс с параметром $\lambda > 0$. Обозначим $Q(\tau) = \mathrm{P}\{X(t+\tau) - X(t) > 0\}$, $\tau \geqslant 0$. Показать, что $Q(\xi + \eta) = Q(\xi) + Q(\eta) - Q(\xi)Q(\eta)$.

18.613. Случайный процесс $X(t)$ есть величина интервала времени между двумя последовательными скачками пуассоновского процесса $N(t)$ с параметром λ. Найти одномерную плотность случайного процесса $X(t)$.

18.614*. Число отказов радиоэлектронной аппаратуры представляет собой пуассоновский поток с интенсивностью $5 \cdot 10^{-4}$ отказа в час. Найти вероятность безотказной работы аппаратуры в течение 200 часов, а также математическое ожидание и дисперсию времени безотказной работы аппаратуры.

18.615. Случайный процесс $X^{(k)}(t)$ есть величина интервала времени между n-м скачком пуассоновского процесса с параметром λ, зарегистрированным в момент времени t и $(n+k)$-м скачком того же процесса. Найти одномерную плотность случайного процесса $X^{(k)}(t)$. Какому закону распределения соответствует полученная плотность?

Каноническим разложением действительной случайной функции $X(t)$ называется ее представление в виде

$$X(t) = m_x(t) + \sum_{k=1}^{n} V_k \varphi_k(t), \tag{3}$$

где $\varphi_k(t)$ $(k = 1, 2, \ldots)$ — неслучайные действительные так называемые «*координатные*» *функции*, V_k $(k = 1, 2, \ldots)$ — центрированные попарно некоррелированные случайные величины с дисперсиями D_k $(k = 1, 2, \ldots, n)$.

Если случайная функция $X(t)$ представлена каноническим разложением (3), то ее автоковариационная функция записывается в виде

$$K_x(t_1, t_2) = \sum_{k=1}^{n} D_k \varphi_k(t_1)\, \varphi_k(t_2). \tag{4}$$

Выражение (4) называется также *каноническим разложением автоковариационной функции.*

Если в представлении случайной функции

$$X(t) = \sum_{k=1}^{n} U_k \psi_k(t) \tag{5}$$

случайные величины U_k коррелированы, то такое представление не является каноническим разложением, и поэтому представление (4) для автоковариационной функции будет несправедливо. Однако с помощью линейного преобразования можно привести выражение (5) к каноническому виду.

В основных чертах указанная задача аналогична приведению билинейной (или квадратичной) формы к каноническому виду, рассматриваемому в курсе линейной алгебры.

Пример 3. Случайный процесс $X(t)$ задан выражением (5), где $\mathrm{M}[U_k] = m_k \neq 0$, $k = 1, 2, \ldots, n$,

$$K = \begin{pmatrix} D_1 & K_{12} & \ldots & K_{1n} \\ & D_2 & \ldots & K_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ & & & D_n \end{pmatrix} \text{ — автоковариационная матрица,}$$

причем $K_{ij} = \mathrm{M}[\mathring{U}_i \mathring{U}_j] \neq 0$ для $i \neq j$.

Очевидно, представление (5) не является каноническим разложением. Требуется с помощью линейного преобразования привести выражение (5) к каноническому виду.

◁ Прежде всего центрируем случайные величины U_k с помощью тождественного преобразования:

$$X(t) = \sum_{k=1}^{n} m_k \psi_k(t) + \sum_{k=1}^{n} (U_k - m_k)\, \psi_k(t) = m_x(t) + \sum_{k=1}^{n} \mathring{U}_k \psi_k(t).$$

Заметим, что данное выражение может быть записано в виде скалярного произведения

$$\mathring{X}(t) = \mathring{U}^{\mathsf{T}} \psi(t), \tag{6}$$

где $\mathring{U}$ и $\psi(t)$ — векторы-столбцы, а символ T означает транспонирование.

В векторном обозначении автоковариационная матрица записывается следующим образом:

$$K_U = \mathrm{M}\,[\mathring{U}\mathring{U}^{\top}].$$

Пусть $V = AU$ — новый вектор случайных величин. Выберем матрицу A таким образом, чтобы случайные компоненты V_k вектора V были попарно некоррелированы. Имеем

$$K_V = \mathrm{M}\,[VV^{\top}] = \mathrm{M}\,[A\mathring{U}\mathring{U}^{\top}A^{\top}] = A\mathrm{M}\,[\mathring{U}\mathring{U}^{\top}]\,A^{\top} = AK_UA^{\top} = D_V, \qquad (7)$$

где D_V — диагональная матрица дисперсий новых компонент:

$$D_V = \begin{pmatrix} \mathrm{D}\,[V_1] & & & \\ & \mathrm{D}\,[V_2] & & \\ & & \ddots & \\ & & & \mathrm{D}\,[V_n] \end{pmatrix}.$$

(При выводе (7) использовано свойство 1* (§ 4, п. 1) математического ожидания.)

Из равенств (7) заключаем, что матрица A преобразования «координат» приводит автоковариационную матрицу K_U к диагональному виду. Поскольку матрица K_U вещественная симметрическая (свойство 1 автоковариационной функции), то, как доказывается в курсе линейной алгебры, искомая матрица

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

строится следующим образом: k-я строка матрицы A представляет собой k-й ортонормированный собственный вектор матрицы K_U, соответствующий собственному значению λ_k. Собственные значения λ_k ($k = 1, 2, \ldots, n$) матрицы K_U являются корнями уравнения

$$\det\,(K_U - \lambda I) = 0,$$

где I — единичная матрица, причем все $\lambda_k \geqslant 0$ ($k = 1, 2, \ldots, n$), поскольку матрица K_U вещественная симметрическая. Если указанная матрица A уже построена, то

$$D_V = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

причем порядок расположения собственных значений λ_k соответствует порядку записи собственных векторов a_k в матрице A. Остается выбрать новый вектор координатных функций $\varphi(t)$ из условия неизменности скалярного произведения (6). Учитывая, что ортогональная матрица неособенная, имеем

$$\overset{\circ}{X}(t) = \overset{\circ}{U}^{\top}\psi(t) = (A^{-1}A\overset{\circ}{U})^{\top}\psi(t) = (A\overset{\circ}{U})^{\top}(A^{-1})^{\top}\psi(t) = V^{\top}\varphi(t),$$

где $V = A\overset{\circ}{U}$ — преобразованный вектор, а

$$\varphi(t) = (A^{-1})^{\top}\psi(t) = (A^{\top})^{\top}\psi(t) = A\psi(t). \tag{8}$$

Здесь учтено, что $A^{-1} = A^{\top}$ в силу ортогональности матрицы A. Заметим, что задача имеет не единственное решение. Результат зависит от способа формирования ортогональной матрицы A. Кроме того, можно положить $\mathrm{D}[V_k] = 1$ $(k = 1, 2, \ldots, n)$, если выбрать новые координатные функции в виде $\varphi'_k(t) = \sqrt{\lambda_k}\,\varphi_k(t)$. $\triangleright$

Пример 4. Случайная функция $X(t)$ задана выражением

$$X(t) = U_1 t + U_2 \cos t, \quad \mathrm{M}[U_1] = 1, \quad \mathrm{M}[U_2] = -1,$$

$$K_U = \begin{pmatrix} 9 & -12 \\ -12 & 25 \end{pmatrix}.$$

Привести данную случайную функцию к каноническому виду.

$\triangleleft$ Центрируем функцию: $X(t) = t - \cos t + \overset{\circ}{U}_1 t + \overset{\circ}{U}_2 \cos t = m_x(t) + (\overset{\circ}{U}^{\top}\psi(t))$, где $\overset{\circ}{U} = (\overset{\circ}{U}_1, \overset{\circ}{U}_2)^{\top}$, $\psi(t) = (t, \cos t)^{\top}$. В данном примере удобнее воспользоваться не ортогональным преобразованием, а методом Лагранжа (аналогичен методу Лагранжа приведения билинейных форм к каноническому виду). Для этого тождественно преобразуем выражение для ковариационной функции:

$$K_x(t_1, t_2) = \sigma_1^2 t_1 t_2 + \sigma_1\sigma_2\rho t_1 \cos t_2 + \sigma_1\sigma_2 t_2 \cos t_1 + \sigma_2^2 \cos t_1 \cos t_2 =$$
$$= (\sigma_1 t_1 + \sigma_2\rho \cos t_1)(\sigma_1 t_2 + \sigma_2\rho \cos t_2) + \sigma_2^2(1 - \rho^2) \cos t_1 \cos t_2.$$

Заметим, что $\rho = \dfrac{K_{12}}{\sigma_1\sigma_2} = -\dfrac{4}{5}$. Обозначим

$$\begin{aligned} \varphi_1(t) &= \sigma_1 t + \sigma_2\rho \cos t, \\ \varphi_2(t) &= \sigma_2\sqrt{1-\rho^2}\,\cos t \end{aligned} \tag{9}$$

и положим $\mathrm{D}[V_1] = \mathrm{D}[V_2] = 1$. Матрица преобразования A, усматриваемая из системы (9), имеет вид

$$A = \begin{pmatrix} \sigma_1 & \sigma_2\rho \\ 0 & \sigma_2\sqrt{1-\rho^2} \end{pmatrix} = \begin{pmatrix} 3 & -4 \\ 0 & 3 \end{pmatrix}.$$

Однако в отличие от случая, рассмотренного в примере 3, матрица A не ортогональна. Потребуем, чтобы скалярное произведение $\mathring{U}\psi$ не изменялось при линейном преобразовании. Тогда аналогично формулам (8) получим, что если $\varphi(t) = A\psi(t)$, то $V = (A)^{-1}\mathring{U}$. Вычисляя обратную матрицу, находим

$$(A^{\top})^{-1} = \frac{1}{9}\begin{pmatrix} 3 & 0 \\ 4 & 3 \end{pmatrix};$$

таким образом, новый вектор случайных величин

$$V = \frac{1}{9}\begin{pmatrix} 3 & 0 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} \mathring{U}_1 \\ \mathring{U}_2 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 3\mathring{U}_1 \\ 4\mathring{U}_1 + 3\mathring{U}_2 \end{pmatrix}. \tag{10}$$

Проверим некоррелированность V_1 и V_2:

$$\mathrm{M}[V_1V_2] = \frac{1}{81}\left(12\mathrm{M}[\mathring{U}_1^2] + 9\mathrm{M}[\mathring{U}_1\mathring{U}_2]\right) = 0.$$

Таким образом, получаем каноническое разложение

$$X(t) = t - \cos t + V_1\varphi_1(t) + V_2\varphi_2(t),$$

где V_1 и V_2 определяются уравнениями (10), а координатные функции $\varphi_1(t)$ и $\varphi_2(t)$ — уравнениями (9). ▷

18.616*. Случайный процесс $X(t)$ задан следующим выражением: $X(t) = U^{\top}\psi(t)$, где $U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$ — случайный вектор с вектором математических ожиданий $m_U = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ и ковариационной матрицей $K_U = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$; вектор координатных функций $\psi = \begin{pmatrix} \psi_1(t) \\ \psi_2(t) \end{pmatrix}$. Найти каноническое разложение процесса $X(t)$ и записать автоковариационную функцию.

18.617*. Случайный процесс $X(t)$ задан представлением $X(t) = t + U_1\cos t + U_2\sin t$, где $\mathrm{M}[U_1] = 1$, $\mathrm{M}[U_2] = 2$, $K_U = \begin{pmatrix} 25 & 0{,}6 \\ 0{,}6 & 36 \end{pmatrix}$. Найти канонические разложения процесса и автоковариационной функции.

18.618*. Случайный процесс задан выражением

$$X(t) = m_X(t) + \sum_{k=1}^{3} U_k\psi_k(t),$$

где U_k — центрированные случайные величины с ковариационной матрицей

$$K_U = \begin{pmatrix} 4 & 4 & 0 \\ 4 & 8 & -4 \\ 0 & -4 & 4 \end{pmatrix}.$$

Найти канонические разложения процесса $X(t)$ и его автоковариационной функции.

2. Дифференцирование и интегрирование случайных функций. Говорят, что *случайная функция $X(t)$ сходится в среднеквадратичном при $t \to t_0$ к случайной величине X_0* (краткое обозначение: $X_0 = \underset{t\to t_0}{\mathrm{l.i.m.}}\, X(t)$), если

$$\lim_{t\to t_0} \mathrm{M}\left\{[X(t) - X_0]^2\right\} = 0.$$

Случайная функция $X(t)$ называется *непрерывной в среднеквадратичном в точке t_0*, если

$$\underset{t\to t_0}{\mathrm{l.i.m.}}\, X(t) = X_0 = X(t_0).$$

Для непрерывности в среднеквадратичном случайной функции $X(t)$ на интервале $t \in (a, b)$ необходимо и достаточно, чтобы $m_x(t)$ было непрерывно при $t \in (a, b)$, а автоковариационная функция $K_x(t_1, t_2)$ была непрерывна на диагонали $t_2 = t_1 = t \in (a, b)$.

Из сходимости в среднеквадратичном следует сходимость по вероятности. Обратное неверно.

Производной случайной функции $X(t)$ называется случайная функция $Z(t) = \dfrac{dX(t)}{dt}$, определяемая как предел в среднеквадратичном отношения приращения случайной функции к приращению неслучайного аргумента:

$$Z(t) = \frac{dX(t)}{dt} = \underset{\Delta t\to 0}{\mathrm{l.i.m.}}\, \frac{X(t+\Delta t) - X(t)}{\Delta t}. \tag{11}$$

Для дифференцируемости случайной функции $X(t)$ на интервале (a, b) необходимо и достаточно, чтобы математическое ожидание $m_x(t)$ было дифференцируемой функцией на (a, b) и ковариационная функция $K_x((t_1, t_2)$ имела вторую смешанную производную по t_1 и t_2 на диагонали $t_1 = t_2 = t \in (a, b)$.

Интегралом в среднеквадратичном от случайной функции $X(t)$ в пределах от a до b называется величина

$$\int_a^b X(t)\,dt = \underset{\substack{n\to\infty \\ \max\limits_k |\Delta s_k|\to 0}}{\mathrm{l.i.m.}} \sum_{k=1}^{n} X(s_k)\,\Delta s_k,$$

где $s_0 = a$, s_1, s_2, ... , $s_n = b$ — точки деления отрезка (a, b), $\Delta s_k = s_k - s_{k-1}$. Если один из пределов интегрирования переменный, то результатом интегрирования будет случайная функция этого переменного предела. Например,

$$Y(t) = \int_a^t X(s)\,ds = \underset{\substack{n \to \infty \\ \max\limits_k |\Delta s_k| \to 0 \\ s_0 = a,\ s_n = t}}{\mathrm{l.i.m.}} \sum_{k=1}^{n} X(s_k)\,\Delta s_k. \tag{12}$$

Если $m_x(t)$ и $K_x(t_1, t_2)$ — соответственно математическое ожидание и автоковариационная функция случайного процесса $X(t)$, то

$$\mathrm{M}\left[\frac{dX(t)}{dt}\right] = \frac{dm_x(t)}{dt}, \quad \mathrm{M}\left[\int_a^t X(s)\,ds\right] = \int_a^t m_x(s)\,ds;$$

$$K_z(t_1, t_2) = \frac{\partial}{\partial t_1}\frac{\partial}{\partial t_2} K_x(t_1, t_2);$$

$$K_Y(t_1, t_2) = \int_a^{t_1} ds_1 \int_a^{t_2} K_x(s_1, s_2)\,ds_2,$$

где случайные функции $Z(t)$ и $Y(t)$ определяются соответственно формулами (11) и (12).

Если $X(t)$ — каноническое разложение вида (3), то вопрос о дифференцируемости или интегрируемости канонического разложения по существу сводится к вопросу о дифференцируемости или интегрируемости координатных функций и математического ожидания в обычном смысле.

Приведенные выше выражения для автоковариационной функции могут быть записаны универсальным образом в следующих обозначениях.

Пусть $Y(t) = L_t^0 X(t)$, где L_t^0 — линейный однородный оператор (это значит, что L_t^0 удовлетворяет двум условиям:

$$L_t^0(x_1(t) + x_2(t)) = L_t^0 x_1(t) + L_t^0 x_2(t) \quad \text{(аддитивность)},$$

$$L_t^0(cx_1(t)) = cL_t^0 x_1(t) \quad \text{(однородность)},$$

где $x_1(t)$ и $x_2(t)$ — любые неслучайные функции (реализации случайного процесса), а c — неслучайная константа). Тогда

$$m_Y(t) = L_t^0 m_x(t), \quad K_Y(t_1, t_2) = L_{t_1}^0 L_{t_2}^0 K_x(t_1, t_2). \tag{13}$$

Пример 5. На RC-цепочку, изображенную на рис. 16, подается случайное напряжение $X(t)$ с характеристиками $m_X(t) = 2t$ и $K_X(t_1, t_2) = 4t_1 t_2$. Найти математическое ожидание, автоковариационную функцию и дисперсию напряжения $Y(t)$ на выходе RC-цепочки.

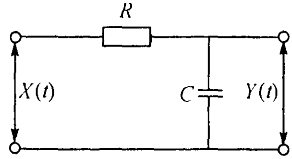


Рис. 16

◁ Дифференциальное уравнение, описывающее связь между входным и выходным напряжениями RC-цепочки, составляется на основе законов Кирхгофа и имеет вид

$$\frac{dy(t)}{dt} + \beta y(t) = \beta x(x), \quad \beta = \frac{1}{RC} > 0.$$

Решая данное дифференциальное уравнение при нулевом начальном условии методом вариации постоянной, получим

$$y(t) = \beta e^{-\beta t} \int_0^t e^{\beta\tau} x(\tau)\, d\tau.$$

Таким образом, случайный процесс $Y(t)$ является результатом применения к случайному процессу $X(t)$ линейного однородного оператора $L_t^0 x = \beta e^{-\beta t} \int_0^t e^{\beta\tau} x\, d\tau$. Поэтому согласно формулам (13)

$$m_Y(t) = \beta e^{-\beta t} \int_0^t e^{\beta\tau} m_X(\tau)\, d\tau = \beta e^{-\beta t} \cdot 2 \int_0^t \tau e^{\beta\tau}\, d\tau = 2\left[t - \frac{1}{\beta}\left(1 - e^{-\beta t}\right)\right];$$

$$K_Y(t_1, t_2) = 4\beta^2 e^{-\beta(t_1+t_2)} \int_0^{t_1} e^{\beta s_1} s_1\, ds_1 \int_0^{t_2} e^{\beta s_2} s_2\, ds_2 = 4\left[t_1 - \frac{1}{\beta}\left(1 - e^{-\beta t_1}\right)\right]\left[t_2 - \frac{1}{\beta}\left(1 - e^{-\beta t_2}\right)\right].$$

Дисперсия процесса на выходе

$$D_Y(t) = K_Y(t, t) = 4\left[t - \frac{1}{\beta}\left(1 - e^{-\beta t}\right)\right]^2. \ ▷$$

18.619. Случайная функция задана каноническим разложением $X(t) = t + V_1 \cos t + V_2 \sin t$, $\quad \mathrm{D}[V_1] = 1$, $\quad \mathrm{D}[V_2] = 2$.

Вычислить математическое ожидание, дисперсию и ковариационную функцию процесса $Z(t) = \dfrac{dX(t)}{dt}$.

18.620 (продолжение). В условиях предыдущей задачи вычислить математическое ожидание, дисперсию и ковариационную функцию процесса $Y(t) = \int\limits_0^t X(s)\,ds$.

18.621. Случайный процесс $X(t)$ задан каноническим разложением $X(t) = 1 + Ut + Vt^2$ с характеристиками $D_U = 3$, $D_V = 1$. Найти математическое ожидание и ковариационную функцию производной данного процесса.

18.622 (продолжение). В условиях предыдущей задачи найти математическое ожидание и дисперсию процесса $Y(t) = \int\limits_0^t X(s)\,ds$.

18.623*. Исследовать вопрос о непрерывности и дифференцируемости в среднеквадратичном пуассоновского процесса, определенного в примере 2 и имеющего характеристики, полученные при решении задач 18.610, 18.612.

18.624. На вход дифференцирующего устройства поступает случайный процесс с математическим ожиданием $m_X(t) = 3t^2 + t$ и автоковариационной функцией

$$K_X(t_1, t_2) = \sigma^2 e^{-\alpha|t_1 - t_2|}(1 + \alpha|t_1 - t_2|).$$

Дифференцируем ли данный процесс в среднеквадратичном? Найти дисперсию процесса на выходе дифференцирующего устройства.

18.625*. Случайная функция $X(t)$ задана выражением

$$X(t) = V\cos\omega t,$$

где V — случайная величина с характеристиками $m_V = 2$, $\sigma_V = 3$. Найти характеристики случайной функции

$$Y(t) = X(t) + 3\,\frac{dX(t)}{dt}.$$

18.626. Ковариационная функция случайного процесса $X(t)$ имеет вид

$$K_X(t_1, t_2) = D_X\, e^{-\alpha(t_2 - t_1)^2}\cos\beta(t_2 - t_1), \quad \alpha > 0,\ \beta > 0.$$

Определить дисперсию производной процесса $X(t)$.

18.627. Известны характеристики случайного процесса:

$$m_X(t) = 3t^2 + 2t + 1, \quad K_X(t_1, t_2) = 2e^{-(t_2-t_1)^2}.$$

Найти математическое ожидание и дисперсию процесса

$$Y(t) = t\frac{dX(t)}{dt} + t^2.$$

18.628. На вход интегратора, работающего по принципу $y(t) = \int\limits_0^t x(s)\,ds$, где $x(s)$ — произвольная реализация случайного процесса на входе, поступает случайный процесс $X(t)$ с автоковариационной функцией $K_X(t_1, t_2) = 4\cos\omega t_1 \cos\omega t_2$ и математическим ожиданием $m_X(t) = 1 + \sin^2\omega t$. Найти математическое ожидание и дисперсию случайного процесса $Y(t)$ на выходе интегратора.

18.629. На вход интегратора, описанного в предыдущей задаче, поступает случайная ступенька (см. задачу 18.601) со следующими характеристиками: A распределена по закону $N(0, \sigma)$, T — по закону $\mathrm{Ex}(\lambda)$. Найти дисперсию процесса $Y(t)$ на выходе интегратора в произвольный момент времени $t > 0$.

18.630. Задана ковариационная функция $K_X(t_1, t_2)$ случайного процесса $X(t)$. Показать, что взаимная ковариационная функция случайных процессов $X(t)$ и $Y(t) = \int\limits_0^t X(s)\,ds$ может быть представлена в виде $K_{XY}(t_1, t_2) = \int\limits_0^{t_2} K_X(t_1, s)\,ds$, а взаимная ковариационная функция процессов $X(t)$ и $Z(t) = \dfrac{dX(t)}{dt}$ в виде $K_{XZ}(t_1, t_2) = \dfrac{\partial}{\partial t_2}K_X(t_1, t_2)$.

18.631. Ковариационная функция случайного процесса $X(t)$ задана в виде $K_X(t_1, t_2) = t_1^2 + t_1 t_2 + 2t_2^2$. Найти взаимную ковариационную функцию $K_{XY}(t_1, t_2)$ случайных процессов $X(t)$ и $Y(t) = \int\limits_0^t X(s)\,ds$.

18.632. Ковариационная функция случайного процесса $X(t)$ задана так же, как в задаче 18.631. Найти взаимную ковариационную функцию процессов $X(t)$ и $Z(t) = \dfrac{dX(t)}{dt}$.

18.633. Задана автоковариационная функция $K_X(t_1, t_2)$ дважды дифференцируемого случайного процесса $X(t)$. Найти ковариационную функцию связи между процессами $X(t)$ и $Y(t)$, если $Y(t) = \varphi(t)\,X(t) + \psi(t)\,X''(t)$, где $\varphi(t)$ и $\psi(t)$ — заданные неслучайные функции времени.

18.634. $X(t)$ — гауссовский процесс с характеристиками

$$m_X(t) = 1, \quad K_X(t_1, t_2) = \frac{1}{1 + (t_2 - t_1)^2}.$$

Записать плотность распределения вероятностей случайного вектора $(X(t), X'(t))$.

3. Стационарные случайные функции. Случайная функция $X(t)$ называется *строго стационарной* (*стационарной в узком смысле*), если ее n-мерные законы распределения инвариантны относительно сдвига во времени на произвольную величину τ. Например, если случайная функция непрерывного типа $X(t)$ строго стационарна, то

$$f_n(x_1, x_2, \ldots, x_n/t_1 + \tau, t_2 + \tau, \ldots, t_n + \tau) = f_n(x_1, \ldots, x_n/t_1, \ldots, t_n).$$

Если случайная функция строго стационарна, то ее математическое ожидание и дисперсия не зависят от времени, а автоковариационная функция $K_X(t, t+\tau)$ зависит лишь от разности моментов времени, т.е. от τ. Обратное заключение в общем случае неверно. Случайная функция $X(t)$ называется *стационарной в широком смысле*, если ее математическое ожидание и дисперсия не зависят от t, а автоковариационная функция $K_X(t, t+\tau)$ зависит лишь от разности аргументов $K_X(t, t+\tau) = K_X(\tau)$. Для нормальных случайных процессов оба понятия стационарности совпадают. Автоковариационная функция стационарной в широком смысле случайной функции обладает следующими свойствами (следствия общих свойств автоковариационной функции):

1) $K_X(\tau) = K_X(-\tau)$, $K_X(0) \geqslant 0$.

2) $|K_X(\tau)| \leqslant K_X(0) = D_X$.

3) Функция $K_X(t_1 - t_2)$ неотрицательно определенная. Если $X(t)$ — стационарная в широком смысле дифференцируемая случайная функция и $Y(t) = \dfrac{dX(t)}{dt}$, то $Y(t)$ — также стационарная в широком смысле случайная функция, причем

$$m_Y = 0, \quad K_Y(\tau) = -\frac{d^2}{d\tau^2} K_X(\tau).$$

Для дифференцируемости стационарной в широком смысле случайной функции необходимо и достаточно существование второй производной автоковариационной функции при $\tau = 0$. Так как для стационарной в широком смысле случайной функции $X(t)$ имеем

$$\mathrm{D}\left[\frac{dX(t)}{dt}\right] = K_Y(0) = -\frac{d^2}{d\tau^2} K_X(\tau)\Big|_{\tau=0},$$

то условие дифференцируемости фактически равносильно условию конечности дисперсии производной от стационарной случайной функции.

Средним по конечному промежутку времени от реализации стационарного случайного процесса $X(t)$ называется число (вообще говоря, случайное), определяемое соотношением

$$\langle x(t)\rangle_0^T = \frac{1}{T}\int_0^T x(t)\,dt.$$

Стационарный случайный процесс $X(t)$ называется *эргодическим относительно математического ожидания*, если выполняется условие

$$\lim_{T\to\infty}\langle x(t)\rangle_0^T = m_X = \int_{-\infty}^{+\infty} x f_1(x/t)\,dx,$$

где предел в левой части понимается в обычном смысле (т.е. если среднее по бесконечному промежутку времени от одной реализации (любой) равно среднему по множеству реализаций (среднему по ансамблю)). Аналогично определяется эргодичность относительно автоковариационной функции.

Стационарный случайный процесс называется *эргодическим относительно автоковариационной функции*, если выполняется условие

$$\lim_{T\to\infty}\langle \mathring{x}(t)\,\mathring{x}(t+\tau)\rangle_0^T = K_X(\tau) =$$
$$= \int_{-\infty}^{+\infty}(x_1 - m_X)(x_2 - m_X)\,f_2(x_1, x_2/t, t+\tau)\,dx_1\,dx_2.$$

Две стационарные случайные функции $X(t)$ и $Y(t)$ называются *стационарно связанными*, если их взаимная ковариационная функция зависит лишь от разности аргументов.

Пример 6. Случайный процесс $X(t)$ представляет собой *случайное гармоническое колебание*

$$X(t) = A\cos(\omega t + \Phi),$$

где $A > 0$ — случайная амплитуда с плотностью распределения вероятностей $f_A(a)$ (равна нулю при $a < 0$) такая, что существует второй начальный момент $\mathrm{M}[A^2]$, Φ — независимая от A случайная фаза колебания, равномерно распределенная на отрезке $[-\pi, \pi]$. Найти автоковариационную функцию процесса и дать ответы на следующие вопросы:

1) Является ли данный процесс стационарным в широком смысле?

2) Дифференцируем ли данный процесс один раз?

3) Является ли он эргодическим относительно математического ожидания?

◁ 1) Запишем $X(t)$ следующим образом:

$$X(t) = A\cos\Phi\,\cos\omega t - A\sin\Phi\,\sin\omega t = V_1\cos\omega t - V_2\sin\omega t. \tag{14}$$

Учитывая независимость A и Φ, имеем

$$\mathrm{M}[V_1] = \mathrm{M}[A\cos\Phi] = \mathrm{M}[A]\,\mathrm{M}[\cos\Phi] = m_A\,\frac{1}{2\pi}\int\limits_{-\pi}^{\pi}\cos\varphi\,d\varphi = 0,$$

где $m_A = \displaystyle\int\limits_{-\infty}^{\infty} a f_A(a)\,da.$

Аналогично устанавливается, что и $\mathrm{M}[V_2] = 0$. Найдем ковариацию случайных величин V_1 и V_2:

$$K_{12} = \mathrm{M}[V_1V_2] = \mathrm{M}[A^2\cos\Phi\,\sin\Phi] = \mathrm{M}[A^2]\cdot\frac{1}{2}\,\mathrm{M}[\sin 2\Phi] = 0.$$

Поэтому выражение (14) является каноническим разложением. Следовательно, по формуле (4)

$$K_X(t_1, t_2) = \mathrm{D}[V_1]\,\cos\omega t_1\cos\omega t_2 + \mathrm{D}[V_2]\,\sin\omega t_1\sin\omega t_2.$$

Применяя свойство 7) дисперсии произведения независимых величин, можем написать

$$\mathrm{D}[V_1] = \mathrm{D}[A\cos\Phi] = \mathrm{D}[A]\mathrm{D}[\cos\Phi] + m_A^2\mathrm{D}[\cos\Phi] + \mathrm{M}^2[\cos\Phi]\,\mathrm{D}[A] =$$
$$= \mathrm{D}[\cos\Phi](\mathrm{D}[A] + m_A^2) = \mathrm{M}[A^2]\mathrm{D}[\cos\Phi].$$

Но

$$\mathrm{D}[\cos\Phi] = \mathrm{M}[\cos^2\Phi] = \frac{1}{2\pi}\int\limits_{-\pi}^{\pi}\cos^2\varphi\,d\varphi = \frac{1}{2}$$

и

$$\mathrm{M}[A^2] = \int\limits_{-\infty}^{+\infty} a^2 f_A(a)\,da.$$

Таким образом, $D[V_1] = \frac{1}{2}M[A^2]$. Аналогично получается $D[V_2] = \frac{1}{2}M[A^2]$. Поэтому автоковариационная функция $K_X(t_1, t_2) = \frac{1}{2}M[A^2]\cos\omega(t_2 - t_1)$ зависит лишь от разности моментов времени. Кроме того, $M[X(t)] = 0$ в силу доказанной центрированности случайных величин V_1 и V_2. Отсюда следует, что процесс $X(t)$ стационарный в широком смысле.

2) Проверим достаточные условия дифференцируемости:

$$K''_X(\tau)_{\tau=0} = -\frac{M[A^2]}{2}\omega^2\cos\omega\tau\Big|_{\tau=0} = -\frac{M[A^2]}{2}\omega^2$$

— конечная величина, поэтому процесс дифференцируем.

3) Вычислим среднее по конечному промежутку времени от одной реализации:

$$\langle x(x)\rangle_0^T = \frac{1}{T}\int\limits_0^T a_1\cos(\omega t + \varphi_1)\,dt = \frac{a_1}{\omega T}(\sin(\omega T + \varphi_1) + \sin\varphi_1).$$

Отсюда

$$\lim_{T\to\infty}\langle x(x)\rangle_0^T = 0 = m_x,$$

поэтому процесс $X(t)$ эргодический относительно математического ожидания.

18.635. Является ли пуассоновский процесс $X(t)$ с параметром λ стационарным в широком смысле?

18.636*. Случайный процесс $X(t)$ задан следующим образом:

$$X(t) = V_1\cos\omega t + V_2\sin\omega t, \quad \omega > 0, \quad t \in \mathbb{R},$$

где V_1 и V_2 — независимые случайные величины, принимающие с равной вероятностью лишь два возможных значения $+1$ и -1. Найти автоковариационную функцию процесса и убедиться, что $X(t)$ стационарен в широком смысле, но не является стационарным в узком смысле.

18.637*. Найти математическое ожидание и ковариационную функцию процесса $Y(t) = X(t+1) - X(t)$, если $X(t)$ — пуассоновский процесс с параметром λ. Является ли $Y(t)$ стационарным в широком смысле?

18.638. Задана автоковариационная функция $K_X(\tau)$ стационарного случайного процесса $X(t)$. Вычислить среднее значение квадрата приращения процесса за время τ.

18.639. Случайный процесс $X(t)$ задан следующим каноническим разложением:

$$X(t) = m_X + \sum_{k=0}^{n} U_k \cos \omega_k t + V_k \sin \omega_k t,$$

где U_k и V_k — центрированные некоррелированные случайные величины, причем $\mathrm{M}[U_i U_j] = \mathrm{M}[V_i V_j] = D_i \delta_{ij}$, $\mathrm{M}[U_i V_j] = 0$ для всех $i, j = 0, 1, \ldots, n$. Показать, что данный процесс является стационарным в широком смысле, и найти автоковариационную функцию и дисперсию процесса.

18.640. Случайный процесс $X(t)$ представляет собой суперпозицию гармонических колебаний вида

$$X(t) = \sum_{k=0}^{n} A_k \cos(\omega_k t + \Phi_k),$$

где $\{A_k\}$, $k = 0, 1, 2, \ldots, n$, — центрированные некоррелированные случайные амплитуды ($\mathrm{M}[A_i A_j] = D_i \delta_{ij}$), $\{\Phi_k\}$ — независимые от любой из случайных величин A_i случайные фазы, распределенные по одному и тому же закону с четной плотностью $f_\Phi(\varphi)$ и некоррелированные между собой. Вычислить дисперсии процессов $X(t)$ и $\dfrac{dX(t)}{dt}$.

18.641 (продолжение). В условиях предыдущей задачи вычислить дисперсию процесса $\displaystyle\int_0^t X(s)\,ds$.

18.642. Показать, что два случайных гармонических колебания $X(t) = A\cos(\omega t + \Phi)$ и $Y(t) = B\cos(\omega t + \Phi)$, где A и B — две случайные амплитуды с ковариацией K, а Φ — независимая от A и B случайная фаза, равномерно распределенная на отрезке $[-\pi, \pi]$, являются стационарно связанными.

18.643. Стационарный случайный процесс $X(t)$ представляет собой случайное гармоническое колебание $X(t) = A\cos(\omega t + \Phi)$, описанное в примере 6. Является ли данный процесс эргодическим относительно автоковариационной функции?

18.644. Является ли случайное гармоническое колебание $X(t) = a\cos(\omega t + \Phi)$, где a — постоянная амплитуда, а Φ — равномерно распределенная на отрезке $[-\pi, \pi]$ случайная величина, эргодическим процессом относительно автоковариационной функции?

18.645. $X(t)$ представляет собой *случайный импульсный сигнал*, реализации которого строятся следующим образом. В некоторый случайный момент времени T_1 появляется прямоугольный импульс длительностью t_0 со случайной амплитудой A_1, распределенной независимо от T_1 с характеристиками $m_{A_1} = 0$, $\sigma_{A_1} = \sigma$. В момент времени $T_1 + t_0$ данный импульс заканчивается, и появляется следующий импульс длительностью t_0 со случайной амплитудой A_2, распределенной одинаково с A_1, но независимо от A_1 и T_1, и т.д. Одна из реализаций данного процесса изображена на рис. 17 (при этом считаем, что начало процесса (момент T_1) в идеальной модели импульсного сигнала находится в минус бесконечности).

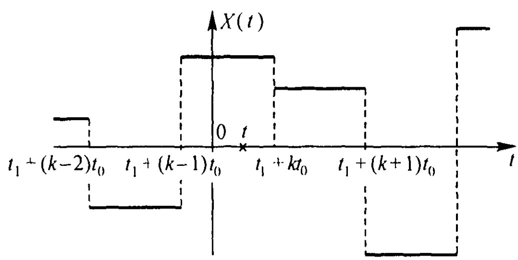


Рис. 17

Найти характеристики процесса: $m_X(t)$ и $K_X(t_1, t_2)$. Является ли данный процесс стационарным в широком смысле?

◁ Зафиксируем момент времени наблюдения t. Пусть $t \in (t_k,\, t_k + t_0)$, где t_k — момент начала k-го импульса. При этих значениях t наблюдается случайная величина A_k, которая по условию центрирована для любых $k \in N$. Поэтому $\mathrm{M}[X(t)] = 0$. Пусть $t_2 = t + \tau$ $(\tau \geqslant 0)$. В силу произвольности момента наблюдения t и случайности моментов времени t_k можно считать, что промежуток времени T от момента t до момента окончания k-го импульса $(t_k + t_0)$ является случайной величиной, распределенной по закону $R(0, t_0)$. Поэтому

$$\mathrm{M}[X(t)\,X(t+\tau)] = \begin{cases} \sigma^2 \mathrm{P}\{\tau < T\} + 0 \cdot \mathrm{P}\{\tau \geqslant T\}, & 0 \leqslant \tau < t_0, \\ 0, & \tau \geqslant t_0. \end{cases}$$

В силу равномерности распределения T получаем

$$K_X(t, t+\tau)] = \begin{cases} \sigma^2 \left(1 - \dfrac{\tau}{t_0}\right), & 0 \leqslant \tau < t_0, \\ 0, & \tau \geqslant t_0. \end{cases}$$

Тот факт, что ковариационная функция зависит лишь от разности моментов времени наблюдения, свидетельствует о стационарности процесса $X(t)$ в широком смысле. ▷

18.646*. *Телеграфным сигналом* называется случайный процесс $X(t)$, который с равной вероятностью может принимать лишь два значения $+1$ и -1, причем число перемен знака за время $\tau = t_2 - t_1$ не зависит от предыстории процесса до момента t_1 и представляет собой пуассоновский процесс $N(\tau)$ с параметром λ. Одна из реализаций телеграфного сигнала изображена на рис. 18. Вычислить автоковариационную функцию телеграфного сигнала.

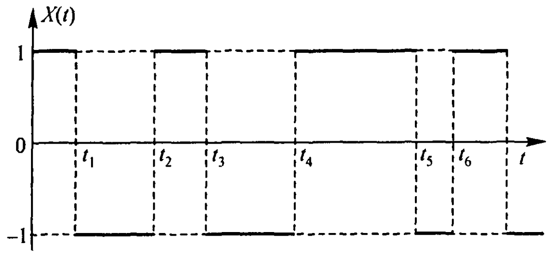


Рис. 18

18.647*. *Фототелеграфным сигналом* называется случайный процесс, реализации которого строятся следующим образом. Значения сигнала изменяются скачком в случайные моменты времени t_k, $k = 0, 1, 2, \ldots$ Число скачков за время $\tau > 0$ представляет собой пуассоновский процесс с параметром λ. В интервале (t_{k-1}, t_k) между двумя скачками $X(t)$ может принимать лишь два значения 0 или 1 с вероятностями $1-p$ и p. Значения $X(t)$ в различных интервалах независимы. Найти автоковариационную функцию сигнала. Является ли фототелеграфный сигнал стационарным в широком смысле?

18.648*. Случайный процесс $X(t)$ задан следующим образом: $X(t)$ изменяет свои значения в случайные моменты времени, причем число «скачков» процесса за время τ представляет собой пуассоновский процесс с параметром λ. В промежутках между каждыми двумя скачками значения процесса $X(t)$ не меняются и представляют собой независимые случайные величины с нулевым математическим ожиданием и одинаковыми дисперсиями σ^2. Найти основные характеристики процесса: $m_X(t)$, $D_X(t)$ и $K_X(t_1, t_2)$.

18.649*. Исследовать вопрос о непрерывности и дифференцируемости в среднеквадратичном телеграфного сигнала, определенного в задаче 18.646.

В задачах 18.650–18.652 заданы автоковариационные функции стационарных случайных процессов. В каждом случае требуется установить, сколько раз дифференцируем данный процесс.

18.650. $K_X(\tau) = \sigma_X^2 e^{-\alpha^2\tau^2}$, $\alpha > 0$.

18.651. $K_X(\tau) = \sigma_X^2 e^{-\alpha|\tau|}\left(1 + \alpha|\tau| + \frac{1}{3}\alpha^2\tau^2\right)$, $\alpha > 0$.

18.652. $K_X(\tau) = \sigma_X^2 e^{-\alpha|\tau|}\left(\cos\beta\tau - \frac{\alpha}{\beta}\sin\beta|\tau|\right)$, $\alpha > 0$, $\beta > 0$.

18.653. На вход радиотехнической цепи, состоящей из последовательно соединненых дифференцирующих устройств и сумматора (рис. 19), поступает случайный стационарный сигнал $X(t)$ с нулевым математическим ожиданием и автоковариационной функцией $K_X(\tau)$. Найти автоковариационную функцию случайного сигнала $Y(t)$ на выходе сумматора.

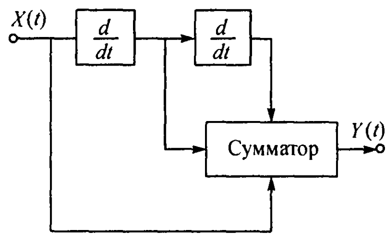


Рис. 19

18.654. Показать, что взаимная ковариационная функция $K_{XY}(t, t')$ стационарной случайной функции $X(t)$ и ее производной $Y(t) = \frac{dX(t)}{dt}$ удовлетворяет условию

$$K_{XY}(t, t') = -K_{XY}(t', t).$$

18.655. Показать, что если стационарный случайный процесс $X(t)$ дифференцируем, то он стационарно связан со своей производной. Найти ковариационную функцию связи $K_{XY}(t_1, t_2)$, где $Y(t) = \frac{dX(t)}{dt}$.

18.656*. Стационарный случайный процесс $X(t)$ имеет автоковариационную функцию

$$K_X(\tau) = \sigma^2 e^{-\alpha|\tau|}\left(\cos\omega\tau + \frac{\alpha}{\beta}\sin\omega|\tau|\right).$$

Найти ковариационную функцию связи процессов

$$X(t) \text{ и } Y(t) = \frac{dX(t)}{dt}.$$

18.657. Известна автоковариационная функция стационарного случайного процесса $X(t)$:

$$K_X(\tau) = \sigma^2 e^{-\alpha|\tau|}\left(1 + \alpha|\tau| + \frac{1}{3}\alpha^2\tau^2\right).$$

Определить ковариационную функцию связи между

$$X(t) \text{ и } Y(t) = \frac{d^2X(t)}{d^2t}.$$

18.658. На вход интегрирующего устройства, работающего по принципу $y(t) = \int\limits_0^t x(s)\,ds$, где $x(t)$ и $y(t)$ — реализации соответственно входного и выходного процессов, поступает стационарный случайный процесс $X(t)$ с автоковариационной функцией $K_X(\tau)$. Показать, что

$$K_{XY}(t_1, t_2) = \int\limits_0^{t_2} K_X(t_1 - \tau)\,d\tau.$$

18.659* (продолжение). В условиях предыдущей задачи показать, что

$$\sigma_Y^2(t) = 2\int\limits_0^t (t-\tau)\,K_X(\tau)\,d\tau.$$

18.660*. Задана автоковариационная функция стационарного случайного процесса $X(t)$: $K_X(\tau) = 4e^{-2|\tau|}$. Найти дисперсию случайной функции $Y(t) = \int\limits_0^t X(s)\,ds$.

18.661*. Стационарный случайный процесс $X(t)$ имеет автоковариационную функцию $K_X(\tau) = \sigma^2 e^{-\alpha|\tau|}$. Выяснить, является ли он стационарно связанным с процессом $Y(t) = \int\limits_0^t X(s)\,ds$.

18.662*. Показать, что если $X(t)$ — нормальный стационарный в широком смысле дифференцируемый случайный процесс, то процесс $Y(t) = \dfrac{dX(t)}{dt}$ также нормальный стационарный в широком смысле. Найти характеристики m_Y, $K_Y(\tau)$ и D_Y, если

$$K_X(\tau) = \sigma_X^2 e^{-\alpha|\tau|}\left(\cos\beta\tau + \frac{\alpha}{\beta}\sin\beta|\tau|\right)$$

(α и β — положительные константы).

18.663. Нормальный стационарный случайный процесс $X(t)$ имеет характеристики

$$m_X = 1, \quad K_X(\tau) = 4e^{-2|\tau|}\left(\cos 3\tau + \frac{2}{3}\sin 3|\tau|\right).$$

Вычислить $\mathrm{P}\left\{\left|\dfrac{dX(t)}{dt}\right| < \sqrt{13}\right\}$.

18.664*. Стационарный нормальный процесс $X(t)$ имеет характеристики $m_X = 0$, $K_X(\tau) = \sigma_X^2 e^{-\alpha^2\tau^2}$ ($\alpha > 0$ — постоянная величина). Найти двумерную плотность совместного распределения вероятностей случайных процессов $X(t)$ и $Y(t) = \dfrac{dX(t)}{dt}$ в один и тот же момент времени.

4. Спектральное разложение стационарных случайных функций. Стационарная в широком смысле случайная функция $X(t)$, заданная во всей области определения $t \in G_t$ каноническим разложением вида

$$X(t) = m_x + \sum_{k=0}^{n} U_k \cos\omega_k t + V_k \sin\omega_k t, \tag{15}$$

где U_t и V_k — центрированные случайные величины, удовлетворяющие условиям

$$\mathrm{M}[V_i V_j] = \mathrm{M}[U_i U_j] = D_i \delta_{ij},$$

$$\mathrm{M}[V_i U_j] = 0 \quad \text{для всех } i, j = 0, 1, \ldots, n,$$

называется *случайной функцией с дискретным спектром.* Автоковариационная функция такого процесса имеет вид (см. задачу 18.639)

$$K_X(\tau) = \sum_{k=0}^{n} D_k \cos\omega_k t. \tag{16}$$

Представления (15) и (16) называются *спектральными разложениями* соответственно *случайного процесса* и *автоковариационной функции.* Дисперсия процесса есть сумма дисперсий отдельных гармоник на частотах ω_k $(k = 0, 1, \ldots, n)$:

$$D_X = K_X(0) = \sum_{k=0}^{n} D_k.$$

Если число слагаемых n бесконечно велико, а частоты ω_k кратны основной частоте $\left(\text{т.е. } \omega_k = k\omega_1 = \dfrac{k\pi}{T}\right)$, то разложение (16) представляет собой, по существу, ряд Фурье по косинусам кратных дуг функции $K_X(\tau)$ на отрезке $[-T, T]$ (в силу четности периодически продолженной на всю ось функции $K_X(\tau)$ ряд содержит только косинусы).

Стационарные случайные функции, рассматриваемые лишь на конечном промежутке $t \in [-T, T]$, всегда могут быть представлены в виде спектральных разложений (15) и (16).

Если автоковариационная функция $K_X(\tau)$ не является периодической, то стационарный случайный процесс $X(t)$ не может быть на всей оси $-\infty < t < \infty$ представлен в виде разложения (15) и (16) и, следовательно, не является при всех действительных t процессом с дискретным спектром.

Стационарная случайная функция $X(t)$ называется *случайной функцией с непрерывным спектром*, если существует такая действительная неотрицательная функция $S_X(\omega)$, определенная на всей оси частот $-\infty < \omega < +\infty$ и называемая *спектральной плотностью*, что справедливы интегральные *формулы Винера–Хинчина*

$$K_X(\tau) = \int_0^{+\infty} S_X(\omega) \cos\omega\tau \, d\omega, \tag{17}$$

$$S_X(\omega) = \frac{2}{\pi} \int_0^{+\infty} K_X(\tau) \cos\omega\tau \, d\tau. \tag{18}$$

Для справедливости представлений (17)–(18) достаточно, чтобы ковариационная функция $K_X(\tau)$ была абсолютно интегрируема на полуоси. Таким образом, автоковариационная функция и спектральная плотность стационарной случайной функции с непрерывным спектром связаны друг с другом *взаимно обратными косинус-преобразованиями Фурье*. Из формул (17)–(18) и свойств ковариационной функции $K_X(\tau)$ вытекает, что $S_X(\omega)$ — четная функция: $S_X(-\omega) = S_X(\omega)$. В силу четности подынтегральных функций в формулах Винера–Хинчина, последние могут быть также записаны в экспоненциальном виде:

$$K_X(\tau) = \frac{1}{2} \int_{-\infty}^{+\infty} S_X(\omega) \, e^{i\omega\tau} \, d\omega, \tag{19}$$

$$S_X(\omega) = \frac{1}{\pi} \int_{-\infty}^{+\infty} K_X(\tau) \, e^{-i\omega\tau} \, d\tau. \tag{20}$$

Как следует из (17) и (19), дисперсия стационарного процесса с непрерывным спектром может быть выражена в виде интеграла от спектральной плотности:

$$D_X = K_X(0) = \int_0^{+\infty} S_X(\omega) \, d\omega = \frac{1}{2} \int_{-\infty}^{+\infty} S_X(\omega) \, d\omega. \tag{21}$$

Условия $S_X(\omega) \geqslant 0$ и $S_X(-\omega) = S_X(\omega)$ для всех действительных ω являются необходимыми условиями стационарности в широком смысле случайного процесса $X(t)$. Полезными характеристиками стационарных случайных функций с непрерывным спектром являются *эффективная ширина спектра* $\Delta\omega$ и *средний интервал корреляции* $\Delta\tau$

(*эффективная длительность автокорреляционной функции*), определяемые следующим образом:

$$\Delta\omega = \frac{1}{\max\limits_{\omega} S_X(\omega)} \int\limits_{-\infty}^{+\infty} S_X(\omega)\,d\omega = \frac{2\sigma_X^2}{\max\limits_{\omega} S_X(\omega)},$$

$$\Delta\tau = 2\int\limits_0^{+\infty} |\,\rho_X(\tau)\,|\,d\tau = \frac{2}{\sigma_X^2}\int\limits_0^{+\infty} |\,K_X(\tau)\,|\,d\tau.$$

Геометрически средний интервал корреляции $\Delta\tau$ (соответственно эффективная ширина спектра $\Delta\omega$) равен основанию прямоугольника с высотой $\rho_X(0) = 1$ $\left(\max\limits_{\omega} S_X(\omega)\right)$, площадь которого равна площади под кривой $|\,\rho_X(\tau)\,|$ при $-\infty < \tau < +\infty$ (площади под кривой $S_X(\omega)$ при $-\infty < \omega < +\infty$). Из этих определений и формулы (20) вытекает, что величина $\Delta\tau$ и $\Delta\omega$ связаны между собой неравенством

$$\Delta\tau\,\Delta\omega \geqslant 2\pi \tag{22}$$

(обычно называемым «соотношением неопределенности»). Смысл соотношения (22) можно кратко выразить в виде следующего правила: *чем у́же ширина спектра стационарного процесса, тем больше интервал корреляции его сечений, и наоборот.*

Пример 7. Автоковариационная функция стационарной случайной функции $X(t)$ задана в виде

$$K_X(\tau) = \sigma_X^2 e^{-\alpha|\tau|}, \quad -\infty < \tau < +\infty,\ \alpha > 0.$$

Найти спектральную плотность $S_X(\omega)$ и эффективные характеристики $\Delta\tau$ и $\Delta\omega$.

◁ При вычислении $S_X(\omega)$ удобнее в данном случае использовать формулу (20):

$$S_X(\omega) = \frac{1}{\pi}\int\limits_{-\infty}^{+\infty} K_X(\tau)\,e^{i\omega\tau}\,d\tau =$$

$$= \frac{1}{\pi}\sigma_X^2\left[\int\limits_0^{+\infty} e^{-\alpha\tau+i\omega\tau}\,d\tau + \int\limits_{-\infty}^{0} e^{\alpha\tau+i\omega\tau}\,d\tau\right] = \frac{\sigma_X^2}{\pi}\,\frac{2\alpha}{\omega^2+\alpha^2}.$$

Пользуясь определениями, находим эффективную ширину спектра

$$\Delta\omega = \frac{2\sigma_X^2}{\max\limits_{\omega} S_X(\omega)} = \frac{2\sigma_X^2}{S_X(0)} = \pi\alpha$$

и средний интервал корреляции

$$\Delta\tau = \frac{2}{\sigma_X^2}\,\sigma_X^2 \int\limits_0^{+\infty} e^{-\alpha\tau}\,d\tau = \frac{2}{\alpha}.$$

Таким образом, в данном примере неравенство (22) обращается в равенство. ▷

Пример 8. *Белым шумом* называется стационарный в широком смысле случайный процесс с постоянной спектральной плотностью на всех частотах $-\infty < \omega < +\infty$. Белый шум физически неосуществим, поскольку его дисперсия в соответствии с формулой (19) бесконечна. Пусть $X\,(t)$ — стационарный в широком смысле процесс со спектральной плотностью $S_X(\omega)$ следующего вида:

$$S_X(\omega) = \begin{cases} \dfrac{D_X}{\omega_0}, & |\omega| < \omega_0, \\ 0, & |\omega| \geqslant \omega_0 \end{cases}$$

(*низкочастотный белый шум*).

Найти автоковариационную функцию данного процесса и выяснить, является ли низкочастотный белый шум дифференцируемым.

◁ По формуле (17) имеем

$$K_X(\tau) = \int\limits_0^{+\infty} S_X(\omega)\,\cos\omega\tau\,d\omega = \frac{D_X}{\omega_0}\int\limits_0^{\omega_0}\cos\omega\tau\,d\omega = D_X\,\frac{\sin\omega_0\tau}{\omega_0\tau}.$$

Так как $|\,K_X''(\tau)|_{\tau=0} = \dfrac{D_X\omega_0^2}{3} < +\infty$, то процесс дифференцируем. ▷

18.665. Пусть $X\,(t)$ — произвольный стационарный в широком смысле случайный процесс. Доказать «соотношение неопределенности» (22).

18.666. Стационарный в широком смысле процесс $X\,(t)$ таков, что его автоковариационная функция принимает лишь положительные значения. Показать, что для такого процесса неравенство (22) обращается в равенство.

В задачах 18.667 и 18.668 заданы автоковариационные функции некоторых случайных процессов. В каждом из этих случаев найти спектральную плотность и, проверив ее свойства, убедиться, что заданный процесс не является стационарным в широком смысле.

18.667. $K_X(\tau) = \begin{cases} a & \text{при} \quad |\tau| \leqslant T, \\ 0 & \text{при} \quad |\tau| > T. \end{cases}$

18.668. $K_X(\tau) = \begin{cases} 1 - \dfrac{\tau^2}{\tau_0^2} & \text{при} \quad |\tau| \leqslant \tau_0, \\ 0 & \text{при} \quad |\tau| > \tau_0. \end{cases}$

18.669. Спектральная плотность стационарного в широком смысле случайного процесса $X(t)$ имеет следующий вид:

$$S_X(\omega) = \begin{cases} \dfrac{\sigma_X^2}{\omega_2 - \omega_1} & \text{при} \quad \omega_1 < |\omega| < \omega_2 \quad (\omega_2 > \omega_1 > 0), \\ 0 & \text{в остальных случаях} \end{cases}$$

(*полосовой белый шум*). Вычислить автоковариационную функцию $K_X(\tau)$.

18.670 (продолжение). В условиях предыдущей задачи рассмотреть случай $\omega_2 \to \omega_1$. Какому случайному процессу соответствует этот предельный случай?

В задачах 18.671–18.674 определить ковариационную функцию, дисперсию и эффективную ширину спектра стационарного процесса, имеющего заданную спектральную плотность.

18.671. $S_X(\omega) = \dfrac{b}{\pi} \dfrac{\alpha}{\omega^2 + \alpha^2}, \quad b > 0, \ \alpha > 0.$

18.672. $S_X(\omega) = \dfrac{b}{\pi} \dfrac{\alpha^3}{(\omega^2 + \alpha^2)^2}, \quad b > 0, \ \alpha > 0.$

18.673. $S_X(\omega) = ae^{-|\omega|/\omega_0}, \quad a > 0, \ \omega_0 > 0.$

18.674. $S_X(\omega) = \begin{cases} a\left(1 - \dfrac{\omega^2}{\omega_0^2}\right), & |\omega| < \omega_0, \\ 0, & |\omega| \geqslant \omega_0. \end{cases}$

18.675. Найти спектральную плотность, эффективную ширину спектра и средний интервал корреляции стационарного случайного процесса $X(t)$ с ковариационной функцией

$$K_X(\tau) = D_X e^{-\alpha^2\tau^2}, \quad \alpha > 0.$$

18.676. Найти спектральную плотность, эффективную ширину спектра стационарного случайного процесса $X(t)$ с ковариационной функцией $K_X(\tau) = D_X e^{-\alpha|\tau|} \cos \beta\tau$.

18.677. Показать, что дисперсия процесса $Y(t) = \dfrac{dX(t)}{dt}$, где $X(t)$ — дифференцируемый стационарный в широком смысле случайный процесс со спектральной плотностью $S_X(\omega)$, может быть записана в виде

$$D_Y = \int\limits_0^{+\infty} \omega^2 S_X(\omega)\, d\omega.$$

18.678*. Доказать, что условие

$$\int_0^{+\infty} \omega^2 S_X(\omega)\,d\omega < c < \infty$$

является необходимым и достаточным условием дифференцируемости стационарного в широком смысле случайного процесса $X(t)$ со спектральной плотностью $S_X(\omega)$. Решить вопрос о дифференцируемости стационарных процессов со спектральными плотностями, заданными в задачах 18.671–18.674.

18.679. Найти дисперсию процесса

$$Y(t) = \frac{dX(t)}{dt},$$

где $X(t)$ — стационарный процесс из задачи 18.672.

18.680. Стационарный процесс $X(t)$ имеет спектральную плотность $S_X(\omega)$. Найти спектральную плотность процесса

$$Y(t) = aX(t) + b\,\frac{dX(t)}{dt}.$$

В задачах 18.681–18.684 найти спектральную плотность, эффективную ширину спектра, а также проверить условие дифференцируемости следующих процессов.

18.681. $X(t)$ — импульсный сигнал длительности t_0 (см. задачу 18.645).

18.682. $X(t)$ — фототелеграфный сигнал (см. задачу 18.647).

18.683. $X(t)$ — стационарный случайный процесс с автоковариационной функцией

$$K_X(\tau) = D_X e^{-\alpha|\tau|}(1 + \alpha|\tau|), \quad \alpha > 0.$$

18.684. $X(t)$ — стационарный случайный процесс с автоковариационной функцией

$$K_X(\tau) = \sigma_X^2 e^{-\alpha|\tau|}\left(\cos\beta\tau + \frac{\alpha}{\beta}\sin\beta|\tau|\right).$$

18.685. На вход линейного устройства, состоящего из линии задержки на время t_1 и вычитающего устройства (рис. 20), поступает стационарный случайный процесс $X(t)$ с математическим ожиданием m_X и ковариационной функцией $K_X(\tau)$. Найти математическое ожидание и ковариационную функцию процесса $Y(t) = X(t + t_1) - X(t)$ на выходе устройства. Сохраняется ли стационарность процесса на выходе?

18.686 (продолжение). На вход линейного устройства, описанного в предыдущей задаче, поступает стационарный случайный процесс, характеризуемый спектральной плотностью $S_X(\omega)$. Найти спектральную плотность $S_Y(\omega)$ процесса на выходе устройства.

18.687. На вход фильтра, состоящего из двух последовательно включенных линейных устройств, описанных в задаче 18.685

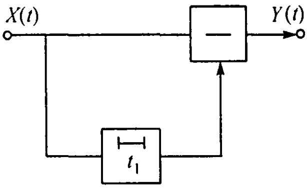


Рис. 20

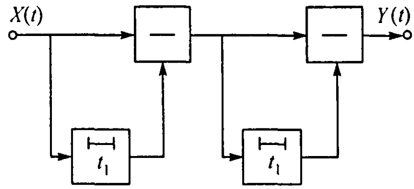


Рис. 21

(рис. 21), поступает стационарный сигнал $X(t)$ с ковариационной функцией $K_X(\tau)$. Определить дисперсию процесса на выходе фильтра.

18.688 (продолжение). В условиях задачи 18.687 входной сигнал $X(t)$ характеризуется спектральной плотностью

$$S_X(\omega) = \frac{D_X}{\alpha\sqrt{\pi}}\, e^{-\omega^2/4\alpha^2}, \quad \alpha > 0.$$

Найти дисперсию сигнала на выходе фильтра.

18.689*. Найти спектральную плотность случайного сигнала $Z(t) = X(t)Y(t)$, где $X(t)$ и $Y(t)$ — два независимых телеграфных сигнала с параметрами λ_1 и λ_2 соответственно.

5. Преобразование стационарных случайных функций линейными динамическими системами с постоянными коэффициентами. *Линейной динамической системой с постоянными коэффициентами* называется система, описываемая линейным дифференциальным уравнением с постоянными коэффициентами

$$\left(a_n\frac{d^n}{dt^n}+a_{n-1}\frac{d^{n-1}}{dt^{n-1}}+\ldots+a_0\right)y(t)=\left(b_m\frac{d^m}{dt^m}+\cdots+b_0\right)x(t), \qquad (23)$$

где $x(t)$ — реализация *входного* случайного стационарного процесса, $y(t)$ — реализация процесса на выходе системы.

Передаточной функцией линейной динамической системы называется функция комплексной переменной p, определяемая формулой

$$b_m p^m + b_{m-1}p^{m-1} + \cdots + b_0$$

Функция $H(p)$, как видно из определения, есть отношение преобразованных по Лапласу выходного сигнала к входному сигналу, определяемых из уравнения (23) при нулевых начальных условиях. Свойства сигнала на выходе линейной динамической системы полностью определяются свойствами передаточной функции $H(p)$ и свойствами входного сигнала. Говорят, что линейная динамическая система удовлетворяет *условию устойчивости*, если функция $H(p)$ не имеет полюсов в правой полуплоскости комплексной плоскости p.

Если на вход устойчивой линейной динамической системы с постоянными коэффициентами подается стационарный входной сигнал, то по прошествии достаточно большого времени с момента начала воздействия (именно, при $t \geqslant \tau_0$, где τ_0 — характерное время релаксации переходных процессов) сигнал на выходе системы будет близок к стационарному в широком смысле процессу.

Если $X(t)$ — входной стационарный сигнал с характеристиками m_X и $S_X(\omega)$, то соответствующие характеристики выходного процесса $Y(t)$ в стационарном режиме (т.е. при $t \gg \tau_0$) будут

$$m_Y = \frac{b_0}{a_0} m_X,$$

$$S_Y(\omega) = |H(i\omega)|^2 S_X(\omega). \tag{25}$$

Функция $|H(i\omega)|^2$ называется *амплитудно-частотной характеристикой системы.*

Дисперсия стационарного процесса на выходе системы в силу формулы (25) равна

$$D_Y = \frac{1}{2}\int_{-\infty}^{+\infty} S_Y(\omega)\,d\omega = \frac{1}{2}\int_{-\infty}^{+\infty} |H(i\omega)|^2 S_X(\omega)\,d\omega. \tag{26}$$

Для конечности дисперсии необходима и достаточна сходимость несобственного интеграла в формуле (26). Достаточным, например, является условие, чтобы порядок оператора дифференцирования входного сигнала в формуле (23) был не выше порядка дифференцирования выходного сигнала (т.е. условие $m \leqslant n$).

Пример 9. На вход линейной динамической системы, описываемой дифференциальным уравнением

$$y''(t) + 5y'(t) + 4y(t) = x'(t) + 4x(t),$$

поступает стационарный случайный процесс с математическим ожиданием $m_X = m$ и ковариационной функцией

$$K_X(\tau) = \sigma_X^2 e^{-\alpha|\tau|} \quad (\alpha > 0).$$

а) Проверить условия конечности дисперсии на выходе и устойчивости системы.

б) Найти математическое ожидание, ковариационную функцию и дисперсию стационарного процесса на выходе системы.

$\lhd$ а) Так как $m = 1 < n = 2$, то условие конечности дисперсии выполнено. Преобразуя по Лапласу данное дифференциальное уравнение с нулевыми начальными условиями, находим передаточную функцию системы:

$$H(p) = \frac{p+4}{p^2+5p+4}.$$

Полюса передаточной функции $p_1 = -4$, $p_2 = -1$ имеют отрицательные действительные части, следовательно, условие устойчивости системы выполнено.

б) Учитывая, что $H(p)$ — дробно-рациональная функция, находим амплитудно-частотную характеристику системы по формуле

$$|H(i\omega)|^2 = H(i\omega)\,H(-i\omega) = \frac{1}{\omega^2+1}.$$

Учитывая, что для процесса с ковариационной функцией $K_X(\tau) = \sigma_X^2 e^{-\alpha|\tau|}$ спектральная плотность $S_X(\omega) = \sigma_X^2 \dfrac{2\alpha}{\pi} \dfrac{1}{\omega^2+a^2}$ (см. пример 7), получим по формуле (25)

$$S_Y(\omega) = \frac{S_X(\omega)}{\omega^2+1} = \sigma_X^2 \frac{2\alpha}{\pi} \frac{1}{(\omega^2+a^2)(\omega^2+1)}.$$

Ковариационную функцию выходного сигнала находим с помощью прямого преобразования Фурье:

$$K_Y(\tau) = \frac{1}{2}\int_{-\infty}^{+\infty} S_Y(\omega)\, e^{i\omega\tau}\, d\omega = \sigma_X^2 \frac{\alpha}{\pi} \int_{-\infty}^{+\infty} \frac{e^{i\omega\tau}\, d\omega}{(\omega^2+a^2)(\omega^2+1)}.$$

Последний интеграл вычисляем с помощью теории вычетов, замыкая контур интегрирования в верхней (при $\tau > 0$) или в нижней (при $\tau < 0$) полуплоскости. Например, при $\tau > 0$

$$K_Y(\tau) = \sigma_X^2 \frac{\alpha}{\pi} 2\pi i \left\{ \text{выч} \left[\frac{e^{ip\tau}}{(p^2+\alpha^2)(p^2+1)};\, i\alpha \right] + \right.$$
$$\left. + \text{выч} \left[\frac{e^{ip\tau}}{(p^2+\alpha^2)(p^2+1)};\, i \right] \right\} = \frac{\sigma_X^2}{1+\alpha} e^{-\alpha\tau}.$$

В силу четности автоковариационной функции получим окончательно

$$K_Y(\tau) = \frac{\sigma^2}{1+\alpha} e^{-\alpha|\tau|}.$$

Дисперсия сигнала на выходе равна

$$D_Y = K_Y(0) = \frac{\sigma_X^2}{1+\alpha}.$$

Таким образом, дисперсия на выходе системы в $1+\alpha$ раз меньше дисперсии на входе. Это объясняется тем, что данная динамическая система является, по существу, интегрирующей цепью, а операция интегрирования процесса сглаживает шум. $\triangleright$

18.690. На вход линейной динамической системы, описываемой уравнением

$$2y'(t) + y(t) = 3x(t),$$

поступает стационарный сигнал с характеристиками $m_X = 1$, $K_X(\tau) = e^{-2|\tau|}$. Найти m_Y и D_Y на выходе системы.

18.691. На вход интегрирующей RC-цепочки (рис. 16), описываемой дифференциальным уравнением

$$y'(t) + \beta y(t) = \beta x(t), \quad \beta = \frac{1}{RC} > 0,$$

поступает стационарный низкочастотный белый шум $X(t)$ с характеристиками $m_X = 0$, $\sigma_X^2 = 2$ и частотой отсечки ω_0. Найти математическое ожидание и дисперсию процесса на выходе системы.

18.692* (продолжение). В условиях предыдущей задачи найти дисперсию D_Y в случае идеального белого шума на входе (т.е. рассматривая входной процесс как предельный случай низкочастотного белого шума с постоянной спектральной плотностью $S_X(\omega) = c^2$ при всех ω).

18.693*. Напряжение на входе интегрирующей RC-цепочки, описанной в задаче 18.692, представляет собой стационарный случайный процесс со спектральной плотностью $S_X(\omega) = \sigma_X^2 e^{-\alpha^2\omega^2}$. Найти отношение дисперсии выходного напряжения $Y(t)$ к дисперсии входного напряжения. Чему равно это отношение в частном случае $\beta = \alpha = 1$?

18.694. Случайный стационарный процесс $X(t)$ со спектральной плотностью $S_X(\omega) = \dfrac{a^2}{(\omega^2+\alpha^2)^2}$ $(a > 0,\ \alpha > 0)$ поступает на вход дифференцирующей цепочки. Рассматривая дифференцирующую цепочку как простейший пример линейной стационарной системы, найти дисперсию стационарного процесса $Y(t) = \dfrac{dX(t)}{dt}$ и отношение дисперсий σ_Y^2/σ_X^2.

18.695. На вход линейной динамической системы, описываемой уравнением

$$y'(t) + 5y(t) = x'(t) + 2x(t),$$

поступает стационарный случайный процесс с ковариационной функцией

$$K_X(\tau) = \sigma_X^2 e^{-|\tau|}$$

и математическим ожиданием $m_X = 5$. Найти математическое ожидание, спектральную плотность и дисперсию стационарного случайного процесса на выходе системы.

18.696. На вход *идеального полосового фильтра* с амплитудно-частотной характеристикой

$$|H(i\omega)|^2 = \begin{cases} 1, & \text{если} \quad \omega_1 \leqslant |\omega| \leqslant \omega_2, \\ 0, & \text{если} \quad \omega| < \omega_1 \text{ или } |\omega| > \omega_2, \end{cases}$$

поступает случайный телеграфный сигнал. Найти дисперсию процесса на выходе фильтра.

18.697*. На вход линейной динамической системы, описываемой уравнением $y(t) = \beta x(t) + \dfrac{dx(t)}{dt}$, поступает стационарный сигнал с ковариационной функцией $K_X(\tau) = \sigma_X^2 e^{-\alpha|\tau|}(1 + \alpha|\tau|)$. Найти спектральную плотность, эффективную ширину спектра и дисперсию, характеризующие стационарный сигнал на выходе системы.

18.698*. Работа электрической цепи (рис. 22) описывается дифференциальным уравнением

$$y'(t) + \frac{R}{L}y(t) = x'(t).$$

На ее вход поступает низкочастотный белый шум, описанный в

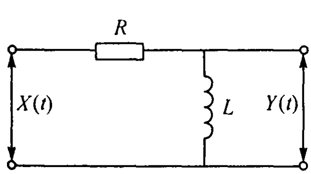


Рис. 22

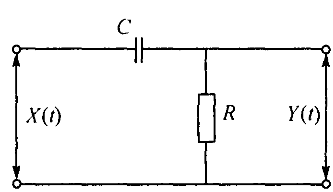


Рис. 23

примере 8. Найти дисперсию случайного процесса на выходе в стационарном режиме работы цепи.

18.699*. Случайный сигнал $Y(t)$, возникающий при прохождении случайного стационарного сигнала $X(t)$ через RC-фильтр (рис. 23), определяется дифференциальным уравнением

$$y'(t) + \beta y(t) = x'(t), \quad \beta = \frac{1}{RC} > 0.$$

Найти дисперсию процесса $Y(t)$, если на вход RC-фильтра поступает импульсный сигнал, определенный в задаче 18.645.

18.699* (продолжение). На вход RC-фильтра, описанного в задаче 18.699, поступает стационарный случайный процесс с ковариационной функцией $K_X(\tau) = e^{-\lambda|\tau|}(1 + \lambda|\tau|)$, $\lambda > 0$, $\lambda \neq \beta$. Найти дисперсию процесса $Y(t)$ на выходе фильтра.

18.700*. На вход резонансного контура, описываемого дифференциальным уравнением

$$y''(t) + 2\beta y'(t) + \omega_0^2 y(t) = x(t) \quad (\omega_0 > \beta),$$

поступает идеальный белый шум ($S_X(\omega) = c^2$ при $-\infty < \omega < \infty$). Определить дисперсию стационарного случайного сигнала $Y(t)$ на выходе системы.

Глава 19

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

§ 1. Методы статистического описания результатов наблюдений

Статистическая обработка данных в настоящее время проводится с использованием компьютеров. Современные пакеты статистических программ (ПСП) позволяют использовать при анализе данных не только огромное число алгоритмов и процедур статистического анализа, но и проводить их проверку, заполнять пропущенные значения, формировать определенные подмножества данных, обрабатывать данные представленные в нечисловом виде, моделировать выборки из генеральной совокупности практически с любым заданным законом распределения. Перечислим наиболее популярные пакеты статистических программ: Statistica, Statgraphics, Stadia, SPSS и другие. Большой набор функций для статистической обработки данных содержат пакеты Matlab и Excell. Описание ПСП и примеров их использования можно найти в специальной литературе, например [13–16].

Использование ПСП значительно упрощает использование статистических методов. Читателю мы рекомендуем использовать ПСП при решении почти всех задач, связанных с выполнением расчетов и построением графиков. Более того, на стадии обучения работе на ПСП следует решить на компьютере примеры из теоретических разделов задачника, сравнить ответы и проанализировать результаты. При использовании ПСП выполняется только расчетная часть решаемой задачи. Выбор метода с учетом конкретных условий, учет ограничений при использовании тех или иных статистических процедур, интерпретация полученных результатов требует, конечно, знания теории математической статистики.

1. Выборка и способы ее представления. Математическая статистика позволяет получать обоснованные выводы о параметрах, видах распределений и других свойствах случайных величин по конечной совокупности наблюдений над ними — *выборке*.

Выборка понимается следующим образом. Пусть случайная величина X наблюдается в случайном эксперименте $\mathcal{E}$. Повторим эксперимент $\mathcal{E}$ n раз, предполагая, что условия проведения эксперимента, а следовательно, и распределение наблюдаемой случайной величины X не изменяются от эксперимента к эксперименту. Этот новый составной эксперимент связан с n-мерной случайной величиной — случайным вектором $(X_1, X_2, \ldots, X_n)$, где X_j — случайная величина, соответствующая j-му эксперименту. Очевидно, X_j, $j = 1, 2, \ldots, n$, — независимые в совокупности случайные величины, каждая из которых имеет тот же закон распределения, что и случайная величина X.

Закон распределения случайной величины X называется *распределением генеральной совокупности*, а случайный вектор $(X_1, \ldots, X_n)$ — *выборочным вектором*. Числа $x_1, \ldots, x_n$, получаемые на практике при n-кратном повторении эксперимента $\mathcal{E}$ в неизменных условиях, представляют собой реализацию выборочного вектора и называются выборкой $(x_1, \ldots, x_n)$ объема n.

Выборку $(x_1, x_2, \ldots, x_n)$ при необходимости можно рассматривать как точку *выборочного пространства*, т.е. множества, на котором задано распределение выборочного вектора $(X_1, X_2, \ldots, X_n)$. Аналогично определяется выборка в случае, когда случайный эксперимент $\mathcal{E}$ связан с несколькими случайными величинами. Например, выборка объема n из двумерной генеральной совокупности есть последовательность $(x_1, y_1), \ldots, (x_n, y_n)$ пар значений случайных величин X и Y, принимаемых ими в n независимых повторениях случайного эксперимента $\mathcal{E}$.

Результаты наблюдений $x_1, x_2, \ldots, x_n$ генеральной совокупности X, записанные в порядке их регистрации, обычно труднообозримы и неудобны для дальнейшего анализа. Задачей статистического описания выборки является получение такого ее представления, которое позволяет выявить характерные особенности совокупности исходных данных.

Вариационным рядом выборки $x_1, x_2, \ldots, x_n$ называется способ ее записи, при котором элементы упорядочиваются по величине, т.е. записываются в виде последовательности $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$, где $x^{(1)} \leqslant x^{(2)} \leqslant \ldots \leqslant x^{(n)}$. Разность между максимальным и минимальным элементами выборки $x^{(n)} - x^{(1)} = \omega$ называется *размахом выборки*.

Пусть выборка $(x_1, x_2, \ldots, x_n)$ содержит k различных чисел $z_1, z_2, \ldots, z_k$, причем z_i встречается n_i раз $(i = 1, 2, \ldots, k)$. Число n_i называется *частотой* элемента выборки z_i. Очевидно, что $\sum\limits_{i=1}^{k} n_i = n$.

Статистическим рядом называется последовательность пар (z_i, n_i). Обычно статистический ряд записывается в виде таблицы, первая строка которой содержит элементы z_i, а вторая — их частоты.

Пример 1. Записать в виде вариационного и статистического рядов выборку 5, 3, 7, 10, 5, 5, 2, 10, 7, 2, 7, 7, 4, 2, 4. Определить размах выборки.

◁ Объем выборки $n = 15$. Упорядочив элементы выборки по величине, получим вариационный ряд

$$2,\ 2,\ 2,\ 3,\ 4,\ 4,\ 5,\ 5,\ 5,\ 7,\ 7,\ 7,\ 7,\ 10,\ 10.$$

Размах выборки

$$\omega = 10 - 2 = 8.$$

Различными в заданной выборке являются элементы $z_1 = 2$, $z_2 = 3$, $z_3 = 4$, $z_4 = 5$, $z_5 = 7$, $z_6 = 10$; их частоты соответственно равны $n_1 = 3$, $n_2 = 1$, $n_3 = 2$, $n_4 = 3$, $n_5 = 4$, $n_6 = 2$. Следовательно, статистический ряд исходной выборки можно записать в виде следующей таблицы:

z_i	2	3	4	5	7	10
n_i	3	1	2	3	4	2

Для контроля правильности записи находим $\sum n_i = 15$ [1]. ▷

Для каждой из приведенных ниже выборок определить размах, а также построить вариационный и статистический ряды.

19.1. 11, 15, 12, 0, 16, 19, 6, 11, 12, 13, 16, 8, 9, 14, 5, 11, 3.

19.2. 17, 18, 16, 16, 17, 18, 19, 17, 15, 17, 19, 18, 16, 16, 18, 18.

19.3. В группе на занятии по статистике проводится эксперимент по регистрации номера месяца рождения каждого из студентов (опрос проводится, например, по списку группы). Построить вариационный и статистический ряды полученной выборки.

19.4. Из художественного произведения некоторого автора (с числом страниц, бо́льшим ста), используя таблицу случайных чисел (таблица П12), выбрать случайным образом 20 страниц. На каждой из выбранных страниц, используя ту же таблицу, выбрать строку, содержащую некоторое предложение либо его начало. Таким образом случайно выбираются 20 предложений, и в каждом из них подсчитывается число слов. Построить вариационный и статистический ряды полученной выборки и определить ее размах.

При большом объеме выборки ее элементы объединяют в группы (разряды), представляя результаты опытов в виде *группированного статистического ряда*. Для этого интервал, содержащий все элементы выборки, разбивается на k непересекающихся интервалов. Вычисления значительно упрощаются, если эти интервалы имеют одинаковую длину $b \approx \dfrac{w}{k}$. Во всем дальнейшем изложении рассматривается именно этот случай. После того как частичные интервалы выбраны, определяют частоты — количество n_i элементов выборки, попавших в i-й интервал (элемент, совпадающий с верхней границей интервала, относится к последующему интервалу). Получающийся статистический ряд в верхней строке содержит середины z_i интервалов группировки, а в нижней — частоты n_i ($i = 1, 2, \ldots, k$).

Наряду с частотами одновременно подсчитываются также накопленные частоты $\sum\limits_{j=1}^{i} n_j$, относительные частоты n_i/n и *накопленные отно-*

[1]) Здесь и в дальнейшем в суммах вида $\sum\limits_{i=1}^{k} n_i$ будем опускать верхний и нижний индексы.

сительные частоты $\sum_{j=1}^{i} n_j/n$, $i = 1, 2, \ldots, k$. Полученные результаты сводятся в таблицу, называемую *таблицей частот группированной выборки.*

Следует помнить, что группировка выборки вносит погрешность в дальнейшие вычисления, которая растет с уменьшением числа интервалов.

Пример 2. Представить выборку 55 наблюдений в виде таблицы частот, используя 7 интервалов группировки. Выборка:

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

◁ Размах выборки $w = 23{,}8 - 10{,}1 - 13{,}7$. Длина интервала группировки $b = 13{,}7/7 \approx 2$. В качестве первого интервала удобно взять интервал 10–12. Результаты группировки сведены в таблицу 1.1. ▷

Таблица 1.1

Номер интервала i	Границы интервала	Середина интервала z_i	Частота n_i частота	Накопленная частота $\sum_{j=1}^{i} n_j$	Относительная n_j/n	Накопленная относительная частота $\sum_{j=1}^{i} n_j/n$
1	10–12	11	2	2	0,0364	0,0364
2	12–14	13	4	6	0,0727	0,1091
3	14–16	15	8	14	0,1455	0,2546
4	16–18	17	12	26	0,2182	0,4728
5	18–20	19	16	42	0,2909	0,7637
6	20–22	21	10	52	0,1818	0,9455
7	22–24	23	3	55	0,0545	1,0000

В задачах 19.5, 19.6 найти размах выборки, число и длину интервалов, а также составить таблицу частот (границы первого интервала указываются).

19.5. Время решения контрольной задачи учениками 4-го класса (в секундах):

38	60	41	51	33	42	45	21	53	60
68	52	47	46	49	49	14	57	54	59
77	47	28	48	58	32	42	58	61	30
61	35	47	72	41	45	44	55	30	40
67	65	39	48	43	60	54	42	59	50

Первый интервал: 14–23.

19.6. Продолжительность работы электронных ламп одного типа (в часах):

13,4	14,7	15,2	15,1	13,0	8,8	14,0	17,9	15,1	16,5	16,6
14,2	16,3	14,6	11,7	16,4	15,1	17,6	14,1	18,8	11,6	13,9
18,0	12,4	17,2	14,5	16,3	13,7	15,5	16,2	8,4	14,7	15,4
11,3	10,7	16,9	15,8	16,1	12,3	14,0	17,7	14,7	16,2	17,1
10,1	15,8	18,3	17,5	12,7	20,7	13,5	14,0	15,7	21,9	14,3
17,7	15,4	10,9	18,2	17,3	15,2	16,7	17,3	12,1	19,2	

Первый интервал: 8,4–10,4.

Пусть $(x_1, x_2, \ldots, x_n)$ — выборка из генеральной совокупности с функцией распределения $F_X(x)$. *Распределением выборки* называется распределение дискретной случайной величины, принимающей значения $x_1, x_2, \ldots, x_n$ с вероятностями $1/n$. Соответствующая функция распределения называется *эмпирической* (*выборочной*) *функцией распределения* и обозначается $F_n^*(x)$.

Эмпирическая функция распределения определяется по значениям накопленных частот соотношением

$$F_n^*(x) = \frac{1}{n} \sum_{z_i < x} n_i, \tag{1}$$

где суммируются частоты тех элементов выборки, для которых выполняется неравенство $z_i < x$. Очевидно, что $F_n^*(x) = 0$ при $x \leqslant x^{(1)}$ и $F_n^*(x) = 1$ при $x > x^{(n)}$. На промежутке $(x^{(1)}, x^{(n)}]$ $F_n^*(x)$ представляет собой неубывающую кусочно постоянную функцию.

Аналогично формуле (1) определяется эмпирическая функция распределения для группированной выборки.

Значение эмпирической функции распределения для статистики определяется следующим утверждением.

Теорема (Гливенко). *Пусть $F_n^*(x)$ — эмпирическая функция распределения, построенная по выборке объема n из генеральной совокупности с функцией распределения $F_X(x)$. Тогда для любого $x \in$ $\in (-\infty, +\infty)$ и любого $\varepsilon > 0$*

$$\lim_{n\to\infty} P(|F_n^*(x) - F_X(x)| < \varepsilon) = 1.$$

Таким образом, при каждом x $F_n^*(x)$ сходится по вероятности к $F_X(x)$ и, следовательно, при большом объеме выборки может служить приближенным значением (оценкой) функции распределения генеральной совокупности в каждой точке x.

В ряде случаев для наглядного представления выборки используют гистограмму и полигон частот.

Гистограммой частот группированной выборки называется кусочно постоянная функция, постоянная на интервалах группировки и принимающая на каждом из них значения $\dfrac{n_i}{b}$, $i = 1, 2, \ldots, k$ соответственно. Площадь ступенчатой фигуры под графиком гистограммы равна объему выборки n.

Аналогично определяется *гистограмма относительных частот*. Площадь соответствующей ступенчатой фигуры для нее равна единице. При увеличении объема выборки и уменьшении интервала группировки гистограмма относительных частот является статистическим аналогом плотности распределения $f_X(x)$ генеральной совокупности.

Полигоном частот называется ломаная с вершинами в точках $\left(z_i, \dfrac{n_i}{b}\right)$, $i = 1, 2, \ldots, k$, а *полигоном относительных частот* — ломаная с вершинами в точках $\left(z_i, \dfrac{n_i}{nb}\right)$, $i = 1, 2, \ldots, k$. Таким образом, полигон относительных частот получается из полигона частот сжатием по оси Oy в n раз.

Если плотность распределения генеральной совокупности является достаточно гладкой функцией, то полигон относительных частот является более хорошим приближением плотности, чем гистограмма.

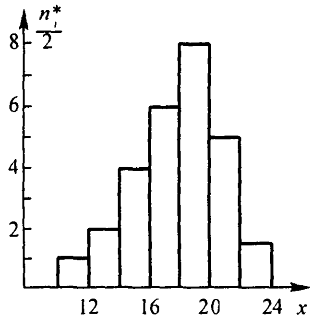


Рис. 24

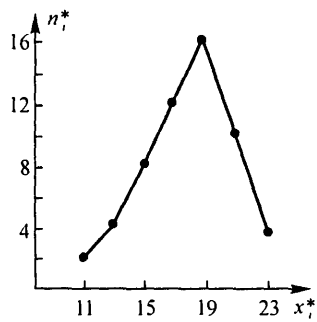


Рис. 25

Пример 3. Построить гистограмму и полигон частот, а также график эмпирической функции распределения группированной выборки из примера 2.

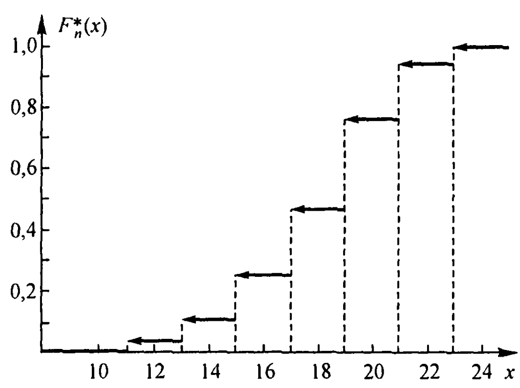


Рис. 26

◁ По результатам группировки (см. таблицу 1.1) строим гистограмму частот (рис. 24). Соединяя отрезками ломаной середины верхних оснований прямоугольников, из которых состоит полученная гистограмма, получаем соответствующий полигон частот (рис. 25).

Так как середина первого интервала группировки $z_1 = 11$, то $F_n^*(x) = 0$ при $x \leqslant 11$. Рассуждая аналогично, находим, что $F_n^*(x) = 1$ при $x > 23$. На полуинтервале $(11, 23]$ эмпирическую функцию распределения строим по данным третьего и последнего столбцов таблицы 1.1. $F_n^*(x)$ имеет скачки в точках, соответствующих серединам интервалов группировки. В результате получаем график $F_n^*(x)$, изображенный на рис. 26. ▷

В задачах 19.7–19.10 построить графики эмпирических функций распределения, гистограммы и полигоны частот для выборок, представленных статистическими рядами.

19.7.

z_i	15	16	17	18	19
n_i	1	4	5	4	2

19.8.

z_i	2	3	4	5	6	7	8
n_i	1	3	4	6	5	2	1

19.9.

Границы интервалов	10–20	20–30	30–40	40–50	50–60	60–70	70–80
Частоты	1	2	7	18	12	8	2

19.10.

Границы интервалов	18–20	20–22	22–24	24–26	26–28	28–30	30–32	32–34
Частоты	4	3	3	2	4	7	12	5

19.11. Измерения емкости затвор — сток у 80 полевых транзисторов дали следующие результаты:

1,9	3,1	1,3	0,7	3,2	1,1	2,9	2,7	2,7	4,0
1,7	3,2	0,9	0,8	3,1	1,2	2,6	1,9	2,3	3,2
4,1	1,3	2,4	4,5	2,5	0,9	1,4	1,6	2,2	3,1
1,5	1,1	2,3	4,3	2,1	0,7	1,2	1,5	1,8	2,9
0,8	0,9	1,7	4,1	4,3	2,6	0,9	0,8	1,2	2,1
3,2	2,9	1,1	3,2	4,5	2,1	3,1	5,1	1,1	1,9
0,9	3,1	0,9	3,1	3,3	2,8	2,5	4,0	4,3	1,1
2,1	3,8	4,6	3,8	2,3	3,9	2,4	4,1	4,2	0,9

Построить гистограмму и полигон относительных частот по этой выборке, предварительно проведя группировку. В качестве длины интервала взять следующие значения: а) $b = 0{,}3$; б) $b = 0{,}6$; в) $b = 1{,}2$.

В задачах статистического анализа сложных систем, например при разработке систем автоматического проектирования (САПР), широко используется *метод моделирования выборки* из генеральной совокупности с заданным законом распределения.

Пусть случайная величина X имеет функцию распределения $F_X(x)$. Как известно из теории вероятностей (см. задачу 18.504), случайная величина $Y = F_X(X)$ имеет равномерное распределение $R(0, 1)$. Отсюда следует, что случайная величина X может быть получена из равномерно распределенной случайной величины Y по формуле $X = F_X^{-1}(Y)$, где F_X^{-1} — функция, обратная к F_X (заведомо существующая для случайных величин непрерывного типа).

Метод моделирования выборки из генеральной совокупности с законом распределения $F_X(x)$ реализуется следующим алгоритмом:

$$x_j = F_X^{-1}(y_j), \quad j = 1, 2, \ldots, n, \tag{2}$$

где $y_1, y_2, \ldots, y_n$ — выборка из генеральной совокупности с равномерным распределением $R(0, 1)$, являющаяся последовательностью случайных чисел.

Алгоритм (2) получения выборки из генеральной совокупности с законом распределения $F_X(x)$ поясняется на рис. 27.

Случайные числа $y_1, y_2, \ldots, y_n$ можно получить, выбрав случайным образом n чисел из таблицы П12 и разделив каждое выбранное число на 100. При наличии любого вычислительного устройства случайные числа $y_1, y_2, \ldots, y_n$ генерируются с помощью формулы

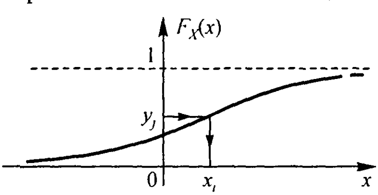


Рис. 27

$$y_{j+1} = \{my_j\}, \quad j \in \mathbb{N}, \tag{3}$$

где $\{a\}$ — дробная часть числа a, а m — простое число, большее десяти. В качестве начального значения y_1 в формуле (3) можно выбрать произвольное число из интервала $(0, 1)$ с ненулевыми разрядами в десятичной системе счисления. От удачного выбора начального значения y_1 зависит качество последовательности $y_1, y_2, \ldots, y_n$.

19.12. Шкала вольтметра имеет цену делений 1 В. При измерении напряжения отсчет делается с точностью до ближайшего целого деления. Считая, что ошибка округления имеет равномерное распределение, получить методом моделирования выборку объемом $n = 20$. Построить вариационный ряд и гистограмму частот полученной выборки.

19.13. Методом моделирования получить выборки объемом $n = 10$ из генеральной совокупности с показательным законом распределения $\mathrm{Ex}(\lambda)$ с $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$.

19.13. Методом моделирования получить выборки объемом $n = 10$ из генеральной совокупности с показательным законом распределения $\mathrm{Ex}(\lambda)$ с $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$.

19.14. Автомобили подъезжают к автозаправочной станции последовательно, причем время между прибытием двух автомобилей имеет показательное распределение с параметром $\lambda_1 = 1$. Если очереди нет, автомобиль заправляется сразу, в противном случае он становится в очередь. Время заправки автомобиля имеет показательное распределение с параметром $\lambda_2 = 2$. Используя выборки, полученные в задаче 19.13, составить таблицу, содержащую время подъезда для каждого из пяти последовательно прибывших автомобилей, время начала и конца заправки, продолжительность ожидания в очереди, общее время на ожидание и обслуживание.

19.15. Пусть t_i — время наработки на отказ i-го элемента схемы. Известно, что t_i распределено по закону $\mathrm{Ex}(\lambda_1)$, $i = 1, 2, 3$. Используя выборки, полученные в задаче 19.13, построить эмпирическую функцию распределения времени наработки на отказ для схемы на рис. 28.

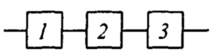


Рис. 28

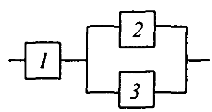


Рис. 29

19.16 (продолжение). В условиях предыдущей задачи построить функцию распределения времени наработки на отказ для схемы на рис. 29.

19.17. Время безотказной работы (в месяцах) телевизионной трубки имеет нормальное распределение $N(24, 3)$. Магазин продал 15 телевизоров. Методом моделирования получить выборку времени безотказной работы трубок у проданных телевизоров. Построить гистограмму и оценить наиболее вероятное число трубок, потребующих замены в течение 10 лет.

Для получения выборки из генеральной совокупности, распределенной по закону $N(0, 1)$, можно воспользоваться алгоритмом, основанным на любом из следующих соотношений:

$$u_j = \sqrt{-2\ln y_{j-1}}\cos 2\pi y_j, \tag{4}$$

$$j = 2, 3, \ldots, n,$$

$$v_j = \sqrt{-2\ln y_{j-1}}\sin 2\pi y_j, \tag{5}$$

где, как и выше, $y_1, y_2, \ldots, y_n$ — выборка из генеральной совокупности с равномерным распределением $R(0, 1)$.

19.18. Случайные величины X на Y независимы и имеют нормальные распределения $N(10, 2)$ и $N(9, 1)$ соответственно. Используя соотношения (4) и (5), получить выборки объема $n = 20$ и оценить вероятность того, что $\min\{X, Y\} < 10$.

19.19. Пусть X и Y — независимые случайные величины, распределенные соответственно по законам $N(0, 1)$ и $N(0, 3)$. Используя соотношения (4) и (5), получить выборки объема $n = 50$ для X и Y.

19.20. Используя результат предыдущей задачи, получить выборку объема $n = 50$ для случайной величины $Z = |X| + |Y|$ и построить гистограмму частот.

19.21. Используя результат задачи 19.19, получить выборку объема $n = 50$ для случайной величины $V = |X - Y|$ и построить гистограмму частот.

19.22. Пусть I — индикаторная случайная величина события $A = \{|X| < 1\}$, т.е.

$$I = \begin{cases} 1, & |X| < 1, \\ 0, & |X| \geqslant 1, \end{cases}$$

где X — случайная нормально распределенная величина из задачи 19.19. Используя выборку, полученную в этой задаче, найти частоту события $\{I = 1\}$ и сравнить полученный результат с точным значением вероятности этого события.

19.23.** Доказать теорему Гливенко.

2. Числовые характеристики выборочного распределения. Пусть x_1, $x_2, \ldots, x_n$ — выборка объема n из генеральной совокупности с функцией распределения $F_X(x)$. Рассмотрим выборочное распределение, т.е. распределение дискретной случайной величины, принимающей значения $x_1, x_2, \ldots, x_n$ с вероятностями, равными $1/n$. Числовые характеристики этого выборочного распределения называются *выборочными* (*эмпирическими*) *числовыми характеристиками.* Следует отметить, что выборочные числовые характеристики являются характеристиками данной выборки, но не являются характеристиками распределения генеральной совокупности. Чтобы подчеркнуть это различие, выборочные характеристики в дальнейшем будем обозначать теми же символами, что и в главе 18, но со значком * наверху. Некоторые выборочные характеристики имеют традиционные обозначения, например, $\overline{x}$ — выборочное среднее.

Пример 4. Получить формулы, определяющие выборочные математическое ожидание и дисперсию для негруппированной выборки объема n.

◁ Математическое ожидание дискретной случайной величины определяется по формуле (см. главу 18, § 2, формулу (4))

$$m_X = \sum_{j=1}^{n} p_j x_j.$$

Так как для выборочного распределения $p_j = 1/n$, то

$$m_x^* = \overline{x} = \frac{1}{n}\sum_{j=1}^{n} x_j. \tag{1}$$

Аналогично

$$D_x^* = \sum_{j=1}^{n}(x_j - \overline{x})^2 p_j = \frac{1}{n}\sum_{j=1}^{n}(x_j - \overline{x})^2 =$$

$$= \frac{\sum\limits_{j=1}^{n} x_j^2 - \left(\sum\limits_{j=1}^{n} x_j\right)^2 \Big/ n}{n} = \frac{1}{n}\left(\sum_{j=1}^{n} x_j^2 - n\overline{x}^2\right). \ \triangleright \tag{2}$$

Выборочной модой d_x^* унимодального (одновершинного) распределения называется элемент выборки, встречающийся с наибольшей частотой.

Выборочной медианой называется число h_x^*, которое делит вариационный ряд на две части, содержащие равное число элементов. Если объем выборки n — нечетное число (т.е. $n = 2l + 1$), то $h_x^* = x^{(l+1)}$, т.е. является элементом вариационного ряда со средним номером. Если же $n = 2l$, то $h_x^* = \frac{1}{2}\left(x^{(l)} + x^{(l+1)}\right)$.

Пример 5. Определить среднее, моду и медиану для выборки 5, 6, 8, 2, 3, 1, 1, 4.

$\triangleleft$ Представим данные в виде вариационного ряда: 1, 1, 2, 3, 4, 5, 6, 8. Выборочное среднее $\overline{x} = \frac{1}{8}(1 + 1 + 2 + 3 + 4 + 5 + 6 + 8) = 3{,}75$. Все элементы входят в выборку по одному разу, кроме 1, следовательно, мода $\tilde{d}_x = 1$. Так как $n = 8$, то медиана $\tilde{h}_x = \frac{1}{2}(3 + 4) = 3{,}5$. $\triangleright$

Вычислить моду, медиану, среднее и дисперсию следующих выборок:

19.24. 7, 3, 3, 6, 4, 5, 1, 2, 1, 3.

19.25. 3,1; 3,0; 1,5; 1,8; 2,5; 3,1; 2,4; 2,8; 1,3.

19.26. а) 1, 2, 3, 4, 5, 5, 9; б) 1, 2, 3, 4, 5, 5, 12. Сравнить полученные числовые результаты для выборок а) и б).

19.27. Доказать, что выборочные начальные и центральные моменты порядка s, $s = 1, 2, \ldots$, для негруппированной выборки объема n определяются следующими формулами:

$$\alpha_s^* = \frac{1}{n}\sum_{j=1}^{n} x_j^s, \quad \mu_s^* = \frac{1}{n}\sum_{j=1}^{n}(x_j - \alpha_1^*)^s.$$

19.28. Доказать, что выборочные начальные и центральные моменты порядка s, $s = 1, 2, \ldots$, для группированной выборки объема n определяются следующими формулами:

$$\alpha_s^* = \frac{1}{n}\sum_{i=1}^{k} n_i z_i^s, \quad \mu_s^* = \frac{1}{n}\sum_{i=1}^{k} n_i (z_i - \alpha_1^*)^s.$$

19.29. Доказать, что для выборочной дисперсии справедлива следующая формула:

$$D_X^* = \alpha_2^* - \overline{x}^2.$$

19.30. Доказать справедливость следующих соотношений:

$$\mu_3^* = \alpha_3^* - 3\alpha_2^*\alpha_1^* + 2\alpha_1^{*3},$$
$$\mu_4^* = \alpha_4^* - 4\alpha_3^*\alpha_1^* + 6\alpha_2^*\alpha_1^{*2} - 3\alpha_1^{*4}.$$

В задачах 19.31–19.33 определить среднее, моду, медиану и дисперсию группированных выборок.

19.31.

Границы интервалов	1–3	3–5	5–7	7–9	9–11	11–13
Частоты	1	2	4	2	1	1

19.32.

Границы интервалов	0–4	4–8	8–12	12–16	16–20	20–24
Частоты	1	1	3	2	1	1

19.33.

Границы интервалов	5–7	7–9	9–11	11–13	13–15	15–17
Частоты	8	14	40	26	6	4

19.34. Доказать следующие свойства выборочного среднего:

а) $\sum (x_i - \overline{x}) = 0$;

б)* $\sum (x_i - \overline{x})^2 < \sum (x_i - a)^2$, где $a \in \mathbb{R}$, $a \neq \overline{x}$.

19.35.** Доказать следующее свойство выборочной медианы:

$$\sum |x_i - h_X^*| < \sum |x_i - a|, \quad \text{где } a \in \mathbb{R}, \quad a \neq h_X^*.$$

19.36. Предположим, что в результатах наблюдений случайной величины X присутствует одна и та же систематическая погрешность a. Какое влияние оказывает эта систематическая погрешность на величины выборочных среднего, моды, медианы и дисперсии?

19.37. Как изменятся выборочные среднее, мода, медиана и дисперсия, если результаты наблюдения подвергнуть преобразованию масштаба, т.е. увеличить или уменьшить одновременно в k раз?

Результаты задач 19.36 и 19.37 используются для упрощения вычислений выборочных среднего и дисперсии группированной выборки. Для этого группированную выборку преобразуют следующим образом:

$$u_i = \frac{1}{b}(z_i - d_x^*), \quad i = 1, 2, \ldots, k, \tag{3}$$

где d_x^* — выборочная мода, а b — длина интервала группировки. Соотношения (3) показывают, что в выборку $z_1, z_2, \ldots, z_k$ внесена систематическая ошибка d_x^*, а результат подвергнут преобразованию масштаба с коэффициентом $k = 1/b$. Полученный в результате набор чисел $u_1, u_2, \ldots, u_k$ можно рассматривать как выборку из генеральной совокупности $U = \frac{1}{b}(x - d_x^*)$. Тогда выборочные среднее $\overline{x}$ и дисперсия D_x^* исходных данных связаны со средним $\overline{u}$ и дисперсией D_U^* преобразованных данных следующими соотношениями:

$$\overline{x} = b\overline{u} + d_x^*, \tag{4}$$

$$D_x^* = b^2 D_U^*. \tag{5}$$

Пример 6. Вычислить среднее и дисперсию группированной выборки

Границы интервалов	134–138	138–142	142–146	146–150	150–154	154–158
Частоты	1	3	15	18	14	2

◁ Длина интервала группировки $b = 4$, значение середины интервала, встречающегося с наибольшей частотой, $d_x^* = 148$. Таким образом, преобразование последовательности середин интервалов выполняется по формуле (3):

$$u_i = \frac{z_i - 148}{4}, \quad i = 1, 2, \ldots, 6.$$

Таблица 1.2

i	z_i	u_i	n_i	$n_i u_i$	$n_i u_i^2$	$n_i(u_i + 1)^2$
1	136	−3	1	−3	9	4
2	140	−2	3	−6	12	3
3	144	−1	15	−15	15	0
4	148	0	18	0	0	18
5	152	1	14	14	14	56
6	156	2	2	4	8	18
$\sum$	–	–	53	−6	58	99

Вычисления удобно свести в таблицу (см. таблицу 1.2). Последний столбец таблицы 1.2 служит для контроля вычислений при помощи тождества

$$\sum n_i(u_i+1)^2 \equiv \sum n_i u_i^2 + 2\sum n_i u_i + \sum n_i.$$

Подставляя в тождество данные последней строки таблицы 1.2, получим

$$58 + 2\cdot(-6) + 53 = 99.$$

Следовательно, вычисления выполнены правильно. По формулам (1) и (2) находим

$$\overline{u} = \frac{-6}{53} \approx -0{,}113,$$

$$D_U^* = \frac{58 - (-6)^2/53}{53} \approx 1{,}108.$$

По формулам (4) и (5) окончательно вычисляем

$$\overline{x} \approx (-0{,}113)\cdot 4 + 148 \approx 147{,}548,$$

$$D_X^* \approx 4^2 \cdot 1{,}103 \approx 17{,}728. \ \triangleright$$

Для приведенных в задачах 19.38–19.40 выборок выполнить следующие задания:

1) вычислить среднее и дисперсию, предварительно проведя группировку выборки с заданной длиной интервала; для упрощения вычислений преобразовать данные по формуле (3);

2) вычислить среднее и дисперсию негруппированной выборки, используя заданные значения.

Сравнить результаты вычислений.

19.38. Положительные отклонения от номинального размера у партии деталей (в мм):

17	21	8	20	23	18	22	20	17	12
20	11	9	19	20	9	19	17	21	13
17	22	22	10	20	20	15	19	20	20
13	21	21	9	14	11	19	18	23	19

$$n = 40, \quad b = 2, \quad \sum x_i = 689, \quad \sum x_i^2 = 12\,635.$$

19.39. Время восстановления диодов из одной партии (в наносекундах):

69	73	70	68	61	73	70	72	67	70
66	70	76	68	71	71	68	70	64	65
72	70	70	69	66	70	77	69	71	74
72	72	72	68	70	67	71	67	72	69
66	75	76	69	71	67	70	73	71	74

$$n = 50, \quad b = 3, \quad \sum x_i = 3492, \quad \sum x_i^2 = 244\,342.$$

19.40. Время реакции (в секундах):

8,5	7,1	6,7	6,2	2,9	4,4	6,0	5,8	5,4
8,2	6,9	6,5	6,1	3,8	6,0	6,0	5,6	5,3
7,7	6,8	6,5	6,1	4,2	4,7	5,6	5,4	5,3
7,4	6,7	6,4	6,1	4,5	6,0	5,8	5,6	5,1

$$n = 36, \quad b = 1, \quad \sum x_i = 213{,}8, \quad \sum x_i^2 = 1316{,}82.$$

19.41. В условиях задачи 19.12 найти размах выборки, а также выборочные медиану, среднее и дисперсию. Сравнить полученные результаты с теоретическими значениями соответствующих характеристик.

19.42. Поезда метро идут строго по расписанию с интервалом в 3 минуты. Методом математического моделирования получить выборку времен ожидания поезда для 10 студентов, каждый из которых выходит на перрон в случайный момент времени. Найти размах выборки, выборочные медиану, среднее и дисперсию времени ожидания. Сравнить полученные результаты со значениями соответствующих теоретических характеристик.

19.43. Для получения значения случайной величины X, имеющей биномиальное распределение $B(n, p)$, можно воспользоваться следующим методом: получить n случайных чисел $y_1, y_2, \ldots, y_n$ и положить X равным числу случаев, когда $y_i < p$, $i = 1, 2, \ldots, n$. Методом математического моделирования найти число выигрышей в игре с игральным автоматом в 10 сеансах, если вероятность выигрыша в каждом сеансе равна 1/2. Получить выборку результатов для пяти серий игр по 10 сеансов. Найти выборочные моду, среднее и дисперсию числа выигрышей. Сравнить полученные результаты с теоретическими значениями соответствующих характеристик.

19.44. В условиях задачи 19.17 найти выборочные среднее и дисперсию и сравнить с соответствующими теоретическими значениями.

19.45. Определить среднее и дисперсию выборочного распределения из задачи 19.19.

19.46. Определить среднее и дисперсию выборочного распределения из задачи 19.16.

19.47. Проводятся испытания 10 приборов одного типа. Время наработки на отказ приборов данного типа подчиняется распределению $\mathrm{Ex}(\lambda)$. Методом математического моделирования найти выборочные среднее и медиану, а также минимальное и максимальное значения времени наработки на отказ при $\lambda = 1$. (Воспользоваться выборкой из задачи 19.13.)

Выборочные коэффициенты асимметрии и эксцесса вычисляются по формулам

$$a_X^* = \frac{\mu_3^*}{(D_X^*)^{3/2}}, \tag{6}$$

$$e_X^* = \frac{\mu_4^*}{(D_X^*)^2} - 3. \tag{7}$$

Вычисление выборочных коэффициентов асимметрии и эксцесса для группированной выборки значительно упрощается, если данные предварительно преобразовать по формуле (3). При этом используют следующие формулы:

$$a_X^* = \frac{1}{(D_U^*)^{3/2}} \left[\frac{\sum n_i u_i^3}{n} - 3\frac{\sum n_i u_i^2}{n}\frac{\sum n_i u_i}{n} + 2\left(\frac{\sum n_i u_i}{n}\right)^3 \right], \tag{8}$$

$$e_X^* = \frac{1}{(D_U^*)^2} \left[\frac{\sum n_i u_i^4}{n} - 4\frac{\sum n_i u_i^3}{n}\frac{\sum n_i u_i}{n} + 6\frac{\sum n_i u_i^2}{n}\left(\frac{\sum n_i u_i}{n}\right)^2 - 3\left(\frac{\sum n_i u_i}{n}\right)^4 \right] - 3. \tag{9}$$

Для контроля подсчета сумм $\sum n_i u_i^l$, $l = 1, 2, 3, 4$, вычисляют отдельно правую и левую части тождества

$$\sum n_i(u_i+1)^4 \equiv \sum n_i + 4\sum n_i u_i + 6\sum n_i u_i^2 + 4\sum n_i u_i^3 + \sum n_i u_i^4. \tag{10}$$

Результаты вычислений оформляют в виде таблицы.

Пример 7. Вычислить среднее, дисперсию, коэффициенты асимметрии и эксцесса для следующей группированной выборки:

Границы интервалов	10–12	12–14	14–16	16–18	18–20	20–22	22–24
Частоты	2	4	8	12	16	10	3

◁ Длина интервала группировки $b = 2$. Значение z_i, встречающееся с наибольшей частотой, $d_X^* = 19$. Поэтому преобразование (3) имеет вид

$$u_i = \frac{z_i - 19}{2}, \quad i = 1, 2, \ldots, 7.$$

Вычисления оформим в виде таблицы 1.3.

Таблица 1.3

i	z_i	u_i	n_i	n_iu_i	u_i^2	$n_iu_i^2$	u_i^2	$n_iu_i^3$	u_i^4	$n_iu_i^4$	u_i+1	$(u_i+1)^4$	$n_i(u_i+1)^4$
1	11	−4	2	−8	16	32	−64	−128	256	512	−3	81	162
2	13	−3	4	−12	9	36	−27	−108	81	324	−2	16	64
3	15	−2	8	−16	4	32	−8	−64	16	128	−1	1	8
4	17	−1	12	−12	1	12	−1	−12	1	12	0	0	0
5	19	0	16	0	0	0	0	0	0	0	1	1	16
6	21	1	10	10	1	10	1	10	1	10	2	16	160
7	23	2	3	6	4	12	8	24	16	48	3	81	243
$\sum$	–	–	55	−32	–	134	–	−278	–	1034	–	–	653
$\sum \frac{n_iu_i^l}{n}$			–	$-0,532$	–	2,436	–	$-5,054$	–	18,8	–	–	–

Контроль вычислений по тождеству (10) дает

$$55 + 4 \cdot (-32) + 6 \cdot 134 + 4 \cdot (-278) + 1034 = 653.$$

Наконец, воспользовавшись формулами (1), (2), (4), (5), (8) и (9), находим характеристики выборочного распределения:

$$x = 19 + 2 \cdot \frac{-32}{55} \approx 17{,}8,$$

$$D_U^* = \frac{134 - (-32)^2/55}{55} \approx 2{,}10, \quad D_X^* \approx 2^2 \cdot 2{,}10 = 8{,}40,$$

$$a_X^* = \frac{1}{2{,}10^{3/2}}[-5{,}054 - 3 \cdot (-0{,}582) \cdot 2{,}436 + 2 \cdot (-0{,}582)^3] \approx -0{,}393,$$

$$e_X^* = \frac{1}{2{,}10^2}[18{,}8 - 4 \cdot (-0{,}582) \cdot (-5{,}054) +$$
$$+ 6 \cdot (-0{,}582)^2 \cdot 2{,}436 - 3 \cdot (-0{,}582)^4] - 3 \approx -0{,}303. \ \triangleright$$

19.48. Из формул (6) и (7) для выборочных коэффициентов асимметрии и эксцесса получить соотношения (8) и (9) для их вычисления.

В задачах 19.49–19.54 вычислить среднее, дисперсию, коэффициенты асимметрии и эксцесса для заданных выборок, предварительно преобразовав их по формуле (3).

19.49. Распределение скорости автомобилей на одном из участков шоссе (км/ч):

Границы интервалов	61–65	65–69	69–73	73–77	77–81
Частоты	1	4	5	8	14

Границы интервалов	81–85	85–89	89–93	93–97	97–101
Частоты	9	6	1	1	1

19.50. Суммарное число набранных баллов в соревнованиях:

Границы интервалов	49–52	52–55	55–58	58–61	61–64	64–67	67–70
Частоты	3	6	11	19	30	21	10

19.51. Распределение предела прочности образцов сварного шва (Н/мм^2):

Границы интервалов	28–30	30–32	32–34	34–36	36–38	38–40	40–42	42–44
Частоты	8	15	15	12	15	20	10	5

19.52. Распределение отклонений напряжения от номинала (мВ):

Границы интервалов	0,00–0,02	0,02–0,04	0,04–0,06	0,06–0,08
Частоты	9	15	29	35

Границы интервалов	0,08–0,10	0,10–0,12	0,12–0,14	0,14–0,16
Частоты	32	19	8	3

19.53. Время выполнения упражнения (с):

Границы интервалов	8,95–9,05	9,05–9,15	9,15–9,25	9,25–9,35
Частоты	4	8	11	7

Границы интервалов	9,35–9,45	9,45–9,55	9,55–10,05
Частоты	5	3	2

19.54. Горизонтальное отклонение от цели (м) для 200 испытаний ракет:

Границы интервалов	−40 – −30	−30 – −20	−20 – −10	−10 – 0
Частоты	7	11	15	24

Границы интервалов	0–10	10–20	20–30	30–40	40–50	50–60
Частоты	49	41	26	17	7	3

Пусть $x^{(1)}, \ldots, x^{(n)}$ — вариационный ряд выборки объема n. Если np — не целое число, то *выборочной квантилью* x_p^* порядка p $(0 < p < 1)$ называется k-й член вариационного ряда, где $k = [np] + 1$. Если же np — целое число, то соответствующая квантиль x_p^* не определена (она может принимать любое значение из интервала $[x^{(k)},\ x^{(k+1)})$).

19.55. Вычислить выборочные квантили порядков $p = 0{,}1;\ 0{,}5;\ 0{,}9$ по данным, приведенным в задаче 19.40.

19.56. Методом моделирования получить выборку объема $n = 55$ из генеральной совокупности с распределением $N(0,\ 1)$. Определить выборочную медиану и выборочные квантили порядков $p = 0{,}1;\ 0{,}25;\ 0{,}75;\ 0{,}9$. Сравнить полученные результаты с соответствующими теоретическими значениями.

19.57 (продолжение). Решить предыдущую задачу для распределения $N(10,\ 2)$.

19.58 (продолжение). Решить задачу 19.56 для распределения $\mathrm{Ex}\,(2)$.

3. Статистическое описание и выборочные характеристики двумерного случайного вектора. Пусть $(x_i,\ y_i)$, $i = 1,\ 2,\ \ldots,\ n$, — выборка объема n из наблюдений случайного двумерного вектора $(X,\ Y)$. Предварительное представление о двумерной генеральной совокупности можно получить, изображая элементы выборки точками на плоскости с выбранной декартовой прямоугольной системой координат. Это представление выборки называется *диаграммой рассеивания*.

Распределением двумерной выборки называется распределение двумерного дискретного случайного вектора, принимающего значения $(x_i,\ y_i)$, $i = 1, 2, \ldots, n$, с вероятностями, равными $1/n$. *Выборочные числовые характеристики* вычисляются как соответствующие числовые характеристики двумерного случайного вектора дискретного типа (см. гл. 18, § 3, п. 1).

Пример 8. Вычислить выборочные средние, дисперсии и коэффициент корреляции для выборки, приведенной в таблице 1.4. Построить диаграмму рассеивания.

◁ Вычисление указанных выборочных характеристик удобно выполнять в следующей последовательности. Сначала вычисляют суммы

$$\sum x_i, \ \sum y_i, \ \sum x_i^2, \ \sum y_i^2, \ \sum x_i y_i, \ \sum (x_i + y_i)^2.$$

Для контроля правильности вычислений используется тождество

$$\sum (x_i + y_i)^2 \equiv \sum x_i^2 + 2 \sum x_i y_i + \sum y_i^2.$$

Таблица 1.4

x	y	x	y	x	y	x	y	x	y
8,35	3,50	10,50	6,00	11,35	9,50	12,15	6,00	12,85	9,50
8,74	1,49	10,75	2,50	11,50	6,00	12,25	8,05	13,15	9,02
9,25	6,40	10,76	5,74	11,50	9,00	12,35	5,01	13,25	6,49
9,50	4,50	11,00	8,50	11,62	8,50	12,50	7,03	13,26	10,50
9,75	5,00	11,00	5,26	11,75	10,00	12,76	7,53	13,40	7,51
10,24	7,00	11,25	8,00	12,00	9,00	12,85	6,01	13,50	10,00
13,65	9,50	14,50	10,00	13,75	8,51	14,75	12,00	14,00	11,00
15,25	12,50	14,23	8,40	16,00	11,50	14,26	10,00	16,00	13,00
14,51	9,50	16,25	12,00						

Выборочные средние отсюда находятся по формулам (см. также задачу 19.59)

$$\overline{x} = \alpha_{1,0}^* = \frac{1}{n} \sum x_i, \quad \overline{y} = \alpha_{0,1}^* = \frac{1}{n} \sum y_i. \tag{11}$$

Затем вычисляются суммы квадратов отклонений от среднего и произведений отклонений от средних:

$$Q_x = \sum (x_i - \overline{x})^2 = \sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}, \tag{12}$$

$$Q_y = \sum (y_i - \overline{y})^2 = \sum y_i^2 - \frac{\left(\sum y_i\right)^2}{n}, \tag{13}$$

$$Q_{xy} = \sum (x_i - \overline{x})(y_i - \overline{y}) = \sum x_i y_i - \frac{\left(\sum x_i\right)\left(\sum y_i\right)}{n}. \tag{14}$$

Отсюда

$$D_X^* = \frac{1}{n} Q_x, \quad D_Y^* = \frac{1}{n} Q_y,$$
$$r = \frac{\mu_{1,1}^*}{\sqrt{D_X^* D_Y^*}} = \frac{Q_{xy}}{\sqrt{Q_x Q_y}}. \tag{15}$$

Объем выборки $n = 42$. Предварительно вычислим

$$\sum x_i = 522{,}23, \quad \sum y_i = 336{,}41, \quad \sum x_i^2 = 6652{,}25,$$

$$\sum y_i^2 = 2987{,}80, \quad \sum x_i y_i = 4358{,}626.$$

Тогда по формуле (11)

$$\overline{x} = 12{,}434, \quad \overline{y} = 8{,}011.$$

По формулам (12)–(14) находим

$$Q_x = 6652{,}25 - \frac{522{,}23^2}{42} \approx 158{,}8182,$$
$$Q_y = 2987{,}80 - \frac{336{,}41^2}{42} \approx 292{,}5958,$$
$$Q_{xy} = 4358{,}626 - \frac{522{,}23 \cdot 336{,}41}{42} \approx 175{,}1912.$$

Окончательно из соотношений (15) получаем

$$D_X^* = \frac{158{,}8182}{42} \approx 3{,}7814, \quad D_Y^* = \frac{292{,}5958}{42} \approx 6{,}9666,$$

$$r = \frac{175{,}1912}{\sqrt{158{,}8182 \cdot 292{,}5958}} \approx 0{,}813.$$

Диаграмма рассеивания приведена на рис. 30. ▷

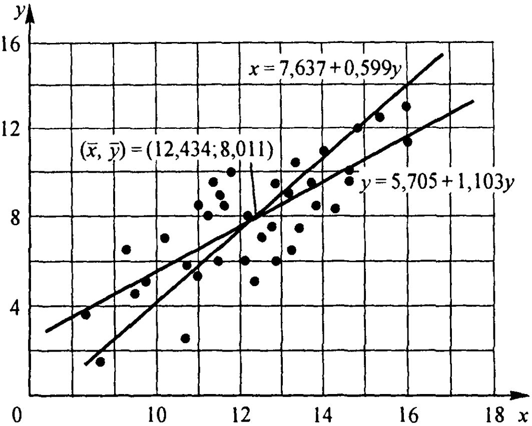


Рис. 30

19.59. Показать, что выборочные начальные и центральные моменты порядка $k+s$ $(k,\ s \geqslant 0)$ по выборке $(x_i,\ y_i)$, $i = 1, \ldots, n$, определяются формулами

$$\alpha^*_{k,\,s} = \frac{1}{n}\sum x_i^k y_i^s,$$

$$\mu^*_{k,\,s} = \frac{1}{n}\sum (x_i - \overline{x})^k (y_i - \overline{y})^s$$

и, в частности,

$$\alpha^*_{1,\,0} = \overline{x} = \frac{1}{n}\sum x_i, \quad \alpha^*_{0,\,1} = \overline{y} = \frac{1}{n}\sum y_i,$$

$$\mu^*_{2,\,0} = D^*_X = \frac{1}{n}\sum (x_i - \overline{x})^2, \quad \mu^*_{0,\,2} = D^*_Y = \frac{1}{n}\sum (y_i - \overline{y})^2,$$

$$\mu^*_{1,\,1} = k^*_{XY} = \frac{1}{n}\sum (x_i - \overline{x})\,(y_i - \overline{y}).$$

19.60. Показать, что для вычисления выборочной ковариации по двумерной выборке объема n можно использовать формулу

$$k^*_{XY} = \frac{1}{n}\sum x_i y_i - \overline{x}\,\overline{y}.$$

19.61. Показать, что для элементов выборки системы двух случайных величин выполняется равенство

$$\sum x_i (y_i - \overline{y}) = \sum y_i (x_i - \overline{x}).$$

19.62. Показать, что выборочный коэффициент корреляции по выборке (x_i, y_i), $i = 1, \ldots, n$, вычисляется по формуле

$$r = \frac{\sum (x_i - \overline{x})\,(y_i - \overline{y})}{\sqrt{\left(\sum (x_i - \overline{x})^2\right)\left(\sum (y_i - \overline{y})^2\right)}}.$$

Вычислить коэффициенты корреляции и построить диаграммы рассеивания для следующих выборок:

19.63.

x	8	10	5	8	9
y	1	3	1	2	3

19.64.

x	9	10	12	5
y	6	4	7	3

19.65.

x	10	2	7	5
y	8	2	6	4

19.66*. Известно, что для некоторой выборки $D_X^* = 16$, $D_Y^* = 9$. Каково наибольшее значение ковариации?

19.67. Пусть над элементами выборки системы двух случайных величин (x_i, y_i), $i = 1, 2, \ldots, n$, выполнено линейное преобразование $u_i = ax_i + b$, $v_i = cy_i + d$, $i = 1, 2, \ldots, n$. Показать, что выборочные ковариации и коэффициент корреляции связаны соотношениями

$$k_{UV}^* = ack_{XY}^* \quad (ac > 0), \quad r_{UV} = r_{XY}.$$

Используя подходящее линейное преобразование, вычислить выборочный коэффициент корреляции для следующих выборок:

19.68.

x	55	71	53	67	81	75	59	89	65	81
y	206	116	221	113	32	128	248	113	284	215

19.69.

x	65,8	68,3	72,7	66,1	73,1	71,8	73,1	66,5
y	166,0	115,2	157,8	152,5	149,3	181,0	173,2	120,4

x	69,3	73,4	67,3	73,6	67,9	69,7	69,7
y	124,5	163,2	125,2	173,3	146,7	157,9	134,5

Выборочная линейная регрессия Y на X по выборке (x_i, y_i), $i = 1, \ldots, n$, определяется уравнением

$$y = \beta_0^* + \beta_1^* x = \overline{y} + r\sqrt{\frac{D_Y^*}{D_X^*}}\,(x - \overline{x}).$$

Коэффициенты β_0^* и β_1^* называются *выборочными коэффициентами регрессии.* Они вычисляются по формулам

$$\beta_1^* = \frac{n\sum x_i y_i - \left(\sum x_i\right)\left(\sum y_i\right)}{n\sum x_i^2 - \left(\sum x_i\right)^2} = \frac{Q_{xy}}{Q_x}, \tag{16}$$

$$\beta_0^* = \overline{y} - \beta_1^* \overline{x}. \tag{17}$$

Аналогично определяется выборочная линейная регрессия X на Y:

$$x = \beta_0'^* + \beta_1'^* y = \overline{x} + r\sqrt{\frac{D_X^*}{D_Y^*}}\,(y - \overline{y}),$$

коэффициенты $\beta_0'^*$ и $\beta_1'^*$ которой находятся по формулам

$$\beta_1'^* = \frac{n\sum x_i y_i - \left(\sum x_i\right)\left(\sum y_i\right)}{n\sum y_i^2 - \left(\sum y_i\right)^2} = \frac{Q_{xy}}{Q_y}, \tag{18}$$

$$\beta_0'^* = \overline{x} - \beta_1'^* y. \tag{19}$$

Для контроля правильности расчетов используют соотношение

$$\sqrt{\beta_1^* \beta_1'^*} = |\, r\,|. \tag{20}$$

Прямые

$$y = \beta_0^* + \beta_1^* x, \quad x = \beta_0'^* + \beta_1'^* y$$

пересекаются в точке с координатами $(\overline{x}, \overline{y})$.

Пример 9. Вычислить выборочные коэффициенты линейной регрессии X на Y и Y на X по выборке примера 8 (таблица 1.4). Нанести прямые регрессии на диаграмму рассеивания.

◁ Воспользуемся результатами вычислений в примере 8. По формулам (16) и (17) находим

$$\beta_1^* = \frac{175{,}1912}{158{,}8182} \approx 1{,}103, \quad \beta_0^* = 8{,}011 - 1{,}103 \cdot 12{,}434 \approx -5{,}705.$$

Таким образом, прямая регрессии Y на X имеет уравнение

$$y = -5{,}705 + 1{,}103x.$$

Аналогично по формулам (18), (19) находим

$$\beta_1'^* = \frac{175{,}1912}{292{,}5958} \approx 0{,}599, \quad \beta_0'^* = 12{,}434 - 0{,}599 \cdot 8{,}011 \approx 7{,}637.$$

Отсюда прямая регрессии X на Y имеет уравнение

$$x = 7{,}637 + 0{,}599y.$$

Проверка по формуле (20) дает

$$\sqrt{1{,}103 \cdot 0{,}599} \approx 0{,}813,$$

что совпадает со значением r, вычисленным в примере 8. Прямые регрессии нанесены на диаграмму рассеивания на рис. 30. ▷

В задачах 19.70–19.72 вычислить коэффициенты корреляции, определить и нанести на диаграмму рассеивания прямые регрессии Y на X и X на Y по данным выборкам.

19.70.

x	8	10	5	8	9
y	1	3	1	2	3

19.71.

x	9	10	12	5
y	6	4	7	3

19.72.

x	10	2	7	5
y	8	2	6	4

19.73. Предел выносливости стали при изгибе Y (Н/мм^2) оценивается на основании другой ее характеристики — предела упругости при кручении X (Н/мм^2). По опытным данным для 12 марок стали найти уравнения линейной регрессии Y на X и X на Y и вычислить коэффициент корреляции между этими характеристиками. Результаты измерений:

x	51	67	84	81	101	109	71	97	109	51	105	89
y	25	30	43	44	57	58	43	46	62	45	55	45

$$\sum x_i = 1015, \quad \sum y_i = 553, \quad \sum x_i^2 = 90667,$$

$$\sum y_i^2 = 26807, \quad \sum x_i y_i = 48888.$$

19.74. По данным измерений двух переменных

x	66	70	75	80	82	85	90	92	95	98
y	60	78	65	87	74	70	78	95	88	90

вычислить коэффициент корреляции и найти уравнение линейной регрессии Y на X.

19.75. Элементы выборки системы двух случайных величин преобразованы так же, как в задаче 19.67 ($ac > 0$). Как изменятся выборочные коэффициенты β_0^* и β_1^* линейной регрессии?

19.76 (продолжение). Записать уравнения регрессии для выборки (u_i, v_i), $i = 1, \ldots, n$, где

$$u_i = \frac{1}{\sqrt{D_X^*}}(x_i - \overline{x}), \quad v_i = \frac{1}{\sqrt{D_Y^*}}(y_i - \overline{y}).$$

19.77. По данным 1953 г. [2]) количество телевизионных точек и численность населения в десяти городах США характеризовались следующими числами (в десятках тысяч), приведенными в таблице 1.5. Нанести данные на диаграмму рассеивания, вычислить коэффициенты выборочной корреляции: а) для первых девяти городов (без Нью-Йорка), б) для всех десяти городов. Сравнить результаты вычислений. Найти коэффициенты линейной регрессии Y на X для десяти городов и нанести уравнение регрессии на диаграмму рассеивания. Объясните полученные результаты.

Таблица 1.5

Города	Денвер	Сан-Антонио	Канзас-Сити	Сиэтл	Цинциннати	Буффало	Новый Орлеан	Милуоки	Хьюстон	Нью-Йорк
Население, x	45	46	47	48	51	58	59	65	67	802
Количество установленных телезизоров y	12	12	29	25	38	35	16	43	22	345

19.78. Пусть Z и X — независимые случайные величины с распределениями $N(0, 1)$ и $N(m_X, \sigma_X)$ соответственно. Доказать, что случайная величина Y, связанная с Z и X соотношением

$$Y = m_Y + \rho_{XY}\frac{\sigma_Y}{\sigma_X}(X - m_X) + \sqrt{1 - \rho_{XY}^2}\,\sigma_Y Z,$$

имеет нормальное распределение $N(m_Y, \sigma_Y)$, причем коэффициент корреляции между X и Y равен ρ_{XY}.

19.79. Используя метод моделирования и результат предыдущей задачи, получить выборку объема $n = 20$ из двумерной нормально распределенной генеральной совокупности с параметрами $m_X = 1$, $\sigma_X = 1$, $m_Y = 4$, $\sigma_Y = 1$, $\rho_{XY} = 0{,}8$. Предварительно вычислив выборочные коэффициенты β_0^* и β_1^* линейной регрессии Y на X, нанести полученные данные и прямую регрессии

[2]) *Миллс Ф.* Статистические методы.—М.: Госстатиздат, 1958.

$y = \beta_0^* + \beta_1^* x$ на диаграмму рассеивания. Найти выборочное среднее и дисперсию остатков, т.е. разностей

$$(y_i - (\beta_0^* + \beta_1^* x_i)), \quad i = 1, 2, \ldots, 20.$$

Двумерную выборку *большого объема* представляют в виде *корреляционной таблицы.* С этой целью группируют реализации величин X и Y по интервалам длины b_x и b_y, а в клетки таблицы записывают число пар исходной выборки (т.е. частоты) для каждой комбинации интервалов. Эту процедуру можно также выполнять непосредственно по диаграмме рассеивания, нанося на нее сетку горизонтальных и вертикальных прямых, взятых с постоянными шагами b_x и b_y. Наблюдения, которые попали на верхнюю и правую границы рассматриваемого прямоугольника, относятся соответственно к соседним верхнему и правому прямоугольникам. В дальнейших вычислениях используются середины интервалов и соответствующие частоты. Обозначим середины интервалов через $\hat{x}_i$, $i = 1, 2, \ldots, k$ и $\hat{y}_j$, $j = 1, 2, \ldots, l$, а соответствующие частоты через n_{ij}; очевидно, $\sum_{i=1}^{k}\sum_{j=1}^{l} n_{ij} = n$.

Полагаем

$$\sum_{i=1}^{k} n_{ij} = n_{\cdot j}, \qquad \sum_{j=1}^{l} n_{ij} = n_{i\cdot}.$$

Для упрощения вычислений вместо середин интервалов $\hat{x}_i$ и $\hat{y}_j$ введем числа

$$u_i = \frac{\hat{x}_i - d_X^*}{b_x}, \quad i = 1, 2, \ldots, k,$$
$$v_j = \frac{\hat{y}_j - d_Y^*}{b_y}, \quad j = 1, 2, \ldots, l,$$

где d_X^* и d_Y^* — середины наиболее часто встречающихся интервалов.

Определение выборочных числовых характеристик распределения по корреляционной таблице выполняется в следующей последовательности. Сначала вычисляют суммы $\sum_{i=1}^{k} n_{i\cdot} u_i$, $\sum_{j=1}^{l} n_{\cdot j} v_j$, $\sum_{i=1}^{k} n_{i\cdot} u_i^2$, $\sum_{j=1}^{l} n_{\cdot j} v_j^2$, $\sum_{i=1}^{k}\sum_{j=1}^{l} n_{ij} u_i v_j$. Затем определяют следующие суммы:

$$Q_u = \sum n_{i\cdot} u_i^2 - \frac{\left(\sum n_{i\cdot} u_i\right)^2}{n}, \tag{21}$$

$$Q_v = \sum n_{\cdot j} v_j^2 - \frac{\left(\sum n_{\cdot j} v_j\right)^2}{n}, \tag{22}$$

$$Q_{uv} = \sum_{i=1}^{k}\sum_{j=1}^{l} n_{ij} u_i v_j - \frac{\left(\sum n_{i\cdot} u_i\right)\left(\sum n_{\cdot j} v_j\right)}{n}. \tag{23}$$

Выборочные средние, дисперсии и коэффициент корреляции находят по формулам

$$\overline{x} = b_x \frac{\sum n_{i\cdot} u_i}{n} + d_X, \quad \overline{y} = b_y \frac{\sum n_{\cdot j} v_j}{n} + d_Y, \tag{24}$$

$$D_X^* = b_x^2 \frac{Q_u}{n}, \quad D_Y^* = b_y^2 \frac{Qv}{n}, \tag{25}$$

$$r = \frac{Q_{uv}}{\sqrt{Q_u Q_v}}. \tag{26}$$

Коэффициенты β_1^* и $\beta_1'^*$ линейной регрессии Y на X и X на Y вычисляют по формулам

$$\beta_1^* = \frac{b_y}{b_x} \frac{Q_{uv}}{Q_u}, \tag{27}$$

$$\beta_1'^* = \frac{b_x}{b_y} \frac{Q_{uv}}{Q_v}. \tag{28}$$

Коэффициенты β_0^* и $\beta_0'^*$ находят по формулам (17) и (19).

Пример 10. Проведя группировку выборки примера 7, вычислить выборочные средние, дисперсии, коэффициент корреляции, а также выборочные коэффициенты линейной регрессии X на Y и Y на X.

$\triangleleft$ Выберем $b_x = 1$, $b_y = 2$. Прямоугольная сетка, соответствующая этим значениям, нанесена на диаграмму рассеивания (рис. 30). Непосредственно по диаграмме строим корреляционную таблицу (таблица 1.6). Находим $d_X^* = 11{,}5$, $d_Y^* = 9$ и вычисляем значения u_i и v_j по формулам

$$u_i = \frac{\hat{x}_i - 11{,}5}{1}, \quad i = 1, 2, \ldots, 9,$$

$$v_j = \frac{\hat{y}_i - 9}{2}, \qquad j = 1, 2, \ldots, 7.$$

Вычисляем следующие суммы:

$$\sum n_{i\cdot} u_i = 43, \quad \sum n_{\cdot j} v_j = -15, \quad \sum n_{i\cdot} u_i^2 = 215,$$

$$\sum n_{\cdot j} v_j^2 = 87, \quad \sum_{i=1}^{9} \sum_{j=1}^{7} n_{ij} u_i v_j = 80.$$

По формулам (21)–(23) находим

$$Q_u = 215 - \frac{43^2}{42} \approx 170{,}976, \quad Q_v = 87 - \frac{(-15)^2}{42} \approx 81{,}643,$$

$$Q_{uv} = 80 - \frac{43 \cdot (-15)}{42} \approx 95{,}357.$$

Таблица 1.6

Границы и середины интервалов для y	v_j	Границы и середины интервалов для x									$n_{\cdot j}$	$n_{\cdot j}v_j$	$n_{\cdot j}v_j^2$
		8–9 8,5	9–10 9,5	10–11 10,5	11–12 11,5	12–13 12,5	13–14 13,5	14–15 14,5	15–16 15,5	16–17 16,5			
		u_i											
		−3	−2	−1	0	1	2	3	4	5			
0–2 1	−4	1	0	0	0	0	0	0	0	0	1	−4	16
2–4 3	−3	1	0	1	0	0	0	0	0	0	2	−6	18
4–6 5	−2	0	2	1	1	1	0	0	0	0	5	−10	20
6–8 7	−1	0	1	2	1	4	2	0	0	0	10	−10	10
8–10 9	0	0	0	0	5	3	3	2	0	0	13	0	0
10–12 11	1	0	0	0	1	0	2	3	0	1	7	7	7
12–14 13	2	0	0	0	0	0	0	1	1	2	4	8	16
$n_{i\cdot}$		2	3	4	8	8	7	6	1	3	$\sum = 42$	$\sum = -15$	$\sum = 87$
$n_{i\cdot}u_j$		2	3	4	8	8	7	6	1	3	$\sum = 43$		
$n_{i\cdot}u_j^2$		2	3	4	8	8	7	6	1	3	$\sum = 215$		

По формулам (24)–(26) вычисляем

$$\overline{x} = 1 \cdot \frac{43}{42} + 11{,}5 \approx 12{,}52, \quad \overline{y} = 2 \cdot \frac{(-15)}{42} + 9 \approx 8{,}28,$$

$$D_X^* = 1^2 \cdot \frac{170{,}976}{42} \approx 4{,}071, \quad D_Y^* = 2^2 \cdot \frac{81{,}643}{42} \approx 7{,}775,$$

$$r = \frac{95{,}357}{\sqrt{170{,}976 \cdot 81{,}643}} \approx 0{,}807.$$

Выборочные коэффициенты регрессии вычисляем по формулам (27), (28), (17) и (19):

$$\beta_1^* = \frac{2}{1} \cdot \frac{95{,}357}{170{,}976} \approx 1{,}12, \qquad \beta_1'^* = \frac{1}{2} \cdot \frac{95{,}357}{81{,}643} \approx 0{,}58,$$

$$\beta_0^* = 8{,}28 - 1{,}12 \cdot 12{,}52 \approx -5{,}74, \quad \beta_0'^* = 12{,}52 - 0{,}58 \cdot 8{,}28 \approx 7{,}72.$$

Таким образом, уравнение линейной регрессии Y на X имеет вид

$$y = -5{,}74 + 1{,}12x,$$

а уравнение линейной регрессии X на Y имеет вид

$$x = 7{,}72 + 0{,}58y.$$

Расхождение полученных результатов с результатами примеров 8 и 9 обусловлено группировкой. ▷

19.80. В таблице 1.7 приводятся результаты лабораторного анализа 64 образцов сланцевых пород на содержание двуокиси кремния (SiO_2) и двуокиси алюминия (Al_2O_3) (в условных единицах). Вычислить коэффициент корреляции между этими признаками, предварительно сгруппировав эти данные в корреляционную таблицу.

Таблица 1.7

SiO_2	Al_2O_3	SiO_2	Al_2O_3	SiO_2	Al_2O_3	SiO_2	Al_2O_3
57,8	17,2	53,8	16,3	48,8	16,4	50,7	21,5
54,6	17,9	53,6	17,2	53,5	15,9	53,1	21,3
54,8	18,8	51,5	15,8	52,8	15,9	52,9	20,3
51,7	19,9	54,0	15,0	52,9	14,8	51,3	20,1
61,1	16,0	50,4	14,4	52,1	19,8	52,7	17,2
62,3	17,8	53,0	15,3	47,3	18,7	46,6	15,6
52,2	18,8	53,3	16,6	49,8	20,2	46,5	16,0
49,2	19,3	51,6	14,9	49,3	17,6	51,3	15,5
53,9	16,1	50,9	14,7	50,1	19,2	51,0	19,2
60,0	14,8	49,6	16,1	54,4	18,2	47,5	18,5
56,2	17,0	52,2	19,5	49,0	16,8	47,7	19,0
55,2	17,8	50,5	15,6	48,9	18,2	44,9	16,6
53,3	19,9	51,1	18,1	51,3	19,7	49,4	16,0
57,9	17,1	52,2	19,5	51,6	19,6	48,9	18,6
54,0	15,5	49,2	15,7	46,2	19,1	48,8	19,4
52,6	17,6	49,3	13,2	50,4	20,2	50,6	18,9

Вычислить коэффициент корреляции и найти уравнения прямых регрессии Y на X и X на Y по данным в следующих корреляционных таблицах (задачи 19.81–19.84):

19.81.

y	x			
	40–50	50–60	60–70	70–80
10–11	2	11	3	2
11–12	1	19	2	4
12–13	3	6	27	6
13–14	2	3	3	8

19.82.

y	x					
	5–15	15–25	25–35	35–45	45–55	55–65
10–20	5	7	0	0	0	0
20–30	0	20	23	0	0	0
30–40	0	0	30	47	2	0
40–50	0	0	10	11	20	6
50–60	0	0	0	9	7	3

19.83.

y	x						
	30–50	50–70	70–90	90–110	110–130	130–150	150–170
50–70	5	0	0	0	0	0	0
70–90	2	3	4	0	0	0	0
90–110	0	1	7	6	0	0	0
110–130	0	0	1	8	4	0	0
130–150	0	0	1	1	5	2	0
150–170	0	0	0	0	0	5	0
170–190	0	0	0	0	0	0	2
190–210	0	0	0	0	0	0	2

19.84.

y	x				
	7,0–7,2	7,2–7,4	7,4–7,6	7,6–7,8	7,8–8,0
2,15–2,45	5	4	0	0	0
2,45–2,75	0	12	8	1	0
2,75–3,05	0	0	5	5	0
3,05–3,35	0	0	4	7	0
3,35–3,65	0	0	0	12	1
3,65–3,95	0	0	0	0	1

В задачах 19.85–19.90 найти числовые характеристики системы двух случайных величин и коэффициенты регрессии Y на X и X на Y по данным выборкам.

19.85.

x	y
0	0,321
0,1	0,282
0,2	0,266
0,3	0,245
0,4	0,232
0,5	0,227
0,6	0,214
0,7	0,208
0,8	0,201
0,9	0,200
1,0	0,199
1,1	0,202
1,2	0,203
1,3	0,209
1,4	0,215
1,5	0,223
1,6	0,234
1,7	0,262

19.86.

x	y
1,0	0,22
1,5	0,23
2,0	0,31
2,5	0,43
3,0	0,56
3,5	0,82
4,0	1,06
4,5	1,25
5,0	1,72
5,5	2,28
6,0	2,67
6,5	3,26
7,0	3,72
7,5	4,32
8,0	5,11
8,5	5,98
9,0	6,64
9,5	7,02
10,0	8,32

19.87.

x	y
0,00	15
0,05	16
0,10	24
0,15	17
0,20	26
0,25	30
0,30	31
0,35	36
0,40	32
0,45	40
0,50	42
0,55	41
0,60	51
0,65	45
0,70	46
0,75	53
0,80	57
0,85	62
0,90	59
0,95	63

19.88.

x	y
0	0,051
1	0,082
2	0,151
3	0,240
4	0,352
5	0,383
6	0,448
7	0,502
8	0,692
9	0,702
10	0,798
11	0,875
12	0,956
13	1,160
14	1,249
15	1,482
16	1,723
17	1,986

19.89.

x	y
3,0	0,422
3,2	0,405
3,4	0,415
3,6	0,438
3,8	0,467
4,0	0,521
4,2	0,546
4,4	0,600
4,6	0,640
4,8	0,712
5,0	0,832
5,2	0,871
5,4	0,988
5,6	1,268
5,8	1,408
6,0	1,500

19.90.

x	y
0,1	0,96
0,2	0,83
0,3	0,81
0,4	0,90
0,5	1,11
0,6	1,12
0,7	1,32
0,8	1,60
0,9	1,70
1,0	1,90
1,1	2,36
1,2	2,77
1,3	3,28
1,4	3,31
1,5	4,35
1,6	4,78

19.91. Построить диаграммы рассеяния для каждого из четырех приведенных в таблице (с. 218) множеств данных [3]. Для каждого набора данных найдите и нанесите на диаграмму рассеяния график линейной регрессии y и x. Прокомментируйте полученные результаты.

19.92. Получить выборку объема $n = 100$ из генеральной совокупности, имеющей биномиальное распределение $D\,(10;\ 0{,}3)$. Полученные данные представить в виде статистического ряда. Построить эмпирическую функцию распределения и гистограмму частот. Найти выборочные среднее, дисперсию, коэффициенты асимметрии и эксцесса.

19.93. Получить выборку объема $n = 100$ из генеральной совокупности, имеющей нормальное распределение $N(5,\ 2)$. Выполнить группировку полученных данных. Построить гистограмму частот. Найти выборочные среднее, дисперсию, коэффициенты асимметрии и эксцесса.

19.94. Решить предыдущую задачу по данным 100 реализаций случайной величины X, имеющей экспоненциальное распределение $\mathrm{Ex}\,(3)$.

[3]) *Anscombe F.J.* Graphs in Statistical Analysis, American Statistician, 17–21, 1973.

Таблица к задаче 19.91.

№п/п	X_1	Y_1	X_2	Y_2	X_3	Y_3	X_4	Y_4
1	10	8,04	10	9,14	10	7,46	8	6,58
2	8	6,95	8	8,14	8	6,77	8	5,76
3	13	7,58	13	8,74	13	12,74	8	7,71
4	9	8,81	9	8,77	9	7,114	8	8,84
5	11	8,33	11	9,26	11	7,814	8	8,47
6	14	9,96	14	8,10	14	8,84	8	7,04
7	6	7,24	6	6,13	6	6,08	8	5,25
8	4	4,26	4	3,10	4	5,39	19	12,5
9	12	10,84	12	9,13	12	8,15	8	5,96
10	7	4,82	7	7,26	7	6,42	8	7,91
11	5	5,68	5	4,74	5	5,73	8	6,89
$\overline{x}$	9,0		9,0		9,0		9,0	
$\overline{y}$	7,50		7,50		7,50		7,50	
Q_x	110,0		110,0		110,0		110,0	
Q_y	41,27		41,27		41,23		41,23	
Q_y	55,01		55,00		54,97		54,99	

19.95. Получить выборку объема $n = 100$ из двумерной нормально распределенной генеральной совокупности с параметрами $m_X = 3$, $m_Y = 5$, $\sigma_X^2 = 1$, $\sigma_Y^2 = 2$, $\rho_{XY} = 0{,}8$. Найти выборочные средние, дисперсии и коэффициент корреляции. Вычислить коэффициенты линейной регрессии X на Y и Y на X.

§ 2. Статистическое оценивание характеристик распределения генеральной совокупности по выборке

1. Точечные оценки и их свойства. Метод подстановки. Основная задача математической статистики состоит в нахождении распределения наблюдаемой случайной величины X по данным выборки. Во многих случаях вид распределения X можно считать известным, и задача сводится к получению приближенных значений неизвестных параметров этого распределения. Пусть $F_X(x, \theta)$ — функция распределения случайной величины X, содержащая один неизвестный параметр θ, а $x_1, x_2, \ldots, x_n$ — выборка наблюдений этой случайной величины. *Точечной оценкой* $\tilde{\theta}$ неизвестного параметра θ называется приближенное значение этого параметра, полученное по выборке.

Очевидно, что оценка $\tilde{\theta}$ есть значение некоторой функции элементов выборки, т.е. $\tilde{\theta} = \tilde{\theta}(x_1, x_2, \ldots, x_n)$. Любую функцию элементов выборки называют *статистикой. Чтобы выяснить, какие свойства должна иметь статистика* $\tilde{\theta}(x_1, x_2, \ldots, x_n)$ *для того, чтобы ее значения могли бы считаться хорошей в некотором смысле оценкой параметра* θ, *ее рассматривают как функцию случайного вектора* $(X_1, X_2, \ldots, X_n)$, *одной из реализаций которого является данная выборка* $x_1, x_2, \ldots, x_n$. Так как закон распределения каждой из случайных величин X_i, $i = 1, 2, \ldots, n$, есть $F_X(x, \theta)$, являющаяся функцией параметра θ, то и распределение статистики $\tilde{\theta}(x_1, x_2, \ldots, x_n)$ также зависит от неизвестного параметра θ.

Качество оценок характеризуется следующими основными свойствами:

1. Состоятельность. Оценка $\tilde{\theta} \equiv \tilde{\theta}_n = \tilde{\theta}(x_1, \ldots, x_n)$ называется *состоятельной оценкой* параметра θ, если $\tilde{\theta}_n$ сходится по вероятности к θ при $n \to \infty$. Последнее означает, что $\forall \varepsilon > 0$

$$\mathrm{P}[|\tilde{\theta}_n - \theta| < \varepsilon] \to 1$$

при $n \to \infty$.

Состоятельность оценки $\tilde{\theta}_n$ во многих случаях может быть установлена с помощью следующей теоремы.

Теорема 1. *Если* $\mathrm{M}[\tilde{\theta}_n] \to \theta$ *и* $\mathrm{D}[\tilde{\theta}_n] \to 0$ *при* $n \to \infty$, *то* $\tilde{\theta}_n$ — *состоятельная оценка параметра* θ.

2. Несмещенность. Оценка $\tilde{\theta}$ называется *несмещенной оценкой* параметра θ, если ее математическое ожидание равно оцениваемому параметру, т.е. $\mathrm{M}[\tilde{\theta}] = \theta$.

Разность $\mathrm{M}[\tilde{\theta}] - \theta$ называется *смещением.* Для несмещенных оценок систематическая ошибка оценивания равна нулю.

Простейший метод статистического оценивания — *метод подстановки* или *аналогии* — состоит в том, что в качестве оценки той или иной числовой характеристики (среднего, дисперсии и др.) генеральной совокупности берут соответствующую характеристику распределения выборки — выборочную характеристику.

Пример 1. Пусть $x_1, x_2, \ldots, x_n$ — выборка из генеральной совокупности с конечными математическим ожиданием m и дисперсией σ^2. Используя метод подстановки, найти оценку m. Проверить несмещенность и состоятельность полученной оценки.

$\lhd$ По методу подстановки в качестве оценки $\tilde{m}$ математического ожидания надо взять математическое ожидание распределения выборки —

выборочное среднее. Таким образом, получаем

$$\tilde{m} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Чтобы проверить несмещенность и состоятельность выборочного среднего как оценки m, рассмотрим эту статистику как функцию выборочного вектора $(X_1, \ldots, X_n)$. По определению выборочного вектора имеем: $\mathrm{M}[X_i] = m$ и $\mathrm{D}[X_i] = \sigma^2$, $i = 1, 2, \ldots, n$, причем X_i — независимые в совокупности случайные величины. Следовательно,

$$\mathrm{M}[\overline{X}] = \mathrm{M}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mathrm{M}[X_i] = \frac{1}{n}nm = m,$$

$$\mathrm{D}[\overline{X}] = \mathrm{D}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{D}[X_i] = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}.$$

Отсюда по определению получаем, что $\overline{X}$ — несмещенная оценка m, и так как $\mathrm{D}[\overline{X}] \to 0$ при $n \to \infty$, то в силу теоремы 1 $\overline{X}$ является состоятельной оценкой математического ожидания m генеральной совокупности. ▷

19.96. Пусть $x_1, x_2, \ldots, x_n$ — выборка наблюдений случайной величины X. Используя метод подстановки, найти оценки следующих числовых характеристик случайной величины X: дисперсии, медианы, асимметрии, эксцесса.

19.97*. Пусть $x_1, x_2, \ldots, x_n$ — выборка из генеральной совокупности с конечным начальным моментом α_{2l}. Используя метод подстановки, найти оценку начального момента α_l. Показать, что полученная оценка является несмещенной и состоятельной.

19.98*. Предположим, что выборка $x_1, x_2, \ldots, x_n$ получена из генеральной совокупности с конечными математическим ожиданием m и дисперсией σ^2. Показать, что выборочная дисперсия

$$D_X^* = \frac{1}{n}\sum (x_i - \overline{x})^2$$

является смещенной оценкой дисперсии генеральной совокупности, и найти ее смещение.

19.99 (продолжение). В условиях предыдущей задачи показать, что несмещенная оценка дисперсии генеральной совокупности задается статистикой

$$s^2 = \frac{1}{n-1}\sum (x_i - \overline{x})^2.$$

19.100.** Показать, что оценки D_X^* и s^2, полученные в задачах 19.98 и 19.99 соответственно, являются состоятельными оценками дисперсии генеральной совокупности.

19.101. Пусть $x_1, \ldots, x_n$ — выборка из генеральной совокупности с известным средним m и неизвестной дисперсией σ^2. Показать, что несмещенной оценкой σ^2 будет статистика

$$s_0^2 = \frac{1}{n}\sum(x_i - m)^2.$$

19.102. В качестве оценки математического ожидания m генеральной совокупности по выборке $x_1, \ldots, x_n$ предлагается взять статистику $\tilde{m}_1 = x_1$. Проверить несмещенность и состоятельность этой оценки.

19.103. Рассмотрим две выборки объемов n_1 и n_2 из одной генеральной совокупности со средним m и дисперсией σ^2. Пусть $\overline{X}_1$, $\overline{X}_2$, S_1^2 и S_2^2 — несмещенные оценки средних и дисперсий, определенные по этим выборкам. Показать, что объединенные оценки, вычисляемые по формулам

$$\overline{X} = \frac{n_1\overline{X}_1 + n_2\overline{X}_2}{n_1 + n_2},$$

$$S^2 = \frac{(n_1 - 1)\,S_1^2 + (n_2 - 1)\,S_2^2}{n_1 + n_2 - 2},$$

будут несмещенными и состоятельными оценками m и σ^2.

19.104*. Случайная величина X имеет распределение с плотностью $f_X(x)$, равной e^{a-x} при $x \geqslant a$ и 0 при $x < a$. Для оценки неизвестного параметра a по выборке $x_1, x_2, \ldots, x_n$ наблюдений случайной величины X предлагается выбрать статистику

$$\tilde{a} = x^{(1)} = \min_{1\leqslant i\leqslant n} x_i.$$

Проверить несмещенность и состоятельность этой оценки.

19.105*. Пусть $x_1, x_2, \ldots, x_n$ — выборка из генеральной совокупности, имеющей равномерное распределение $R\,(0,\,1)$. Показать что статистика

$$\tilde{m} = \frac{1}{2}\,(x^{(1)} + x^{(n)}),$$

где $x^{(1)}$ и $x^{(n)}$ — соответственно наименьший и наибольший элементы выборки, является несмещенной и состоятельной оценкой математического ожидания m.

19.106*. Пусть $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ — выборка из двумерной генеральной совокупности. Методом подстановки найти оценку ковариации. Показать, что полученная оценка является смещенной и состоятельной. Найти несмещенную оценку.

19.107*. Пусть $\tilde{\theta}$ — несмещенная оценка параметра θ, $\mathrm{D}[\tilde{\theta}] < \infty$. Показать, что $\tilde{\theta}^2$ является смещенной оценкой θ^2, и вычислить смещение.

19.108. Показать, что выборочное среднее, вычисленное по выборке из генеральной совокупности, имеющей распределение Пуассона с параметром λ, будет несмещенной и состоятельной оценкой этого параметра.

19.109. В результате проведения n независимых экспериментов в одних и тех же условиях случайное событие A произошло x раз.

а) Показать, что относительная частота $h = \dfrac{x}{n}$ появления события A будет несмещенной и состоятельной оценкой вероятности события A: $\mathrm{P}(A) = p$ в одном эксперименте.

б) Определить такое значение p, при котором дисперсия σ^2 будет максимальна.

19.110. Пусть $\tilde{\theta}$ — состоятельная оценка параметра θ, а $\psi(\tilde{\theta})$ — непрерывная функция в области изменения $\tilde{\theta}$. Доказать, что $\psi(\tilde{\theta})$ — состоятельная оценка $\psi(\theta)$.

19.111*. Пусть $x_1, x_2, \ldots, x_n$ — выборка из генеральной совокупности с конечным математическим ожиданием m и дисперсией σ^2. Показать, что

$$s = \sqrt{\frac{1}{n-1}\sum (x_i - \overline{x})^2}$$

является смещенной и состоятельной оценкой параметра σ.

19.112*. Доказать теорему 1 о состоятельности оценки.

Для оценки параметра θ может быть предложено несколько несмещенных оценок. Мерой точности несмещенной оценки $\tilde{\theta}$ считают ее дисперсию $\mathrm{D}[\tilde{\theta}]$.

Пусть $\tilde{\theta}_1$ и $\tilde{\theta}_2$ — две различные несмещенные оценки параметра θ. Если $\mathrm{D}[\tilde{\theta}_1] < \mathrm{D}[\tilde{\theta}_2]$, то говорят, что оценка $\tilde{\theta}_1$ *более эффективна*, чем оценка $\tilde{\theta}_2$.

В предположении, что распределение случайной величины X и статистика $\tilde{\theta}$ удовлетворяют некоторым условиям регулярности (A*) [4], для дисперсии несмещенной оценки $\tilde{\theta}$ параметра θ выполняется *неравенство Крамера-Рао*:

$$\mathrm{D}[\tilde{\theta}] \geqslant \frac{1}{I_n(\theta)}, \tag{1}$$

[4]) См., например: *Чистяков В. И.* Курс теории вероятностей.— М.: Наука, 1982. С. 180.

где $I_n(\theta)$ — *информация Фишера*, содержащаяся в выборке объема n относительно неизвестного параметра θ. Для непрерывной случайной величины X с плотностью распределения $f_X(x,\theta)$

$$I_n(\theta) = n\mathrm{M}\left[\left(\frac{\partial}{\partial\theta}\ln f_X(X,\theta)\right)^2\right]. \tag{2}$$

Если же X — дискретная случайная величина, то

$$I_n(\theta) = n\mathrm{M}\left[\left(\frac{\partial}{\partial\theta}\ln p(X,\theta)\right)^2\right], \tag{3}$$

где $p(x,\theta) = \mathrm{P}[X = x]$.

Условия регулярности (A*) выполняются для обычно используемых статистик нормального, биномиального и пуассоновского распределений.

Несмещенная оценка $\tilde{\theta}_0$ параметра θ, дисперсия которой достигает своего наименьшего возможного значения $\dfrac{1}{I_n(\theta)}$, называется *эффективной*:

$$\mathrm{D}[\tilde{\theta}_0] = \frac{1}{I_n(\theta)}. \tag{4}$$

Несмещенная оценка $\tilde{\theta} = \tilde{\theta}_n$ называется *асимптотически эффективной* оценкой параметра θ, если

$$\lim_{n\to\infty}\frac{1}{I_n(\theta)\,\mathrm{D}[\tilde{\theta}_n]} = 1. \tag{5}$$

Если условия регулярности (A*) не выполняются, то может существовать несмещенная оценка параметра θ, дисперсия которой меньше, чем нижняя граница в неравенстве (1). Такая оценка называется *сверхэффективной*.

Пример 2. Пусть $x_1, x_2, \ldots, x_n$ — выборка из нормально распределенной генеральной совокупности $N(m,\sigma)$. Показать, что выборочное среднее $\overline{x}$ является эффективной оценкой параметра m.

◁ В примере 1 было показано, что $\overline{x}$ — несмещенная оценка параметра m, причем $\mathrm{D}[\overline{X}] = \dfrac{\sigma^2}{n}$. Используя (2), найдем информацию Фишера $I_n(m)$. Имеем

$$f_X(x,m) = \frac{1}{\sqrt{2\pi}\,\sigma}\exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\};$$

$$\ln f_X(x,m) = -\frac{(x-m)^2}{2\sigma^2} - \ln(\sqrt{2\pi}\,\sigma).$$

Следовательно,

$$\frac{\partial \ln f_X(x, m)}{\partial m} = \frac{x - m}{\sigma^2}.$$

Математическое ожидание случайной величины $\left(\dfrac{X-m}{\sigma^2}\right)^2$ равно:

$$\mathrm{M}\left[\frac{(X-m)^2}{\sigma^4}\right] = \frac{1}{\sigma^4}\,\mathrm{M}\left[(X-m)^2\right] = \frac{1}{\sigma^4}\,\sigma^2 = \frac{1}{\sigma^2}.$$

Таким образом,

$$I_n(m) = n\mathrm{M}\left[\frac{(X-m)^2}{\sigma^4}\right] = \frac{n}{\sigma^2}.$$

Так как условие (4) выполнено, т.е. $\mathrm{D}\left[\overline{X}\right] = \dfrac{\sigma^2}{n} = \dfrac{1}{I_n(m)}$, то для нормально распределенной генеральной совокупности $\overline{x}$ является эффективной оценкой математического ожидания m. ▷

19.113. Пусть $\tilde{\theta} = \tilde{\theta}(x_1, \ldots, x_n)$ является оценкой неизвестного параметра θ по выборке объема n. В качестве меры близости оценки $\tilde{\theta}$ к истинному значению θ выберем величину средней квадратической ошибки $\mathrm{M}\left[(\tilde{\theta}-\theta)^2\right]$. Показать, что

$$\mathrm{M}\left[(\tilde{\theta}-\theta)^2\right] = \mathrm{D}\left[\tilde{\theta}\right] + \left(\mathrm{M}\left[\tilde{\theta}\right] - \theta\right)^2.$$

19.114. Показать, что выборочное среднее является эффективной оценкой параметра λ распределения Пуассона.

19.115. Показать, что относительная частота появления события A в n независимых испытаниях является эффективной оценкой вероятности p появления события A в одном испытании.

19.116. Пусть $x_1, \ldots, x_n$ — выборка из нормально распределенной генеральной совокупности $N(m, \sigma)$. Найти информацию Фишера $I_n(\sigma^2)$.

19.117 (продолжение). В условиях предыдущей задачи при известном математическом ожидании m оценивается дисперсия σ^2. Показать, что статистика

$$s_0^2 = \frac{1}{n}\sum (x_i - m)^2$$

является эффективной оценкой σ^2.

19.118*. Пусть $x_1, \ldots, x_n$ — выборка из генеральной совокупности, имеющей равномерное распределение $R(0, 1)$. Показать, что статистика $\tilde{m} = \dfrac{1}{2}\left(x^{(1)} + x^{(n)}\right)$ является более эффективной оценкой математического ожидания, чем выборочное среднее.

2. Метод максимального правдоподобия. Метод максимального правдоподобия является одним из наиболее распространенных методов нахождения оценок неизвестных параметров распределения генеральной совокупности. Пусть X — непрерывная случайная величина с плотностью распределения $f_X(x, \theta)$, зависящей от неизвестного параметра θ, значение которого требуется оценить по выборке объема n. Плотность распределения выборочного вектора $(X_1, \ldots, X_n)$ можно записать в виде

$$f_{X_i, \ldots, X_n}(x_1, \ldots, x_n, \theta) = \prod_{i=1}^{n} f_{X_i}(x_i, \theta). \tag{6}$$

Пусть, наконец, $x_1, \ldots, x_n$ — выборка наблюдений случайной величины X, по которой находится оценка неизвестного параметра.

Функцией правдоподобия $L(\theta)$ выборки объема n называется плотность выборочного вектора (6), рассматриваемая при фиксированных значениях переменных $x_1, \ldots, x_n$. Функция правдоподобия является, таким образом, функцией только неизвестного параметра θ, т.е.

$$L(\theta) = \prod_{i=1}^{n} f_{X_i}(x_i, \theta). \tag{7}$$

Аналогично определяется функция правдоподобия выборки дискретной случайной величины X. Пусть X — дискретная случайная величина, причем вероятность $\mathrm{P}[X = x] = p(x, \theta)$ есть функция неизвестного параметра θ. Предположим, что для оценки параметра θ получена конкретная выборка наблюдений случайной величины X объема n: $x_1, \ldots, x_n$. Функция правдоподобия $L(\theta)$ выборки объема n равна вероятности того, что компоненты выборочного вектора $X_1, \ldots, X_n$ примут фиксированные значения $x_1, \ldots, x_n$, т.е.

$$L(\theta) = \prod_{i=1}^{n} \mathrm{P}[X_i = x_i] = \prod_{i=1}^{n} p(x_i, \theta). \tag{8}$$

Метод максимального правдоподобия состоит в том, что в качестве оценки неизвестного параметра θ принимается значение $\tilde{\theta}$, доставляющее максимум функции правдоподобия. Такую оценку называют *МП-оценкой.* В случае дискретного распределения наблюдаемой случайной величины X МП-оценка неизвестного параметра θ есть такое значение $\tilde{\theta}$, при котором вероятность появления данной конкретной выборки максимальна. Аналогичную интерпретацию МП-оценки можно дать и в случае оценки параметра распределения непрерывной случайной величины.

Для упрощения вычислений, связанных с получением МП-оценок, в некоторых случаях удобно использовать логарифмическую функцию правдоподобия, т.е. $\ln L(\theta)$.

При выполнении некоторых достаточно общих условий *МП-оценки состоятельны, асимптотически эффективны и асимптотически нормально распределены*. Последнее означает, что при увеличении объема выборки n для МП-оценки $\tilde{\theta}_n$ неизвестного параметра θ выполняется условие

$$\lim_{n\to\infty} \mathrm{P}\left[\frac{\theta - \tilde{\theta}_n}{\sqrt{\mathrm{D}\,[\tilde{\theta}_n]}} < x\right] = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-t^2/2}\,dt = \Phi\,(x).$$

где $\Phi(x)$ — функция распределения нормального закона $N(0,\,1)$.

Если для параметра θ существует эффективная оценка, то метод максимального правдоподобия дает именно эту оценку и другой МП-оценки не существует.

Пример 3. Найти МП-оценки математического ожидания m и дисперсии σ^2 нормально распределенной генеральной совокупности.

◁ Пусть $x_1,\, x_2,\, \ldots,\, x_n$ — выборка наблюдений случайной величины X с плотностью распределения

$$f_X(x,\, m,\, \sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}.$$

Найдем функцию правдоподобия $L\,(m,\,\sigma^2)$. По формуле (7) имеем

$$L\,(m,\,\sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{(x_i-m)^2}{2\sigma^2}\right\} = \\ = \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left\{-\frac{\sum(x_i-m)^2}{2\sigma^2}\right\}.$$

Логарифмическая функция правдоподобия отсюда равна

$$\ln L\,(m,\,\sigma^2) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \frac{\sum(x_i-m)^2}{2\sigma^2}.$$

Используя необходимые условия максимума $\ln L\,(m,\,\sigma^2)$, получим систему уравнений для нахождения МП-оценок:

$$\begin{cases} \dfrac{\partial \ln L\,(m,\,\sigma^2)}{\partial m} = \dfrac{1}{\sigma^2}\sum(x_i-m) = 0, \\ \dfrac{\partial \ln L\,(m,\,\sigma^2)}{\partial \sigma^2} = -\dfrac{n}{2\sigma^2} + \dfrac{1}{2\sigma^4}\sum(x_i-m)^2 = 0. \end{cases} \qquad (9)$$

Из первого уравнения системы (9) находим $\tilde{m} = \frac{1}{n}\sum x_i = \overline{x}$. Подставляя это значение во второе уравнение, получаем $\tilde{\sigma}^2 = \frac{1}{n}\sum(x_i - \overline{x})^2 = D_X^*$.

Отметим, что выборочное среднее $\overline{x}$ является несмещенной и состоятельной оценкой m (см. пример 1), а также эффективной оценкой в случае нормально распределенной генеральной совокупности (см. пример 2). Выборочная дисперсия D_X^* является состоятельной и смещенной оценкой σ^2 (см. задачи 19.98 и 19.100). ▷

Пример 4. Найти МП-оценку параметра λ распределения Пуассона.

◁ Пусть $x_1, \ldots, x_n$ — выборка наблюдений случайной величины X, имеющей распределение Пуассона с неизвестным параметром λ, т.е.

$$\mathrm{P}[X = x] = \frac{\lambda^x}{x!} e^{-\lambda},$$

где x принимает неотрицательные целочисленные значения, $x = 0, 1, 2, \ldots$ Функция правдоподобия $L(\lambda)$ выборки объема n определяется по формуле (8):

$$L(\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} = \frac{\lambda^{\sum x_i}}{x_1!\, x_2! \ldots x_n!} e^{-\lambda n}.$$

Найдем логарифмическую функцию правдоподобия:

$$\ln L(\lambda) = -\ln(x_1! \cdot \cdots \cdot x_n!) + \left(\sum x_i\right) \ln\lambda - \lambda n.$$

Используя необходимое условие экстремума, получим уравнение для определения МП-оценки:

$$\frac{d \ln L(\lambda)}{d\lambda} = \frac{\sum x_i}{\lambda} - n = 0.$$

Отсюда следует, что $\tilde{\lambda} = \frac{1}{n}\sum x_i = \overline{x}$.

Полученная МП-оценка является несмещенной и состоятельной оценкой λ (пример 1), а также эффективной оценкой этого параметра (задача 19.114). ▷

19.119*. Найти МП-оценку параметра σ по выборке объема n из нормально распределенной генеральной совокупности с известным математическим ожиданием m. Показать, что полученная оценка является смещенной.

19.120. Пусть x — наблюдаемое значение случайной величины, имеющей биномиальное распределение $B(n, p)$, или, другими словами, x — число «успехов» в n независимых испытаниях, причем p — вероятность «успеха» в одном испытании. Найти МП-оценку параметра p. Показать, что полученная оценка является несмещенной, состоятельной и эффективной.

19.121. Пусть x — число автолюбителей, заправившихся на данной станции в течение n часов. Предположим, что число автолюбителей, подъезжающих на заправку, есть случайная величина X, имеющая распределение Пуассона с параметром $n\lambda$, где λ — ожидаемое число заправляющихся автолюбителей в течение одного часа. Найти МП-оценку параметра λ. Показать, что полученная оценка является несмещенной, состоятельной и эффективной.

В задачах 19.122 и 19.123 по выборке $x_1, x_2, \ldots, x_n$ объема n найти МП-оценки параметров указанных распределений. Показать, что полученные оценки являются несмещенными и состоятельными.

19.122. Показательное распределение $\mathrm{Ex}(1/\lambda)$.

19.123. Нормальное распределение $N(m, 1)$.

19.124. Независимые случайные величины $X_1, X_2, \ldots, X_k$ имеют биномиальные распределения соответственно $B(n_1, p)$, $B(n_2, p), \ldots, B(n_k, p)$. Пусть $x_1, x_2, \ldots, x_k$ — значения, которые приняли эти случайные величины в некотором эксперименте. Найти МП-оценку параметра p. Показать, что полученная оценка является несмещенной, и вычислить ее дисперсию.

19.125*. Пусть $x_1, \ldots, x_n$ — выборка из генеральной совокупности, имеющей равномерное распределение $R(a, b)$. Найти МП-оценки параметров a и b по выборке.

19.126. Случайная величина X имеет плотность распределения $f_X(x)$:

$$f_X(x) = \begin{cases} kx & \text{при} \quad x \in [0, \sqrt{2/k}\,], \\ 0 & \text{при} \quad x \notin [0, \sqrt{2/k}\,]. \end{cases}$$

Найти МП-оценку математического ожидания X по выборке объема n.

19.127. При помощи n различных приборов получены n измерений случайной величины X. В предположении, что X имеет нормальное распределение, а дисперсия i-го измерения известна и равна σ_i^2, $i = 1, 2, \ldots, n$, найти МП-оценку математического ожидания m случайной величины X. Показать, что полученная оценка является несмещенной, и вычислить ее дисперсию.

19.128. Длина l объекта измерялась независимо друг от друга двумя приборами. Оба прибора дают при измерении случайные ошибки, имеющие нормальное распределение со средним, равным нулю, и дисперсиями σ_1^2 и σ_2^2. Найти МП-оценку l, если первым прибором сделано n_1 измерений, а вторым — n_2 измерений.

19.129. Отказ прибора произошел при k-м испытании. Найти МП-оценку вероятности отказа p при одном испытании и вычислить ее математическое ожидание.

19.130*. Для того чтобы событие A произошло ровно x раз, было проведено n испытаний ($n \geqslant x$). Найти МП-оценку вероятности p появления события A в одном испытании.

3. Метод моментов. Для получения оценок неизвестных параметров $\theta_1, \theta_2, \ldots, \theta_s$ распределения генеральной совокупности X часто используется метод моментов, состоящий в следующем.

Пусть $f_X(x, \theta_1, \theta_2, \ldots, \theta_s)$ — плотность распределения случайной величины X. Определим с помощью этой плотности s каких-либо моментов случайной величины X, например первые s начальных моментов, по формулам

$$\alpha_m(\theta_1, \ldots, \theta_s) = \mathrm{M}[X^m] = \int\limits_{-\infty}^{+\infty} x^m f_X(x, \theta_1, \ldots, \theta_s)\,dx, \quad m = 1, 2, \ldots, s.$$

По выборке наблюдений случайной величины найдем значения соответствующих выборочных моментов:

$$\alpha_m^* = \frac{1}{n}\sum_{i=1}^{n} x_i^m, \quad m = 1, 2, \ldots, s.$$

Попарно приравнивая теоретические моменты α_m случайной величины X их выборочным значениям α_m^*, получаем систему s уравнений с неизвестными $\theta_1, \ldots, \theta_s$:

$$\alpha_m(\theta_1, \ldots, \theta_s) = \alpha_m^*, \quad m = 1, 2, \ldots, s.$$

Решая полученную систему относительно неизвестных $\theta_1, \ldots, \theta_s$, находим оценки $\tilde{\theta}_1, \ldots, \tilde{\theta}_s$ неизвестных параметров.

Аналогично находятся оценки неизвестных параметров по выборке наблюдений дискретной случайной величины.

Пример 5. Методом моментов найти оценки неизвестных параметров a и b для Г-распределения с плотностью

$$f_X(x) = \begin{cases} 0, & x \leqslant 0, \\ \dfrac{b^a}{\Gamma(a)}\, x^{a-1} e^{-bx}, & x > 0. \end{cases}$$

◁ Для нахождения оценок параметров a и b по методу моментов воспользуемся начальным моментом первого порядка (математическим ожиданием) и центральным моментом второго порядка (дисперсией):

$$\alpha_1(a, b) = m = \frac{a}{b}, \tag{10}$$

$$\mu_2(a, b) = \sigma^2 = \frac{a}{b^2}. \tag{11}$$

По выборке $x_1, \ldots, x_n$ из генеральной совокупности, имеющей Г-распределение, находим значения соответствующих выборочных моментов:

$$\alpha_1^* = \overline{x} = \frac{1}{n}\sum x_i, \tag{12}$$

$$\mu_2^* = D_{\text{x}}^* = \frac{1}{n}\sum (x_i - \overline{x})^2. \tag{13}$$

Приравнивая (10) и (12), (11) и (13) соответственно, получаем следующую систему уравнений:

$$\frac{a}{b} = \overline{x}, \quad \frac{a}{b^2} = D_{\text{x}}^*,$$

решая которую, находим $\tilde{a} = \dfrac{\overline{x}^2}{D_{\text{x}}^*}$, $\tilde{b} = \dfrac{\overline{x}}{D_{\text{x}}^*}$. ▷

19.131. В n независимых испытаниях событие A произошло x раз. Методом моментов найти оценку вероятности p появления события A в одном испытании.

В задачах 19.132–19.135 по выборке $x_1, x_2, \ldots, x_n$ объема n найти оценки параметров указанных распределений, используя метод моментов.

19.132. Пуассоновское распределение с параметром λ.

19.133. Нормальное распределение $N(m, \sigma)$.

19.134. Показательное распределение $\mathrm{Ex}(\lambda)$.

19.135. Распределение $\chi^2(k)$.

19.136. Используя таблицу случайных чисел (таблица П4), либо метод моделирования, получить 50 равномерно распределенных чисел из интервала $[0, 10]$. Методом моментов найти оценки параметров равномерного распределения, используя эти данные.

19.137*. Рассмотрим n систем с временами работы до первого отказа соответственно $X_1, \ldots, X_n$. Предположим, что $X_1, \ldots$ $\ldots, X_n$ — независимые в совокупности и одинаково распределенные случайные величины с показательным распределением $\mathrm{Ex}(\lambda)$.

Пусть, наконец, x_i, $i = 1, 2, \ldots, n$, — измеренные значения времени отказа i-й системы (в часах). Используя метод моментов, найти оценку вероятности $\mathrm{P}[X_1 \geqslant 1]$ того, что первая система будет работать бесперебойно в течение часа.

4. Распределения χ^2, Стьюдента и Фишера. Распределения основных статистик, вычисляемых по выборке из нормально распределенной генеральной совокупности, связаны с распределениями $\chi^2(k)$, Стьюдента $T(k)$ и Фишера $F(k_1, k_2)$. Квантили этих распределений приведены в Приложении (таблицы П5, П6, П7). Приведем определения и некоторые свойства этих распределений.

Распределением χ^2 *с* k *степенями свободы* называется распределение случайной величины $\chi^2(k)$, равной сумме квадратов k независимых нормально распределенных по закону $N(0, 1)$ случайных величин U_i, $i = 1, 2, \ldots, k$, т.е. распределение случайной величины

$$\chi^2(k) = U_1^2 + \cdots + U_k^2.$$

Распределение χ^2 с k степенями свободы там, где это не вызывает недоразумений, будет обозначаться также $\chi^2(k)$.

Плотность распределения $f_{\chi^2}(x)$ определяется формулой

$$f_{\chi^2}(x) = \begin{cases} 0, & x \leqslant 0, \\ \dfrac{1}{2^{k/2}\Gamma\left(\dfrac{k}{2}\right)} x^{(k-2)/2} e^{-x/2}, & x > 0. \end{cases}$$

График функции $f_{\chi^2}(x)$ приведен на рис. 31. Среднее и дисперсия распределения $\chi^2(k)$ равны соответственно: $\mathrm{M}[\chi^2(k)] = k$, $\mathrm{D}[\chi^2(k)] = 2k$.

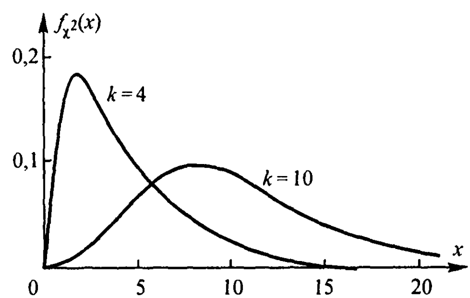


Рис. 31

Распределение χ^2 часто используется в статистических вычислениях, в частности, в связи со следующей теоремой.

Теорема 2. *Пусть $x_1, x_2, \ldots, x_n$ — выборка из нормально распределенной генеральной совокупности $N(m, \sigma)$, а $\overline{x} = \frac{1}{n}\sum x_i$ и $s^2 = \frac{1}{n-1}\sum (x_i - \overline{x})^2$ — соответственно выборочное среднее и выборочная дисперсия. Тогда статистики $\overline{X}$ и S^2 — независимые случайные величины, причем статистика $\frac{n-1}{\sigma^2}S^2$ имеет распределение $\chi^2(n-1)$.*

Заметим, что если $\chi^2(k_1)$ и $\chi^2(k_2)$ — независимые случайные величины, имеющие распределение χ^2 с k_1 и k_2 степенями свободы соответственно, то сумма этих случайных величин имеет распределение χ^2 с $k_1 + k_2$ степенями свободы:

$$\chi^2(k_1) + \chi^2(k_2) = \chi^2(k_1 + k_2).$$

Распределение $\chi^2(k)$ при больших значениях k ($k > 30$) с достаточной для практических расчетов точностью аппроксимируется нормальным распределением.

Это свойство используется для приближенного выражения квантилей $\chi^2_p(k)$ распределения $\chi^2(k)$ через квантили u_p нормального распределения $N(0, 1)$. Обычно используют следующие две формулы:

$$\chi^2_p(k) \approx \frac{1}{2}\left(u_p + \sqrt{2k-1}\right)^2, \tag{14}$$

$$\chi^2_p(k) \approx k\left(1 - \frac{2}{9k} + u_p\sqrt{\frac{2}{9k}}\right)^3. \tag{15}$$

Формула (14), применяемая при $k \geqslant 30$ и $p \geqslant 0{,}5$, дает относительную погрешность в пределах 1 %, а формула (15) применяется для вычисления квантилей малого порядка.

Пример 6. Вычислить квантили $\chi^2_{0,01}(10)$, $\chi^2_{0,95}(40)$, $\chi^2_{0,01}(40)$.

◁ Из таблицы П5 находим $\chi^2_{0,01}(10) = 2{,}56$. Для вычисления квантили $\chi^2_{0,95}(40)$ воспользуемся формулой (14). Так как $u_{0,95} = 1{,}645$ (см. таблицу П1), то $\chi^2_{0,95}(40) \approx \frac{1}{2}\left(1{,}645 + \sqrt{2\cdot 40 - 1}\right)^2 \approx 55{,}47$.

По формуле (15), используя значение $u_{0,01} = -u_{0,99} = -2{,}326$, получаем $\chi^2_{0,01}(40) \approx 40\left(1 - \frac{2}{9\cdot 40} - 2{,}326\sqrt{\frac{2}{9\cdot 40}}\right)^3 \approx 22{,}14$. ▷

Распределением Стьюдента с k степенями свободы называется распределение случайной величины $T(k)$, равной отношению двух независимых случайных величин U и $\sqrt{\chi^2(k)/k}$, т.е.

$$T(k) = \frac{U}{\sqrt{\chi^2(k)/k}},$$

где U имеет нормальное распределение $N(0, 1)$. Распределение Стьюдента с k степенями свободы будет также обозначаться $T(k)$. Распреде-

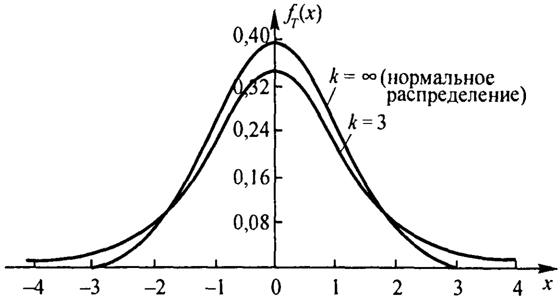


Рис. 32

ление Стьюдента с k степенями свободы имеет плотность $f_T(x)$ (рис. 32):

$$f_T(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{\pi k}}\left(1 + \frac{x^2}{k}\right)^{-(k+1)/2}, \quad -\infty < x < +\infty,$$

среднее $\mathrm{M}[T(k)] = 0$ и дисперсию $\mathrm{D}[T(k)] = \dfrac{k}{k-2}$, $k > 2$. Плотность распределения Стьюдента симметрична относительно оси ординат, следовательно, для квантилей $t_p(k)$ имеет место соотношение $t_p(k) = -t_{1-p}(k)$.

При больших k ($k > 30$) для квантилей $t_p(k)$ распределения Стьюдента выполнено приближенное равенство $t_p(k) \approx u_p$. Более точная формула имеет вид

$$t_p(k) \approx u_p\left(\left(1 - \frac{1}{4k}\right)^2 - \frac{u_p^2}{2k}\right)^{-1/2}. \tag{16}$$

Пример 7. Найти квантили $t_{0,05}(8)$ и $t_{0,90}(40)$.

По таблице П6 находим $t_{0,95}(8) = 1{,}86$; $t_{0,05}(8) = -t_{0,95}(8) = -1{,}86$. Квантиль $t_{0,90}(40)$ определим, используя формулу (16). Так как $u_{0,90} = 1{,}28$ по таблице П1, то

$$t_{0,90}(40) \approx 1{,}28\left(\left(1 - \frac{1}{4\cdot 40}\right)^2 - \frac{1{,}28^2}{2\cdot 40}\right)^{-1/2} \approx 1{,}307.$$

Точное значение квантили $t_{0,90}(40)$ по таблице П6 равно 1,303. ▷

Распределением Фишера с k_1 и k_2 степенями свободы называется распределение случайной величины $F(k_1, k_2)$, равной отношению двух независимых случайных величин $\chi^2(k_1)/k_1$ и $\chi^2(k_2)/k_2$, т.е.

$$F(k_1, k_2) = \frac{\chi^2(k_1)/k_1}{\chi^2(k_2)/k_2}.$$

Распределение Фишера с k_1 и k_2 степенями свободы будет также обозна-

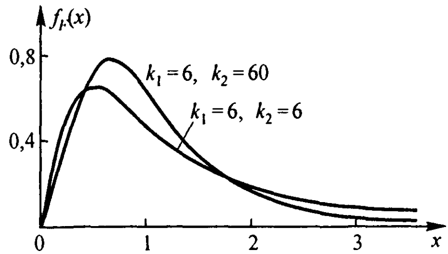


Рис. 33

чаться $F(k_1, k_2)$. Распределение Фишера с k_1 и k_2 степенями свободы имеет плотность $f_F(x)$ (рис. 33):

$$f_F(x) = \begin{cases} 0, & x \leqslant 0, \\ \dfrac{\Gamma\left(\dfrac{k_1+k_2}{2}\right)}{\Gamma\left(\dfrac{k_1}{2}\right)\Gamma\left(\dfrac{k_2}{2}\right)} \left(\dfrac{k_1}{k_2}\right)^{k_1/2} \dfrac{x^{(k_1/2)-1}}{\left(1+\dfrac{k_1}{k_2}\right)^{(k_1+k_2)/2}}, & x > 0, \end{cases}$$

среднее $\mathrm{M}[F] = \dfrac{k_2}{k_2-2}$, $k_2 > 2$.

Квантили распределения Фишера порядка p и $1-p$ связаны следующей формулой:

$$F_{1-p}(k_1, k_2) = \frac{1}{F_p(k_2, k_1)}. \tag{17}$$

Между случайными величинами, имеющими нормальное распределение, распределения χ^2, Стьюдента и Фишера, имеют место соотношения

$$T^2(k) = F(1, k), \tag{18}$$

$$F(k, \infty) = \frac{\chi^2(k)}{k}, \tag{19}$$

$$\chi^2(1) = U^2. \tag{20}$$

При $k_1 \gg 1$ и $k_2 \gg 1$ квантили распределения Фишера можно вычислить, используя приближенную формулу

$$F_p(k_1,\, k_2) \approx \frac{k_2}{k_2-2}\sqrt{\frac{2\,(k_1+k_2-2)}{k_1(k_2-4)}}\, u_p + \frac{k_2}{k_2-2}. \tag{21}$$

Пример 8. Вычислить квантили $F_{0,01}(3{,}5)$, $F_{0,90}(4{,}100)$ и $F_{0,05}(60,\ 120)$.

◁ Используя соотношение (17) и таблицу П7, получаем

$$F_{0,01}(3{,}5) = \frac{1}{F_{0,99}(5{,}3)} = \frac{1}{28{,}24} \approx 0{,}035.$$

Далее, используя соотношение (19) и таблицу П5, находим

$$F_{0,90}(4{,}100) \approx \frac{\chi^2_{0,9}(4)}{4} = \frac{7{,}78}{4} = 1{,}945.$$

Наконец, по формуле (21), используя значение $u_{0,05} = -u_{0,95} = -1{,}645$, получаем

$$F_{0,05}(60{,}120) \approx \frac{120}{120-2}\sqrt{\frac{2\,(60+120-2)}{60\,(120-4)}}\,(-1{,}645) + \frac{120}{120-2} \approx 0{,}639.$$

По таблице П7 значение квантили $F_{0,05}(60,\ 120)$ равно

$$F_{0,05}(60,\ 120) = \frac{1}{F_{0,95}(120,\ 60)} = \frac{1}{1{,}43} \approx 0{,}699. \ \triangleright$$

Используя таблицы квантилей и свойства распределений, определить квантили:

19.138. $\chi^2_{0,05}(8)$ и $\chi^2_{0,99}(130)$.

19.139. $t_{0,01}(7)$ и $t_{0,95}(110)$.

19.140. $F_{0,05}(2,\ 3)$, $F_{0,99}(100,\ 5)$ и $F_{0,01}(60,\ 90)$.

19.141*. Используя свойства распределений χ^2, Стьюдента и Фишера, доказать соотношение (18).

В задачах 19.142–19.145 изучаются свойства статистик, вычисляемых по выборке $x_1,\, x_2,\, \ldots,\, x_n$ из генеральной совокупности, имеющей нормальное распределение $N(m,\, \sigma)$.

19.142. Показать, что выборочное среднее $\overline{X} = \dfrac{1}{n}\sum X_i$ имеет нормальное распределение $N(m,\, \sigma/\sqrt{n}\,)$.

19.143. Найти распределение статистики $S_0^2 = \frac{1}{n}\sum(X_i - m)^2$.

19.144*. Найти математическое ожидание и дисперсию статистики

$$S^2 = \frac{1}{n-1}\sum(X_i - \overline{X})^2.$$

Показать, что статистика S^2 является асимптотически эффективной оценкой дисперсии σ^2.

19.145. Показать, что статистика $T(n-1) = \dfrac{\overline{X} - m}{S/\sqrt{n}}$ имеет распределение Стьюдента с $n-1$ степенью свободы. Найти математическое ожидание и дисперсию $T(n-1)$.

В задачах 19.146–19.149 рассматриваются две независимые выборки объемов n_1 и n_2 из генеральных совокупностей с распределениями $N(m_1, \sigma_1)$ и $N(m_2, \sigma_2)$. $\overline{X}_1$ и $\overline{X}_2$ — выборочные средние, а S_1^2 и S_2^2 — выборочные дисперсии, вычисляемые по этим выборкам.

19.146. Показать, что статистика $F(n_1 - 1,\ n_2 - 1) = \dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$ имеет распределение Фишера с $n_1 - 1$ и $n_2 - 1$ степенями свободы.

19.147. Показать, что если математические ожидания m_1 и m_2 генеральных совокупностей известны, а оценками дисперсий σ_1^2 и σ_2^2 являются статистики S_{01}^2 и S_{02}^2, где $S_{0i}^2 = \frac{1}{n_i}\sum(X_i - m_i)^2$, $i = 1, 2$, то статистика $F(n_1, n_2) = \dfrac{S_{01}^2/\sigma_1^2}{S_{02}^2/\sigma_2^2}$ имеет распределение Фишера с n_1 и n_2 степенями свободы.

19.148. Показать, что если дисперсии генеральных совокупностей известны и равны $\sigma_1^2 = \sigma_2^2 = \sigma^2$, то статистика

$$\frac{\overline{X}_1 - \overline{X}_2 - (m_1 - m_2)}{\sigma\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

имеет распределение $N(0, 1)$.

19.149. Предположим, что дисперсии обеих генеральных совокупностей равны $\sigma_1^2 = \sigma_2^2 = \sigma^2$, но значение σ^2 неизвестно и оценивается при помощи статистики $S^2 = \dfrac{(n_1 - 1)\,S_1^2 + (n_2 - 1)\,S_2^2}{n_1 + n_2 - 2}$. Найти распределение статистики S^2 и вычислить ее дисперсию.

Показать, что оценка неизвестной дисперсии с помощью статистики S^2 более эффективна, чем оценка, вычисляемая по одной из выборок.

19.150. Найти распределение статистики $\overline{X} = \dfrac{n_1\overline{X}_1 + n_2\overline{X}_2}{n_1 + n_2}$.

19.151. В условиях задачи 19.149 показать, что статистика

$$\frac{\overline{X}_1 - \overline{X}_2 - (m_1 - m_2)}{S\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

имеет распределение Стьюдента с $n_1 + n_2 - 2$ степенями свободы.

19.152. Методом моделирования получить 25 выборок объема 15 из генеральной совокупности с нормальным распределением $\mathcal{N}(3, 3)$. Для каждой выборки найти выборочное среднее $\overline{x}$. Полученные данные представляют 25 выборочных значений статистики $\overline{X}$. Выполнить следующие задания:

1) Найти распределение рассматриваемой статистики.

2) Представить выборочные значения в виде гистограммы частот.

3) Найти оценки математического ожидания и дисперсии данной статистики и сравнить их с теоретическими значениями.

19.153 (продолжение). Для каждой из выборок предыдущей задачи найти выборочную дисперсию по формуле $S_0^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n}(x_i - m)^2$, где $n = 15$, а $m = 3$. Используя выборочные значения статистики S_0^2, выполнить задания к задаче 19.152.

19.154 (продолжение). Решить задачу 19.153 для выборочной дисперсии $s^2 = \dfrac{1}{n-1}\sum(x_i - \overline{x})^2$.

19.155 (продолжение). Решить задачу 19.153 для статистики $\dfrac{\overline{X} - m}{S/\sqrt{n}}$.

§ 3. Интервальные оценки

1. Доверительные интервалы и доверительная вероятность. Доверительные интервалы для параметров нормально распределенной генеральной совокупности. При статистической обработке результатов наблюдений часто необходимо не только найти оценку $\tilde{\theta}$ неизвестного параметра θ, но и охарактеризовать точность этой оценки. С этой целью вводится понятие доверительного интервала.

Доверительным интервалом для параметра θ называется интервал (θ_1, θ_2), содержащий (накрывающий) истинное значение θ с заданной вероятностью $p = 1 - \alpha$, т.е.

$$\mathrm{P}\left[\theta_1 < \theta < \theta_2\right] = 1 - \alpha. \tag{1}$$

Число $1 - \alpha$ называется *доверительной вероятностью*, а значение α — *уровнем значимости*. Статистики $\theta_1 = \theta_1(x_1, \ldots, x_n)$ и $\theta_2 = \theta_2(x_1, \ldots, x_n)$, определяемые по выборке $x_1, \ldots, x_n$ из генеральной совокупности с неизвестным параметром θ, называются соответственно *нижней и верхней границами доверительного интервала*.

Условие (1) означает, что в большой серии независимых экспериментов, в каждом из которых получена выборка объема n, в среднем $(1 - \alpha)\,100\,\%$ из общего числа построенных доверительных интервалов содержат истинное значение параметра θ.

Длина доверительного интервала, характеризующая точность интервального оценивания, зависит от объема выборки n и доверительной вероятности $1 - \alpha$: при увеличении объема выборки длина доверительного интервала уменьшается, а с приближением доверительной вероятности к единице — увеличивается. Выбор доверительной вероятности определяется конкретными условиями. Обычно используются значения $1 - \alpha$, равные 0,90; 0,95; 0,99.

При решении некоторых задач применяются односторонние доверительные интервалы, границы которых определяются из условий

$$\mathrm{P}\left[\theta < \theta_2\right] = 1 - \alpha \quad \text{или} \quad \mathrm{P}\left[\theta_1 < \theta\right] = 1 - \alpha.$$

Эти интервалы называются соответственно *левосторонними и правосторонними доверительными интервалами*.

Чтобы найти доверительный интервал для параметра θ, необходимо знать закон распределения статистики $\tilde{\theta} = \tilde{\theta}\,(x_1, \ldots, x_n)$, значение которой является оценкой параметра θ. При этом для получения доверительного интервала наименьшей длины при данном объеме выборки n и заданной доверительной вероятности $1 - \alpha$ в качестве оценки $\tilde{\theta}$ параметра θ следует брать эффективную либо асимптотически эффективную оценку.

Один из методов построения доверительных интервалов состоит в следующем. Предположим, что существует статистика $Y = Y\,(\tilde{\theta}, \theta)$ такая, что:

а) закон распределения Y известен и не зависит от θ;

б) функция $Y\,(\tilde{\theta}, \theta)$ непрерывна и строго монотонна по θ.

Пусть, далее, $(1 - \alpha)$ — заданная доверительная вероятность, а $y_{\alpha/2}$ и $y_{1-\alpha/2}$ — квантили распределения статистики Y порядков $\alpha/2$ и $1 - \alpha/2$ соответственно. Тогда с вероятностью $1 - \alpha$ выполняется неравенство

$$y_{\alpha/2} < Y\,(\tilde{\theta}, \theta) < y_{1-\alpha/2}. \tag{2}$$

Решая неравенство (2) относительно θ, найдем границы θ_1 и θ_2 доверительного интервала для θ. Если плотность распределения статистики Y

симметрична относительно оси Oy, то доверительный интервал имеет наименьшую длину, а если это распределение несимметрично, то длину, близкую к наименьшей.

Пример 1. Пусть $x_1, x_2, \ldots, x_n$ — выборка из нормально распределенной генеральной совокупности. Найти доверительный интервал для математического ожидания m при условии, что дисперсия генеральной совокупности известна и равна σ^2, а доверительная вероятность равна $1 - \alpha$.

◁ В качестве оценки математического ожидания m возьмем выборочное среднее $\overline{x} = \dfrac{1}{n}\sum x_i$. Для нормально распределенной генеральной совокупности выборочное среднее является эффективной оценкой m (см. пример 3 из § 2). Выборочное среднее $\overline{X}$ в данном случае имеет нормальное распределение $N(m, \sigma/\sqrt{n})$.

Рассмотрим статистику $U = \dfrac{\overline{X} - m}{\sigma/\sqrt{n}}$, имеющую нормальное распределение $N(0, 1)$ независимо от значения параметра m. Кроме того, U как функция m непрерывна и строго монотонна. Следовательно,

$$\mathrm{P}\left[u_{\alpha/2} < U < u_{1-\alpha/2}\right] = 1 - \alpha,$$

где $u_{\alpha/2}$ и $u_{1-\alpha/2}$ — квантили нормального распределения $N(0, 1)$.

Решая неравенство

$$u_{\alpha/2} < \frac{\overline{X} - m}{\sigma/\sqrt{n}} < u_{1-\alpha/2}$$

относительно m, получим, что с вероятностью $1 - \alpha$ выполняется следующее условие:

$$\overline{X} - \frac{\sigma}{\sqrt{n}}\, u_{1-\alpha/2} < m < \overline{X} - \frac{\sigma}{\sqrt{n}}\, u_{\alpha/2}.$$

Так как квантили нормального распределения связаны соотношением $u_{\alpha/2} = -u_{1-\alpha/2}$, полученный доверительный интервал для m можно записать следующим образом:

$$\overline{X} - \frac{\sigma}{\sqrt{n}}\, u_{1-\alpha/2} < m < \overline{X} + \frac{\sigma}{\sqrt{n}}\, u_{1-\alpha/2}. \ \triangleright$$

В задачах 19.156–19.175 предполагается, что выборка объема n получена из генеральной совокупности, имеющей либо нормальное распределение $N(m, \sigma)$, либо распределение, достаточно близкое к нормальному.

19.156. Показать, что если дисперсия генеральной совокупности σ^2 неизвестна, а в качестве оценки дисперсии используется статистика $s^2 = \dfrac{1}{n-1}\sum(x_i - \overline{x})^2$, то при доверительной вероятности $1-\alpha$ доверительный интервал для математического ожидания m имеет вид

$$\overline{x} - \frac{s}{\sqrt{n}}t_{1-\alpha/2}(n-1) < m < \overline{x} + \frac{s}{\sqrt{n}}t_{1-\alpha/2}(n-1),$$

где $t_{1-\alpha/2}(n-1)$ — квантиль распределения Стьюдента с $n-1$ степенью свободы.

Выборочные оценки в задачах 19.157–19.160 определялись по результатам n наблюдений. Используя эти данные, а также результаты примера 1 и задачи 19.156, найти 90 % и 99 %-ные доверительные интервалы для математического ожидания (среднего) следующих характеристик:

19.157. Емкость конденсатора, если $\overline{x} = 20$ мкФ, $n = 16$, среднеквадратичное отклонение известно и равно 4 мкФ.

19.158. Время безотказной работы электронной лампы, если $\overline{x} = 500$, $n = 100$, среднеквадратичное отклонение известно и равно 10 ч.

19.159. Диаметр вала, если $\overline{x} = 30$ мм, $n = 9$, $s^2 = 9$ мм^2.

19.160. Содержание углерода в единице продукта, если $\overline{x} = 18$ г, $n = 25$, $s^2 = 16$ г^2.

19.161. Методом моделирования получить 10 выборок объема 25 из генеральной совокупности, имеющей нормальное распределение $N(5,\ 1)$. Для каждой выборки найти доверительный интервал для математического ожидания m, считая, что дисперсия генеральной совокупности известна и равна 1. Доверительную вероятность принять равной 0,9.

Какая часть из полученных интервалов накроет параметр $m = 5$?

19.162 (продолжение). Решить предыдущую задачу, считая, что дисперсия σ^2 генеральной совокупности неизвестна.

19.163 (продолжение). Решить задачи 19.161 и 19.162 при доверительной вероятности 0,99.

19.164*. Пусть из одной генеральной совокупности получены две выборки объемов n_1 и n_2 соответственно. Выборочные оценки средних и дисперсий по этим выборкам равны $\overline{X}_1$, $\overline{X}_2$, S_1^2, S_2^2. Объединенные оценки среднего и дисперсии по выборке объема $n_1 + n_2$ вычисляются по формулам

$$\overline{X} = \frac{n_1\overline{X}_1 + n_2\overline{X}_2}{n_1 + n_2}, \quad S^2 = \frac{(n_1-1)\,S_1^2 + (n_2-1)\,S_2^2}{n_1 + n_2 - 2}.$$

Показать, что если дисперсия генеральной совокупности известна и равна σ^2, то доверительный интервал для среднего определяется следующим образом:

$$\overline{X} - \frac{\sigma}{\sqrt{n_1+n_2}}u_{1-\alpha/2} < m < \overline{X} + \frac{\sigma}{\sqrt{n_1+n_2}}u_{1-\alpha/2};$$

если дисперсия генеральной совокупности неизвестна, $\tilde{\sigma}^2 = S^2$, то доверительный интервал для среднего определяется так:

$$\overline{X} - \frac{S}{\sqrt{n_1+n_2}}t_{1-\alpha/2}(n_1+n_2-2) < m < \\ < \overline{X} + \frac{S}{\sqrt{n_1+n_2}}t_{1-\alpha/2}(n_1+n_2-2).$$

Для уточнения характеристик, приведенных взадачах 19.157–19.160 проделаны повторные эксперименты и получены новые выборочные оценки. Используя объединенные выборочные оценки (см. задачу 19.164), найти 90 % и 99 %-ные доверительные интервалы для среднего (задачи 19.165–19.168).

19.165. Емкость конденсатора, если $n = 9$, $\overline{x} = 18$ мкФ.

19.166. Время безотказной работы электронной лампы, если $n = 64$, $\overline{x} = 480$ ч.

19.167. Диаметр вала, если $n = 16$, $\overline{x} = 29$ мм, $s^2 = 4{,}5$ мм^2.

19.168. Содержание углерода в единице продукта, если $n = 9$, $\overline{x} = 18{,}8$ г, $s^2 = 20$ г^2.

19.169*. Показать, что если m известно, а оценка дисперсии равна $\tilde{\sigma}^2 = s_0^2 = \frac{1}{n}\sum(x_i - m)^2$, то доверительный интервал для дисперсии при доверительной вероятности $1-\alpha$ имеет вид

$$\frac{ns_0^2}{\chi^2_{1-\alpha/2}(n)} < \sigma^2 < \frac{ns_0^2}{\chi^2_{\alpha/2}(n)},$$

где $\chi^2_p(n)$ — квантиль распределения χ^2 с n степенями свободы.

19.170. Показать, что если m неизвестно, $\tilde{m} = \overline{x}$, а $\tilde{\sigma}^2 = s^2 = \frac{1}{n-1}\sum(x_i - \overline{x})^2$, то доверительный интервал для дисперсии при доверительной вероятности $1-\alpha$ имеет вид

$$\frac{(n-1)\,s^2}{\chi^2_{1-\alpha/2}(n-1)} < \sigma^2 < \frac{(n-1)\,s^2}{\chi^2_{\alpha/2}(n-1)}.$$

При решении задач 19.171–19.173 используются доверительные интервалы для дисперсии, полученные в задачах 19.169 и 19.170.

19.171. По данным задачи 19.167 найти 90 % и 95 %-ный доверительные интервалы для дисперсии.

19.172. По данным задачи 19.168 найти 90 % и 99 %-ный доверительные интервалы для дисперсии.

19.173. Результаты 10 измерений емкости конденсатора прибором, не имеющим систематической ошибки, дали такие отклонения от номинала (пкФ): 5,4; −13,9; −11; 7,2; −15,6; 29,2; 1,4; −0,3; 6,6; −9,9. Найти 90 %-ный доверительный интервал для дисперсии и среднего квадратического отклонения.

19.174. Оценка величины сопротивления для большой партии однотипных резисторов, определенная по результатам измерений 100 случайно отобранных экземпляров, равна $\overline{x} = 10$ кОм.

а) Считая, что дисперсия измерений известна: $\sigma^2 = 1$ кОм^2, найти вероятность того, что для резисторов всей партии величина сопротивления лежит в пределах $10 \pm 0{,}1$ кОм.

в) Сколько измерений нужно произвести, чтобы с вероятностью 0,95 утверждать, что для всей партии резисторов величина сопротивления лежит в пределах $10 \pm 0{,}1$ кОм?

19.175. Для определения вертикального угла ориентира используют среднее арифметическое нескольких замеров угла при помощи секстанта. Для углов, измеряемых секстантом, с. к. о. принимается равным $\sigma = 1{,}5'$. Найти количество замеров, которое нужно произвести, чтобы:

а) погрешность результата с вероятностью 0,99 не превосходила $1'$;

б) погрешность результата с вероятностью 0,95 не превосходила $1{,}5'$.

В задачах 19.176–19.182 рассматриваются две независимые выборки объемов n_1 и n_2 из генеральных совокупностей с распределениями $N(m_1, \sigma_1)$ и $N(m_2, \sigma_2)$. $\overline{x}_1$ и $\overline{x}_2$ — выборочные средние, а s_1^2 и s_2^2 — выборочные дисперсии, вычисляемые по этим выборкам.

19.176*. Показать, что если дисперсии обеих совокупностей известны, то доверительный интервал для разности средних $m_1 - m_2$ определяется формулой:

$$(\overline{x}_1 - \overline{x}_2) - u_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < m_1 - m_2 < (\overline{x}_1 - \overline{x}_2) + u_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

19.177*. Пусть $\sigma_1^2 = \sigma_2^2 = \sigma^2$, величина σ^2 неизвестна, а в качестве оценки σ^2 используется статистика

$$s^2 = \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2}.$$

Показать, что доверительный интервал для разности средних $m_1 - m_2$ определяется формулой

$$(\overline{x}_1 - \overline{x}_2) - t_{1-\alpha/2}(n_1 + n_2 - 2)s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} <$$
$$< m_1 - m_2 < (\overline{x}_1 - \overline{x}_2) + t_{1-\alpha/2}(n_1 + n_2 - 2)s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}.$$

19.178. Рассматривается случайная величина $Z = X - Y$, где X и Y — независимые случайные величины. Выборочные оценки для X и Y определялись по результатам $n_1 = 16$ и $n_2 = 36$ наблюдений соответственно. Найти 95%-ный доверительный интервал для математического ожидания Z, если $\overline{x} = 10$, $\overline{y} = 4$, σ_X^2 и σ_Y^2 известны и таковы: $\sigma_X^2 = 1$, $\sigma_Y^2 = 4$.

19.179. Амплитуда колебаний определялась двумя лаборантами. Первый лаборант по 10 наблюдениям получил среднее значение амплитуды $\overline{x}_1 = 81$ мм, а второй по 15 наблюдениям — среднее значение $\overline{x}_2 = 84$ мм. В предположении, что дисперсии измерений известны и равны $\sigma_1^2 = 64$ мм^2 и $\sigma_2^2 = 81$ мм^2 для первого и второго лаборанта соответственно, найти 99%-ный доверительный интервал для разности средних $\overline{X}_1$ и $\overline{X}_2$. Можно ли считать, что результаты лаборантов действительно различаются?

19.180. Из большой партии диодов были взяты две выборки с интервалом в один месяц. Результаты измерения времени восстановления у диодов первой выборки (нс): 51, 62, 53, 52, 63. Выборочное среднее и дисперсия времени восстановления для 7 диодов второй выборки: $\overline{x}_2 = 60{,}3$, $s_2^2 = 36{,}06$.

а) Найти 95%-ный доверительный интервал для среднего по данным первой выборки.

б) Найти 99%-ный доверительный интервал для изменения среднего в течение месяца в предположении, что дисперсия времени восстановления диодов за этот период времени не изменилась. Можно ли считать, что среднее времени восстановления не изменилось?

19.181*. Показать, что доверительный интервал для отношения дисперсий определяется формулой

$$\frac{s_1^2}{s_2^2} F_{\alpha/2}(n_2 - 1,\, n_1 - 1) < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} F_{1-\alpha/2}(n_2 - 1,\, n_1 - 1),$$

где $F_p(n_2 - 1,\, n_1 - 1)$ — квантиль распределения Фишера с $n_2 - 1$ и $n_1 - 1$ степенями свободы порядка p.

19.182. Найти 90%-ный доверительный интервал для отношения дисперсий по данным задачи 19.180.

2. Доверительные интервалы для вероятности успеха в схеме Бернулли и параметра λ распределения Пуассона. Если распределение генеральной совокупности не является нормальным, то в отдельных случаях по выборкам большого объема можно построить доверительные интервалы для неизвестных параметров приближенно, используя при этом предельные теоремы теории вероятностей (см. главу 18, § 5) и вытекающие из них асимптотические распределения и оценки.

Пример 2. Пусть в n независимых испытаниях успех наступил x раз. Найти доверительный интервал для вероятности p успеха в одном испытании.

$\triangleleft$ Эффективной оценкой вероятности успеха p в одном испытании является относительная частота $\tilde{p} = h = x/h$ (см. задачу 19.115). По теореме Муавра–Лапласа (гл. 18, § 5, п. 2) относительная частота h имеет асимптотически нормальное распределение $N(p,\ \sqrt{pq/n})$, где $q = 1-p$.

Рассмотрим статистику $U = (h-p)/\sqrt{pq/n}$, которая, следовательно, имеет асимптотически нормальное распределение $N(0,\ 1)$ независимо от значения p. При больших n тогда имеем

$$\mathrm{P}\left[\left|\frac{h-p}{\sqrt{pq/n}}\right| < u_{1-\alpha/2}\right] \approx 1-\alpha.$$

Отсюда получаем, что с вероятностью $\approx 1-\alpha$ выполняется неравенство

$$h - u_{1-\alpha/2}\sqrt{\frac{pq}{n}} < p < h + u_{1-\alpha/2}\sqrt{\frac{pq}{n}}. \tag{3}$$

Заменяя значения p и q в левой и правой частях неравенства (3) их оценками $\tilde{p} = h$ и $\tilde{q} = 1-h$, получаем, что доверительный интервал для вероятности успеха в схеме Бернулли приближенно имеет вид

$$h - u_{1-\alpha/2}\sqrt{\frac{h(1-h)}{n}} < p < h + u_{1-\alpha/2}\sqrt{\frac{h(1-h)}{n}}. \ \triangleright \tag{4}$$

Пример 3. При проверке 100 деталей из большой партии обнаружено 10 бракованных деталей.

а) Найти 95 %-ный приближенный доверительный интервал для доли бракованных деталей во всей партии.

б) Какой минимальный объем выборки следует взять для того, чтобы с вероятностью 0,95 можно было утверждать, что доля бракованных деталей по всей партии отличается от частоты появления бракованных деталей в выборке не более чем на 1 %?

$\triangleleft$ а) Оценка доли бракованных деталей в партии по выборке равна $\tilde{p} = h = 10/100 = 0{,}1$. По таблице П1 находим квантиль $u_{1-\alpha/2} = u_{0,975} = 1{,}96$. По формуле (4) 95 %-ный доверительный интервал для доли бракованных деталей в партии приближенно имеет вид $0{,}041 < p < 0{,}159$.

б) Представим доверительный интервал (4) в виде неравенства

$$|h-p| < u_{1-\alpha/2}\sqrt{\frac{h(1-h)}{n}},$$

которое выполняется с вероятностью $\approx 1-\alpha = 0{,}95$. Так как по условию задачи $|h-p| \leqslant 0{,}01$, то для определения n получим неравенство

$$u_{0,975}\sqrt{\frac{h(1-h)}{n}} \leqslant 0{,}01.$$

Отсюда следует, что

$$1{,}96\sqrt{\frac{0{,}1(1-0{,}1)}{n}} \leqslant 0{,}01$$

и $n \geqslant (0{,}3 \cdot 196)^2 = 3457{,}44$. Значит, минимальный объем выборки $n = 3458$. $\triangleright$

19.183. Из большой партии транзисторов одного типа были случайным образом отобраны и проверены 100 штук. У 36 транзисторов коэффициент усиления оказался меньше 10. Найти 95 %-ный доверительный интервал для доли таких транзисторов во всей партии.

19.184. С автоматической линии, производящей подшипники, было отобрано 400 штук, причем 10 оказалось бракованными. Найти 90 %-ный доверительный интервал для вероятности появления бракованного подшипника. Сколько подшипников надо проверить, чтобы с вероятностью 0,9973 можно было утверждать, что вероятность появления бракованного подшипника не отличается от частоты более чем на 5 %?

19.185. В 10 000 сеансах игры с автоматом выигрыш появился 4000 раз. Найти 95 %-ный доверительный интервал для вероятности выигрыша. Сколько сеансов игры следует провести, чтобы с вероятностью 0,99 вероятность выигрыша отличалась от частоты не более чем на 1 %?

19.186. При осмотре 60 ящиков обнаружено 10 поврежденных. Найти 90 %-ный доверительный интервал для доли поврежденных ящиков во всей партии.

19.187. Из урны, содержащей неотличимые на ощупь черные и белые шары в неизвестной пропорции, случайным образом извлекается 100 шаров (с возвращением). Найти: а) 90 %-ный и б) 95 %-ный доверительные интервалы для доли черных шаров, если среди вынутых шаров оказалось 30 черных.

19.188. Для проверки утверждения о том, что вероятность отказа прибора p равна 0,01, было проведено испытание 100 приборов, при этом один из приборов отказал. Построить 95%-ную верхнюю границу одностороннего доверительного интервала для p по этим данным.

19.189*. Пусть $x_1, \ldots, x_n$ — выборка из генеральной совокупности с конечным математическим ожиданием m и дисперсией σ^2. Показать, что если σ^2 известна, то доверительный интервал для m при достаточно больших n приближенно имеет вид

$$\overline{x} - u_{-\alpha/2}\,\frac{\sigma}{\sqrt{n}} < m < \overline{x} + u_{1-\alpha/2}\,\frac{\sigma}{\sqrt{n}}.$$

19.190*. Пусть $x_1, \ldots, x_n$ — выборка из генеральной совокупности, имеющей распределение Пуассона с неизвестным параметром λ. Показать, что при достаточно больших n доверительный интервал для параметра λ приближенно имеет вид

$$\overline{x} - u_{1-\alpha/2}\sqrt{\frac{\overline{x}}{n}} < \lambda < \overline{x} + u_{1-\alpha/2}\sqrt{\frac{\overline{x}}{n}}.$$

19.191. На каждой из 36 АТС города в период с двух до трех часов было зафиксировано в среднем 2 вызова. Считая, что число вызовов для каждой АТС имеет распределение Пуассона с одним и тем же параметром λ, приближенно найти доверительный интервал для λ с доверительной вероятностью 0,9.

19.192. Среднее число сбоев в сутки для 100 ЭВМ одного типа равно 2,3. В предположении, что число сбоев имеет распределение Пуассона с параметром λ, приближенно найти 95%-ный доверительный интервал для λ.

3. Доверительные интервалы для коэффициента корреляции ρ. Пусть выборка (x_i, y_i), $i = 1, 2, \ldots, n$, получена из генеральной совокупности, имеющей двумерное нормальное распределение, и r — выборочный коэффициент корреляции, вычисляемый по формуле (12) § 1. При достаточно больших n статистика

$$Z = \frac{1}{2}\ln\frac{1+r}{1-r} = \operatorname{Arth} r$$

имеет приближенно нормальное распределение

$$N\left(\operatorname{Arth} r,\ \frac{1}{\sqrt{n-3}}\right).$$

Доверительный интервал для $\operatorname{Arth}\rho$ имеет вид

$$\operatorname{Arth} r - \frac{u_{1-\alpha/2}}{\sqrt{n-3}} < \operatorname{Arth}\rho < \operatorname{Arth} r + \frac{u_{1-\alpha/2}}{\sqrt{n-3}}. \tag{5}$$

Доверительный интервал для ρ вычисляется с помощью таблиц гиперболического тангенса $\rho = \operatorname{th} z$ (см. таблицу П8).

Пример 4. Выборочный коэффициент корреляции, вычисленный по выборке объема 10, $r = -0{,}64$. Найти 90%-ный доверительный интервал для коэффициента корреляции ρ.

◁ Из таблицы П8 находим $\operatorname{Arth}(-0{,}64) = -\operatorname{Arth} 0{,}64 = -0{,}76$. Так как $u_{0,95} = 1{,}645$, то доверительный интервал для $\operatorname{Arth}\rho$ по формуле (5) имеет вид

$$-0{,}76 - \frac{1{,}645}{\sqrt{10-3}} < \operatorname{Arth}\rho < -0{,}76 + \frac{1{,}645}{\sqrt{10-3}},$$

т.е.

$$-1{,}38 < \operatorname{Arth}\rho < -0{,}14.$$

Снова обращаясь к таблице П8, получим 90%-ный доверительный интервал для коэффициента корреляции:

$$-0{,}881 < \rho < -0{,}139. \quad ▷$$

Построить доверительные интервалы для коэффициентов корреляции двумерной нормально распределенной совокупности по следующим данным:

19.193. $r = 0{,}687$, $n = 50$, $1 - \alpha = 0{,}95$.

19.194. $r = 0{,}71$, $n = 28$, $1 - \alpha = 0{,}95$.

19.195. $r = -0{,}65$, $n = 12$, $1 - \alpha = 0{,}99$ и $1 - \alpha = 0{,}95$.

19.196. $r = 0{,}14$, $n = 300$, $1 - \alpha = 0{,}99$ и $1 - \alpha = 0{,}95$.

19.197. $r = -0{,}36$, $n = 28$, $1 - \alpha = 0{,}99$ и $1 - \alpha = 0{,}95$.

19.198. Построить доверительный интервал для коэффициента корреляции по выборке, полученной в задаче 19.79 при $1-\alpha = 0{,}9$.

§ 4. Проверка статистических гипотез

1. Основные понятия. Проверка гипотез о параметрах нормально распределенной генеральной совокупности. Во многих случаях результаты наблюдений используются для проверки предположений (гипотез) относительно тех или иных свойств распределения генеральной совокупности. В частности, такого рода задачи возникают при сравнении различных технологических процессов или методов обработки по определенным измеряемым признакам, например, по точности, производительности и т.д.

Пусть X — наблюдаемая дискретная или непрерывная случайная величина. *Статистической гипотезой* H называется предположение относительно параметров или вида распределения случайной величины X. Статистическая гипотеза H называется *простой*, если она однозначно определяет распределение случайной величины X; в противном случае гипотеза H называется *сложной*. Например, простой гипотезой является предположение о том, что случайная величина X распределена по нормальному закону $N(0,\,1)$, если же высказывается предположение, что случайная величина X имеет нормальное распределение $N(m,\,1)$, где

$a \leqslant m \leqslant b$, то это сложная гипотеза. Другим примером сложной гипотезы является предположение о том, что непрерывная случайная величина X с вероятностью $1/3$ принимает значение из интервала $(1; 5)$; в этом случае распределение случайной величины X может быть любым из класса непрерывных распределений.

Часто распределение случайной величины X известно, и по выборке наблюдений необходимо проверить предположения о значении параметров этого распределения. Такие гипотезы называются *параметрическими*. В этом параграфе (за исключением п. 5) рассматривается проверка параметрических гипотез. Методы проверки гипотез другого типа (например, о виде распределения, независимости и др.) приводятся в § 6 и § 8.

Проверяемая гипотеза называется *нулевой гипотезой* и обозначается H_0. Наряду с гипотезой H_0 рассматривают одну из альтернативных (конкурирующих) гипотез H_1. Например, если проверяется гипотеза о равенстве параметра θ некоторому заданному значению θ_0, т.е. H_0: $\theta = \theta_0$, то в качестве альтернативной гипотезы можно рассмотреть одну из следующих гипотез: $H_1^{(1)}$: $\theta > \theta_0$; $H_1^{(2)}$: $\theta < \theta_0$; $H_1^{(3)}$: $\theta \neq \theta_0$; $H_1^{(4)}$: $\theta = \theta_1$, где θ_1 — заданное значение, $\theta_1 \neq \theta_0$. Выбор альтернативной гипотезы определяется конкретной формулировкой задачи.

Правило, по которому принимается решение принять или отклонить гипотезу H_0, называется *критерием* K. Так как решение принимается на основе выборки наблюдений случайной величины X, необходимо выбрать подходящую статистику, называемую в этом случае *статистикой Z критерия* K. При проверке простой параметрической гипотезы H_0: $\theta = \theta_0$ в качестве статистики критерия выбирают ту же статистику, что и для оценки параметра θ, т.е. $\tilde{\theta}$.

Проверка статистической гипотезы основывается на принципе, в соответствии с которым маловероятные события считаются невозможными, а события, имеющие большую вероятность, считаются достоверными. Этот принцип можно реализовать следующим образом. Перед анализом выборки фиксируется некоторая малая вероятность α, называемая *уровнем значимости*. Пусть V — множество значений статистики Z, а $V_к \subseteq V$ — такое подмножество, что при условии истинности гипотезы H_0 вероятность попадания статистики критерия в $V_к$ равна α, т.е. $\mathrm{P}\,[Z \in V_к / H_0] = \alpha$.

Обозначим $z_в$ выборочное значение статистики Z, вычисленное по выборке наблюдений. Критерий формулируется следующим образом: отклонить гипотезу H_0, если $z_в \in V_к$; принять гипотезу H_0, если $z_в \in$ $\in V \backslash V_к$. Критерий, основанный на использовании заранее заданного уровня значимости, называют *критерием значимости*. Множество $V_к$ всех значений статистики критерия Z, при которых принимается решение отклонить гипотезу H_0, называется *критической областью*; область $V \backslash V_к$ называется *областью принятия* гипотезы H_0.

Уровень значимости α определяет «размер» критической области $V_к$. Положение критической области на множестве значений статистики Z зависит от формулировки альтернативной гипотезы H_1. Например, если проверяется гипотеза $H_0: \theta = \theta_0$, а альтернативная гипотеза H_1 формулируется как $H_1: \theta > \theta_0$ $(\theta < \theta_0)$, то критическая область размещается на правом (левом) «хвосте» распределения статистики Z, т.е. имеет вид неравенства $Z > z_{1-\alpha}(Z < z_\alpha)$, где $z_{1-\alpha}$ и z_α — квантили распределения статистики Z при условии, что верна гипотеза H_0. В этом случае критерий называется *односторонним*, соответственно правосторонним и левосторонним. Если альтернативная гипотеза формулируется

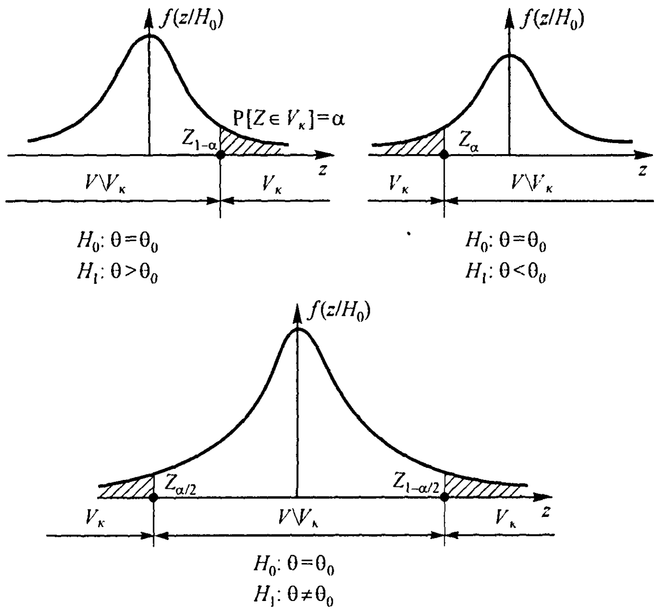


Рис. 34

как $H_1: \theta \neq \theta_0$, то критическая область размещается на обоих «хвостах» распределения Z, т.е. определяется совокупностью неравенств $Z < z_{\alpha/2}$ и $Z > z_{1-\alpha/2}$; в этом случае критерий называется *двусторонним*.

На рис. 34 показано расположение критической области V_κ для различных альтернативных гипотез. Здесь $f_Z(z/H_0)$ — плотность распределения статистики Z критерия при условии, что верна гипотеза H_0, $V \backslash V_\kappa$ — область принятия гипотезы, $\mathrm{P}\left[Z \in V \backslash V_\kappa\right] = 1 - \alpha$.

Таким образом, проверка параметрической статистической гипотезы при помощи критерия значимости может быть разбита на следующие этапы:

1) сформулировать проверяемую (H_0) и альтернативную (H_1) гипотезы;

2) назначить уровень значимости α;

3) выбрать статистику Z критерия для проверки гипотезы H_0;

4) определить выборочное распределение статистики Z при условии, что верна гипотеза H_0;

5) в зависимости от формулировки альтернативной гипотезы определить критическую область V_κ одним из неравенств $Z > z_{1-\alpha}$, $Z < z_\alpha$ или совокупностью неравенств $Z > z_{1-\alpha/2}$ и $Z < z_{\alpha/2}$;

6) получить выборку наблюдений и вычислить выборочное значение $z_в$ статистики критерия;

7) принять статистическое решение:

если $z_в \in V_к$, то отклонить гипотезу H_0 как не согласующуюся с результатами наблюдений;

если $z_в \in V \backslash V_к$, то принять гипотезу H_0, т.е. считать, что гипотеза H_0 не противоречит результатам наблюдений.

Замечание. Обычно на этапах 4–7 используют статистику, квантили которой табулированы: статистику с нормальным распределением $N(0, 1)$, статистику Стьюдента, статистику χ^2 или статистику Фишера. Однако интерпретацию решения и вычисление вероятностей ошибок, допускаемых при проверке гипотез, удобно проводить для статистики, являющейся непосредственной оценкой параметра θ, т.е. статистики $\tilde{\theta}$.

Пример 1. По паспортным данным автомобильного двигателя расход топлива на 100 км пробега составляет 10 л. В результате изменения конструкции двигателя ожидается, что расход топлива уменьшится. Для проверки проводятся испытания 25 случайно отобранных автомобилей с модернизированным двигателем, причем выборочное среднее расходов топлива на 100 км пробега по результатам испытаний составило $\overline{x} = 9{,}3$ л. Предположим, что выборка расходов топлива получена из нормально распределенной генеральной совокупности со средним m и дисперсией $\sigma^2 = 4$ л^2. Используя критерий значимости, проверить гипотезу, утверждающую, что изменение конструкции двигателя не повлияло на расход топлива.

◁ Проверяется гипотеза о среднем (m) нормально распределенной генеральной совокупности. Проверку гипотезы проведем по этапам:

1) проверяемая гипотеза $H_0: \ m = 10$, альтернативная гипотеза $H_1: \ m < 10$;

2) выберем уровень значимости $\alpha = 0{,}05$;

3) в качестве статистики критерия используем оценку математического ожидания — выборочное среднее $\overline{X}$;

4) так как выборка получена из нормально распределенной генеральной совокупности, выборочное среднее также имеет нормальное распределение с дисперсией $\dfrac{\sigma^2}{n} = \dfrac{4}{25}$. При условии, что верна гипотеза H_0, математическое ожидание этого распределения равно 10. Нормированная статистика критерия $U = \dfrac{\overline{X} - 10}{\sqrt{4/25}}$ имеет нормальное распределение $N(0, 1)$.

5) альтернативная гипотеза $H_1: m < 10$ предполагает уменьшение расхода топлива, следовательно, нужно использовать односторонний критерий. Критическая область определяется неравенством $U < u_\alpha$. По таблице П1 находим $u_{0,05} = -u_{0,95} = -1{,}645$;

6) выборочное значение нормированной статистики критерия равно $u_в = \dfrac{9{,}3 - 10}{\sqrt{4/25}} = -1{,}75$;

7) статистическое решение: так как выборочное значение статистики критерия принадлежит критической области, гипотеза H_0 отклоняется: следует считать, что изменение конструкции двигателя привело к уменьшению расхода топлива.

Граница $\overline{x}_к$ критической области для исходной статистики X критерия может быть получена из соотношения

$$\frac{\overline{x}_к - 10}{\sqrt{4/25}} = -1{,}645,$$

откуда получаем $\overline{x}_к = 9{,}342$, т.е. критическая область для статистики $\overline{X}$ определяется неравенством $\overline{X} < 9{,}342$. $\rhd$

Статистическое решение может быть ошибочным. При этом различают ошибки первого и второго рода. *Ошибкой первого рода* называется ошибка, состоящая в том, что гипотеза H_0 отклоняется, в то время как она верна. Вероятность ошибки первого рода равна вероятности попадания статистики критерия в критическую область при условии, что верна гипотеза H_0, т.е. равна уровню значимости α:

$$\mathrm{P}\,[Z \in V_к/H_0] = \alpha. \tag{1}$$

Так в примере 1 вероятность ошибки первого рода равна 0,05.

Ошибка второго рода происходит в том случае, если гипотеза H_0 принимается, но в действительности верна альтернативная гипотеза H_1. Вероятность ошибки второго рода β можно вычислить (при простой альтернативной гипотезе H_1) по формуле

$$\beta = \mathrm{P}\,[Z \in V \backslash V_к/H_1]. \tag{2}$$

Пример 2. В условиях примера 1 предположим, что наряду с гипотезой H_0: $m = 10$ л рассматривается альтернативная гипотеза H_1: $m = 9$ л. В качестве статистики критерия снова возьмем выборочное среднее $\overline{X}$. Предположим, что критическая область задана следующим неравенством $\overline{X} < 9{,}44$ л. Найти вероятности ошибок первого и второго рода для критерия с такой критической областью.

$\lhd$ Найдем вероятность ошибки первого рода. Статистика $\overline{X}$ критерия при условии, что верна гипотеза H_0: $m = 10$, имеет нормальное распределение $N(10, \sqrt{4/25})$. По формуле (1), используя таблицу П1, находим

$$\alpha = \mathrm{P}\,[\overline{X} < 9{,}44/H_0\colon m = 10] = \Phi\left(\frac{9{,}44 - 10}{\sqrt{4/25}}\right) =$$
$$= \Phi\,(-1{,}4) = 1 - \Phi\,(1{,}4) \approx 0{,}08.$$

Это означает, что принятый критерий классифицирует $\sim 8\,\%$ автомобилей, имеющих расход 10 л на 100 км пробега, как автомобили, имеющие меньший расход топлива.

При условии, что верна гипотеза H_1: $m = 9$, статистика $\overline{X}$ имеет нормальное распределение $N(9, \sqrt{4/25})$. Вероятность ошибки второго рода по формуле (2) равна

$$\beta = \mathrm{P}\,[\overline{X} \geqslant 9{,}44/H_1\colon m = 9] = 1 - \Phi\left[\frac{9{,}44 - 9}{\sqrt{4/25}}\right] = 1 - \Phi\,(1{,}1) \approx 0{,}136.$$

Следовательно, в соответствии с принятым критерием 13,6 % автомобилей, имеющих расход топлива 9 л на 100 км пробега, классифицируются

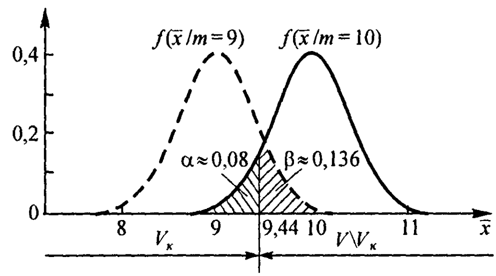


Рис. 35

как автомобили, имеющие расход 10 л. Вероятности ошибок первого и второго рода показаны в виде заштрихованных областей под кривыми плотностей распределения статистики критерия на рис. 35. ▷

При заданной вероятности α ошибки первого рода вероятность ошибки второго рода может быть уменьшена путем увеличения объема выборки. Если при этом вероятность ошибки второго рода не должна превышать заданного значения β, то минимальный объем выборки n можно найти из решения системы:

$$\begin{cases} \mathrm{P}\,[Z \in V_к/H_0] = \alpha, \\ \mathrm{P}\,[Z \in V\backslash V_к/H_1] \leqslant \beta. \end{cases}$$

Аналитическое решение этой системы возможно только в простейших случаях.

Пример 3. Какой минимальный объем выборки n следует взять в условиях примера 1, чтобы при проверке гипотезы H_0: $m = 10$ л против альтернативной гипотезы H_1: $m = 9$ л ошибка первого рода была равна $\alpha = 0{,}01$, а ошибка второго рода не превышала 0,1? Какова критическая область в этом случае?

◁ Так как в альтернативной гипотезе H_1 предполагается меньшее значение параметра m, то критическая область $V_к$ определяется неравенством $\overline{X} < \overline{x}_\kappa$. По условию задачи имеем:

$$\begin{cases} \mathrm{P}\,[\overline{X} < \overline{x}_\kappa / H_0\colon\ m = 10] = \Phi\left(\dfrac{\overline{x}_\kappa - 10}{\sqrt{4/n}}\right) = 0{,}01, \\[2ex] \mathrm{P}\,[\overline{X} \geqslant \overline{x}_\kappa / H_1\colon\ m = 9] = 1 - \Phi\left(\dfrac{\overline{x}_\kappa - 9}{\sqrt{4/n}}\right) \leqslant 0{,}1. \end{cases}$$

Эту систему можно записать так:

$$\begin{cases} \dfrac{\overline{x}_\kappa - 10}{2}\sqrt{n} = u_{0{,}01} = -2{,}326, \\[2ex] \dfrac{\overline{x}_\kappa - 9}{2}\sqrt{n} \geqslant u_{0{,}9} = 1{,}282. \end{cases}$$

Исключая $\overline{x}_\kappa$, получим, что $n \geqslant 53$. Подставляя наименьшее значение n в первое уравнение системы, найдем границу критической области:

$$\overline{x}_\kappa = 10 - \frac{2 \cdot 2{,}326}{\sqrt{53}} = 9{,}361.$$

Следовательно, критическая область V_κ определяется неравенством $\overline{X} < 9{,}361$. ▷

19.199. Станок-автомат изготовляет шарики диаметром 10 мм. Продукция станка контролируется по величине X — отклонению диаметра шарика от номинального размера 10 мм. Предположим, что X — нормально распределенная случайная величина с математическим ожиданием m и дисперсией $\sigma^2 = 0{,}1\,\text{мм}^2$. Рассмотреть следующие гипотезы:

$H^{(1)}$: $m = 0$,

$H^{(2)}$: $m \neq 0$,

$H^{(3)}$: $-1 \leqslant m \leqslant 1$,

$H^{(4)}$: материал, используемый для изготовления шариков, содержит специальные присадки.

Определить, какие из гипотез $H^{(1)}$–$H^{(4)}$ являются статистическими, какие статистические гипотезы являются простыми, а какие сложными?

19.200. При подбрасывании монеты 10 раз герб выпал X раз. Классифицировать следующие гипотезы:

$H^{(1)}$: X имеет биномиальное распределение $B\,(10,\ 1/2)$;

$H^{(2)}$: X имеет биномиальное распределение $B\,(10,\ p)$, причем $1/3 \leqslant p \leqslant 2/3$;

$H^{(3)}$: $\mathrm{P}\,[X \leqslant 3] > 1/2$;

$H^{(4)}$: монета не симметрична.

19.201. Считается, что новое антикоррозийное покрытие имеет эффективность 99 %, если среди 20 испытанных образцов нет ни одного с признаками коррозии; в противном случае эффективность покрытия принимается равной 90 %. Пусть p — вероятность появления признаков коррозии у одного образца. Предположим, что образцы обрабатываются и испытываются независимо один от другого. Рассмотрим нулевую гипотезу $H_0 : p = 0{,}10$ и альтернативную гипотезу $H_1 : p = 0{,}01$. Ответить на следующие вопросы:

а) Какая статистика критерия используется в данной задаче, каковы ее распределение и область изменения?

б) Какова критическая область критерия?

в) В чем состоят ошибки первого и второго рода и чему равны их вероятности?

19.202. В каких случаях и какого рода ошибка допущена при проверке гипотезы H_0 при помощи некоторого критерия, если:

а) H_0 верна, но должна быть отвергнута согласно критерию;

б) H_0 неверна, но должна быть принята согласно критерию;

в) H_0 верна и должна быть принята согласно критерию;

г) H_0 неверна и должна быть отвергнута согласно критерию?

19.203. Наблюдаемый объект может быть либо своим, либо объектом противника. Система обнаружения относит объект к одному из классов по результатам нескольких замеров определенных характеристик. Сформулировать нулевую и альтернативную гипотезы, если в результате ошибки первого рода происходит «пропуск цели». В чем состоит ошибка второго рода?

19.204. Проверка функционирования устройства осуществляется специальным тестом. Если устройство функционирует правильно, то вероятность прохождения теста равна 0,99; в противном случае вероятность прохождения теста равна 0,40. Устройство допускается к работе, если тест проходит 5 раз подряд. В предположении, что число прохождений теста подчиняется биномиальному распределению, ответить на следующие вопросы:

а) Какова область изменения и критическая область статистики критерия? Какое распределение имеет статистика критерия?

б) Как сформулировать нулевую гипотезу, если ошибка первого рода состоит в отклонении правильно функционирующего устройства?

в) Какова альтернативная гипотеза и в чем состоит ошибка второго рода?

г) Чему равны вероятности ошибок первого и второго рода?

19.205. Большая партия изделий может содержать некоторую долю дефектных. Поставщик утверждает, что эта доля составляет 5 %; покупатель предполагает, что доля дефектных изделий равна 10 %. Условия поставки: из партии случайным образом отбирается и проверяется 10 изделий; партия принимается на условиях поставщика, если при проверке обнаружено не более одного дефектного изделия; в противном случае партия принимается на условиях покупателя. Сформулировать эту задачу в терминах теории проверки статистических гипотез и ответить на следующие вопросы:

а) Каковы статистика критерия, область ее значений, критическая область?

б) Какое распределение имеет статистика критерия?

в) В чем состоят проверяемая и альтернативная гипотезы?

г) В чем состоят ошибки первого и второго рода и каковы их вероятности?

19.206. Из продукции автомата, обрабатывающего болты с номинальным значением контролируемого размера $m_0 = 40$ мм, была взята выборка болтов объема $n = 36$. Выборочное среднее контролируемого размера $\overline{x} = 40{,}2$ мм. Результаты предыдущих измерений дают основание предполагать, что действительные размеры болтов образуют нормально распределенную совокупность с дисперсией $\sigma^2 = 1$ мм^2. Можно ли по результатам проведенного выборочного обследования утверждать, что контролируемый размер в продукции автомата не имеет положительного смещения по отношению к номинальному размеру? Принять $\alpha = 0{,}01$. Какова критическая область в этом случае?

19.207. Предположим, что в условиях задачи 19.206 партия болтов с номинальным размером $m_0 = 40$ мм бракуется, если выборочное среднее контролируемого размера будет больше 40,1 мм. Найти вероятности ошибок первого и второго рода при альтернативной гипотезе $H_1 : m = 40{,}3$, если решение принимается по выборке объема $n = 36$.

19.208. Решить задачу 19.207, если партия болтов бракуется при выполнении одного из неравенств $\overline{x} > 40{,}1$ мм и $\overline{x} < 39{,}9$ мм, где $\overline{x}$ — выборочное среднее контролируемого размера.

19.209. В условиях задачи 19.207 для проверки гипотезы $H_0 : m = 40$ мм против альтернативной гипотезы: $H_1 : m = 40{,}3$ мм предлагается выбрать такую критическую область $V_к : 40{,}15 < \overline{x} < 40{,}20$. Найти вероятности ошибок первого и второго рода. Показать эти ошибки на графиках плотностей распределения статистики $\overline{X}$: $f(\overline{x}/H_0)$ и $f(\overline{x}/H_1)$.

19.210. В условиях задачи 19.206 какой минимальный объем выборки n следует взять, чтобы при проверке гипотезы $H_0 : m = 40$ мм против альтернативной гипотезы $H_1 : m = 40{,}3$ мм при вероятности ошибки первого рода $\alpha = 0{,}10$ вероятность ошибки второго рода не превосходила 0,10? Какая критическая область соответствует этим условиям при объеме выборки n?

Проверка статистических гипотез с использованием критериев значимости может быть проведена на основе *доверительных интервалов*. При этом одностороннему критерию значимости соответствует односторонний доверительный интервал, а двустороннему критерию значимости — двусторонний доверительный интервал. Гипотеза H_0 принимается, если значение θ_0 накрывается соответствующим доверительным интервалом; в противном случае гипотеза H_0 отклоняется.

Если проверяется гипотеза H_0: $\theta_1 = \theta_2$, то рассматривается доверительный интервал для разности $\theta_1 - \theta_2$. Гипотеза H_0 принимается, если доверительный интервал для разности параметров $\theta_1 - \theta_2$ накрывает нулевое значение. Исключение составляет проверка гипотезы о равенстве дисперсий $H_0 : \sigma_1^2 = \sigma_2^2$, так как доверительный интервал строится для отношения дисперсий, то гипотеза H_0 в этом случае принимается, если доверительный интервал накрывает значение, равное единице.

Пример 4. В условиях примера 1 проверить гипотезу $H_0 : m = 10$ л при альтернативной гипотезе $H_1 : m < 10$ л на уровне значимости $\alpha = 0{,}05$, используя доверительный интервал для параметра m.

$\lhd$ Найдем границу m_2 левостороннего доверительного интервала $(-\infty, m_2)$ для параметра m при доверительной вероятности $1 - \alpha = 0{,}95$ (§ 3, п. 1). Используя выборочное среднее $\overline{x} = 9{,}3$ и значение квантили $u_{0,95} = 1{,}645$, получим

$$m_2 = \overline{x} + \frac{\sigma}{\sqrt{n}} u_{1-\alpha} = 9{,}3 + \frac{2}{\sqrt{25}} \cdot 1{,}645 = 9{,}958.$$

Так как значение $m = 10$ не накрывается интервалом $(-\infty; 9{,}958)$, то гипотезу H_0 следует отклонить, что совпадает с результатом, полученным при решении примера 1. $\rhd$

19.211. Решить задачу 19.206, используя доверительный интервал для параметра m.

В задачах 19.212–19.244 предполагается, что выборки получены из генеральных совокупностей, имеющих нормальное распределение, либо распределение, достаточно близкое к нормальному. При решении этих задач следует использовать данные о критериях значимости, приведенные в таблицах 4.1 и 4.2.

19.212. В соответствии с техническими условиями среднее время безотказной работы для приборов из большой партии должно составлять не менее 1000 часов со среднеквадратичным отклонением (с. к. о.) 100 часов. Выборочное среднее времени безотказной работы для случайно отобранных 25 приборов оказалось равным 970 часам. Предположим, что с. к. о. времени безотказной работы для приборов в выборке совпадает с с. к. о. во всей партии. Можно ли считать, что вся партия приборов не удовлетворяет техническим условиям, если: а) $\alpha = 0{,}10$; б) $\alpha = 0{,}01$?

Таблица 4.1

Критерии зависимости для проверки гипотез о дисперсиях нормально распределенной генеральной совокупности

Проверяемая гипотеза H_0	Предположение относительно m	Статистика Z критерия [1])	Распределение $Z: f(z/H_0)$	Область принятия гипотезы H_0 для двустороннего критерия	Альтернативная гипотеза и область принятия гипотезы H_0 для правостороннего критерия
$\sigma^2 = \sigma_0^2$	m известно	$\dfrac{nS_0^2}{\sigma_0^2}$	$\chi^2(n)$	$\chi^2_{\alpha/2}(n) < \dfrac{nS_0^2}{\sigma_0^2} < \chi^2_{1-\alpha/2}(n)$	$H_1: \sigma^2 > \sigma_0^2$ $\dfrac{nS_0^2}{\sigma_0^2} < \chi^2_{1-\alpha/2}(n)$
	m известно, $\tilde{m} = \overline{X}$	$\dfrac{(n-1)S^2}{\sigma_0^2}$	$\chi^2(n-1)$	$\chi^2_{\alpha/2}(n-1) < \dfrac{(n-1)s^2}{\sigma_0^2} <$ $< \chi^2_{1-\alpha/2}(n-1)$	$H_1: \sigma^2 > \sigma_0^2$ $\dfrac{(n-1)s^2}{\sigma_0^2} < F_{1-\alpha}(n_1 n_2)$
$\sigma_1^2 = \sigma_2^2$	m_1 и m_2 известны	S_{01}^2/S_{02}^2 $s_{01}^2 > s_{02}^2$	$F(n_1 n_2)$	$\dfrac{s_{01}^2}{s_{02}^2} < F_{1-\alpha/2}(n_1, n_2)$	$H_1: \sigma_1^2 > \sigma_2^2$ $\dfrac{s_{01}^2}{s_{02}^2} < F_{1-\alpha/2}(n_1, n_2)$
$\sigma_1^2 = \sigma_2^2$	m_1 и m_2 известны $\tilde{m}_1 = \overline{X}_1$ $\tilde{m}_2 = \overline{X}_2$	S_{01}^2/S_{02}^2 $s_{01}^2 > s_{02}^2$	$F(n_1-1, n_2-1)$	$\dfrac{s_{01}^2}{s_{02}^2} < F_{1-\alpha/2}(n_1-1, n_2-1)$	$H_1: \sigma_1^2 > \sigma_2^2$ $\dfrac{s_{01}^2}{s_{02}^2} < F_{1-\alpha/2}(n_1-1, n_2-1)$

Таблица 4.1 (*продолжение*)

Критерий Бартлетта для сравнения дисперсий нескольких совокупностей

H_0	Предположение относительно m_i	Статистика Z критерия [2])	$f(z/H_0)$	Область принятия гипотезы
$\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_l^2$	m_i неизвестны $i = 1, 2, \ldots, l$	$\frac{1}{c}\left[\sum_{i=1}^{l}(n_i - 1)\ln S^2 - \sum_{i=1}^{l}(n_i - 1)\ln S_i^2\right]$, где $c = 1 + \frac{1}{3(l-1)}\left[\sum_{i=1}^{l}\frac{1}{n_i - 1} - \frac{1}{\sum_{i=1}^{l}(n_i - 1)}\right]$, $S^2 = \frac{\sum(n_i - 1)S_i^2}{\sum(n_i - 1)}$	$\chi^2(l-1)$	$z_в < \chi^2_{1-\alpha}(l-1)$

[1,2]) Значение статистики Z приводится при условии, что верна гипотеза H_0.

9*

Таблица 4.2

Критерии зависимости для проверки гипотез о средних нормально распределенной генеральной совокупности

Проверяемая гипотеза H_0	Предположение относительно σ^2	Статистика Z критерия [3])	Распределение Z: $f(z/H_0)$	Область принятия гипотезы H_0 для двустороннего критерия	Альтернативная гипотеза и область принятия гипотезы H_0 для правостороннего критерия
$m = m_0$	σ^2 известна	$\dfrac{\overline{X} - m_0}{\sigma/\sqrt{n}}$	$N(0, 1)$	$\dfrac{\lvert\overline{x} - m_0\rvert}{\sigma/\sqrt{n}} < u_{1-\alpha/2}$	$H_1 : m > m_0$ $\dfrac{\lvert\overline{x} - m_0\rvert}{\sigma/\sqrt{n}} < u_{1-\alpha}$
	σ^2 известна,	$\dfrac{\overline{X} - m_0}{S/\sqrt{n}}$	$T(n-1)$	$\dfrac{\lvert\overline{x} - m_0\rvert}{s/\sqrt{n}} < t_{1-\alpha/2}(n-1)$	$H_1 : m > m_0$ $\dfrac{\lvert\overline{x} - m_0\rvert}{s/\sqrt{n}} < t_{1-\alpha}(n-1)$
$m_1 = m_2$	σ_1^2 и σ_2^2 известны	$\dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$	$N(0, 1)$	$\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} < u_{1-\alpha/2}$	$H_1 : m_1 > m_2$ $\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} < u_{1-\alpha}$

Таблица 4.2 (*продолжение*)

Проверяемая гипотеза H_0	Предположение относительно σ^2	Статистика Z критерия [4]	Распределение Z: $f(z/H_0)$	Область принятия гипотезы H_0 для двустороннего критерия	Альтернативная гипотеза и область принятия гипотезы H_0 для правостороннего критерия
$m_1 = m_2$	σ_1^2 и σ_2^2 известны причем гипотеза H_0: $\sigma_1^2 = \sigma_2^2$ принимается, $\sigma_1^2 = S_1^2$, $\sigma_2^2 = S_2^2$	$\dfrac{\overline{X}_1 - \overline{X}_2}{S\sqrt{1/n_1 + 1/n_2}}$, где $S = \left(\dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}\right)$	$T(n_1+n_2-2)$	$\dfrac{\overline{x}_1 - \overline{x}_2}{s\sqrt{1/n_1 + 1/n_2}} <$ $< t_{1-\alpha/2} \times$ $\times (n_1+n_2-2)$	$H_1: m_1 > m_2$; $\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{1/n_1 + 1/n_2}} <$ $< t_{1-\alpha}(n_1+n_2-2)$
$m_1 = m_2$	σ_1^2 и σ_2^2 известны причем гипотеза H_0: $\sigma_1^2 = \sigma_2^2$ отклоняется, $\sigma_1^2 = S_1^2$, $\sigma_2^2 = S_2^2$	$\dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$	$T(k)$, где $k = \dfrac{(s_1^2/n_1 + s_2^2/n_2)}{\dfrac{s_1^2/n_1}{n_1-1} + \dfrac{s_2^2/n_2}{n_2-1}}$	$\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} <$ $< t_{1-\alpha/2}(k)$	$H_1: m_1 > m_2$; $\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} <$ $< t_{1-\alpha}(k)$

[3, 4]) Значение статистики Z приводится при условии, что верна гипотеза H_0.

19.213. Решить задачу 19.212 при условии, что оценка с. к. о. времени безотказной работы, вычисленная по выборке, равна $s = 115$ часов.

19.214. Утверждается, что шарики, изготовленные станком-автоматом, имеют средний диаметр $d_0 = 10$ мм. Используя односторонний критерий при $\alpha = 0{,}05$, проверить эту гипотезу, если в выборке из $n = 16$ шариков средний диаметр оказался равным 10,3 мм, считая, что: а) дисперсия известна и равна $\sigma^2 = 1$ мм^2; б) оценка дисперсии, определенная по выборке, $s^2 = 1{,}21$ мм^2.

19.215. Из большой партии резисторов одного типа и номинала случайным образом отобраны 36 штук. Выборочное среднее величины сопротивления при этом оказалось равным 9,3 кОм. Используя двусторонний критерий при $\alpha = 0{,}05$, проверить гипотезу о том, что выборка взята из партии с номиналом 10 кОм, если:

а) дисперсия величины сопротивления известна и равна 4 кОм^2;

б) дисперсия величины сопротивления неизвестна, а выборочная дисперсия равна 6,25 кОм^2.

19.216. Решить задачу 19.215, используя доверительные интервалы для среднего значения величины сопротивления.

19.217. Технология производства некоторого вещества дает в среднем 1000 кг вещества в сутки с с. к. о. среднего, равным 80 кг. Новая технология производства в среднем дает 1100 кг вещества в сутки с тем же с. к. о. Можно ли считать, что новая технология обеспечивает повышение производительности, если: а) $\alpha = 0{,}05$; б) $\alpha = 0{,}10$?

19.218. В задаче 19.217 вычислить вероятность ошибки второго рода при альтернативной гипотезе, утверждающей, что производительность при новой технологии возросла и составляет 1200 кг вещества в сутки.

19.219. Ожидается, что добавление специальных веществ уменьшает жесткость воды. Оценки жесткости воды до и после добавления специальных веществ по 40 и 50 пробам соответственно показали средние значения жесткости (в градусах жесткости), равные 4,0 и 3,8 градуса. Дисперсия измерений в обоих случаях предполагается равной 0,25 град^2. Подтверждают ли эти результаты ожидаемый эффект? Принять $\alpha = 0{,}05$.

19.220. Решить задачу 19.219, используя метод доверительных интервалов.

19.221. Два штурмана определили пеленг маяка по нескольким замерам, используя различные пеленгаторы. Результаты замеров: $\overline{x}_1 = 70{,}2°$ при $n_1 = 4$ и $\overline{x}_2 = 70{,}5°$ при $n_2 = 9$. При помощи двустороннего критерия проверить при $\alpha = 0{,}05$ гипотезу о том, что различие результатов вызвано только случайными ошибками, если с. к. о. для обоих пеленгаторов известны и равны $\sigma_1 = 0{,}5°$ и $\sigma_2 = 1°$.

19.222. Решить задачу 19.221, используя метод доверительных интервалов.

19.223*. Из генеральных совокупностей с распределениями $N(m_1, \sigma)$ и $N(m_2, \sigma)$ получены две выборки объемов n_1 и n_2. Предлагается отклонить гипотезу $H_0 : m_1 = m_2$, если доверительные интервалы для m_1 и m_2 не пересекаются. Показать, что при доверительной вероятности $1 - \alpha$ уровень значимости этого критерия меньше α.

19.224. Точность наладки станка-автомата, производящего некоторые детали, характеризуется дисперсией длины деталей. Если эта величина будет больше $400\,\text{мкм}^2$, станок останавливается для наладки. Выборочная дисперсия длины 15 случайно отобранных деталей из продукции станка оказалась равной $s^2 = 680\,\text{мкм}^2$. Нужно ли производить наладку станка, если: а) уровень значимости $\alpha = 0{,}01$; б) уровень значимости $\alpha = 0{,}10$?

19.225. Новый метод измерения длины деталей был опробован на эталоне, причем дисперсия результатов измерений, определенная по 10 замерам, составила $100\,\text{мкм}^2$. Согласуется ли этот результат с утверждением: «дисперсия результатов измерений по предложенному методу не превосходит $50\,\text{мкм}^2$»? Принять $\alpha = 0{,}05$.

19.226. При применении определенной процедуры проверки коэффициента трения шины по мокрому асфальту установлено, что дисперсия результатов измерений этого коэффициента составляет 0,1. Выборочное значение дисперсии, вычисленное по результатам 25 измерений коэффициента трения, оказалось равным 0,20.

а) Используя двусторонний критерий, проверить гипотезу о том, что дисперсия результатов измерений коэффициента трения равна 0,1 при $\alpha = 0{,}1$.

б) Решить задачу, испольуя метод доверительных интервалов.

19.227. Два токарных автомата изготовляют детали по одному чертежу. Из продукции первого станка было отобрано $n_1 = 9$ деталей, а из продукции второго $n_2 = 11$ деталей. Выборочные дисперсии контрольного размера, определенные по этим выборкам, $s_1^2 = 5{,}9\,\text{мкм}^2$ и $s_2^2 = 23{,}2\,\text{мкм}^2$. Проверить гипотезу о равенстве дисперсий при $\alpha = 0{,}05$, если альтернативная гипотеза утверждает, что: а) дисперсии не равны; б) дисперсия размера для второго станка больше, чем для первого.

19.228. До наладки станка была проверена точность изготовления 10 втулок и найдено значение оценки дисперсии диаметра $s^2 = 9{,}6\,\text{мкм}^2$. После наладки подверглись контролю еще 15 втулок и получено новое значение оценки дисперсии $s^2 = 5{,}7\,\text{мкм}^2$. Можно ли считать, что в результате наладки станка точность изготовления деталей увеличилась? Принять $\alpha = 0{,}05$.

19.229 (сравнение средних). При измерении производительности двух агрегатов получены следующие результаты (в кг вещества за час работы):

№ замера	1	2	3	4	5
Агрегат A	14,1	10,1	14,7	13,7	14,0
Агрегат B	14,0	14,5	13,7	12,7	14,1

Можно ли считать, что производительности агрегатов A и B одинаковы, в предположении, что обе выборки получены из нормально распределенных генеральных совокупностей? Принять $\alpha = 0{,}10$.

◁ Проверяется гипотеза $H_0\colon\ m_1 = m_2$ при альтернативной гипотезе $H_1\colon\ m_1 \neq m_2$. Вычислим оценки средних и дисперсий:

$$\overline{x}_1 = 13,32, \quad \overline{x}_2 = 13{,}80, \quad s_1^2 \approx 3{,}37, \quad s_2^2 \approx 0{,}46.$$

Предварительно проверим гипотезу о равенстве дисперсий H_0: $\sigma_1^2 = \sigma_2^2$ (таблица 4.1, четвертая строка):

$$\frac{s_1^2}{s_2^2} \approx \frac{3{,}37}{0{,}46} \approx 7{,}33;$$

так как $F_{1-\alpha/2}(n_1 - 1,\, n_2 - 1) = F_{0{,}95}(4{,}4) = 6{,}39$ (таблица П7), то гипотеза о равенстве дисперсий отклоняется. Для проверки гипотезы о равенстве средних используем критерий из таблицы 4.2 (нижняя строка). Вычислим выборочное значение статистики критерия:

$$\frac{|\,x_1 - x_2\,|}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} = \frac{|\,13{,}32 - 13{,}80\,|}{\sqrt{\dfrac{3{,}37}{5} + \dfrac{0{,}46}{5}}} \approx 0{,}55.$$

Число степеней свободы $k \approx \dfrac{\left(\dfrac{3{,}37}{5} + \dfrac{0{,}46}{5}\right)^2}{\dfrac{\left(\dfrac{3{,}37}{5}\right)^2}{6} + \dfrac{\left(\dfrac{0{,}46}{5}\right)^2}{6}} - 2 \approx 6.$ Так как по таблице П6 $t_{0{,}95}(6) = 1{,}943$, гипотеза о равенстве средних принимается. ▷

19.230. Давление в камере контролируется по двум манометрам. Для сравнения точности этих приборов одновременно фиксируются их показания. По результатам 10 замеров выборочные оценки (в единицах шкалы приборов) оказались следующими:

$\overline{x}_1 = 15{,}3$, $\overline{x}_2 = 16{,}1$, $s_1^2 = 0{,}2$ и $s_2^2 = 0{,}15$. Используя двусторонний и односторонний критерии, проверить при $\alpha = 0{,}1$: а) гипотезу о равенстве дисперсий; б) гипотезу о равенстве средних.

19.231. На двух станках A и B производят одну и ту же продукцию, контролируемую по внутреннему диаметру изделия. Из продукции станка A была взята выборка из 16 изделий, а из продукции станка B — выборка из 25 изделий. Выборочные оценки средних и дисперсий контролируемых размеров $\overline{x}_A = 37{,}5$ мм при $s_A^2 = 1{,}21$ мм^2 и $\overline{x}_B = 36{,}8$ мм при $s_B^2 = 1{,}44$ мм^2. Используя двусторонний критерий, проверить гипотезу о равенстве математических ожиданий контролируемых размеров в продукции обоих станков, если: а) $\alpha = 0{,}05$; б) $\alpha = 0{,}10$.

19.232. Для третьего станка C, установленного в том же цехе, что и станки A и B из задачи 19.231, и производящего такую же продукцию, оценки среднего и дисперсии контролируемого размера, вычисленные по выборке из 13 изделий: $\overline{x}_C = 38{,}3$ мм, $s_C^2 = 3{,}08$ мм^2. Используя двусторонний критерий на уровне значимости $\alpha = 0{,}10$, проверить гипотезы о равенстве математических ожиданий контролируемого размера в продукции станков: а) A и C; б) B и C.

19.233. При исследовании влияния двух типов покрытия на удельную проводимость телевизионных трубок получены следующие результаты (в условных единицах):

№ трубки	1	2	3	4	5	6	$\sum x_i$	$\sum x_i^2$
I тип	6	5	12	9	10	–	42	386
II тип	14	11	0	5	6	8	44	442

Можно ли на основании этих данных считать, что тип покрытия влияет на удельную проводимость трубок? Принять $\alpha = 0{,}10$.

19.234. Чтобы определить, какое влияние оказывает температура окружающей среды на систематическую ошибку угломерного инструмента, проведены измерения горизонтального угла объекта δ утром ($t = 10°$ C) и днем ($t = 26°$ C). Результаты измерений δ (в угловых секундах) следующие:

Утром	38,2	36,4	37,7	36,1	37,9	37,8	–	–
Днем	39,5	38,7	37,8	38,6	39,2	39,1	38,9	39,2

Можно ли считать, что температура окружающей среды влияет на систематическую ошибку угломерного инструмента? Принять $\alpha = 0{,}05$.

19.235. Во время испытания радиодальномера проведено 16 независимых измерений дальности до контрольного объекта. Обработка результатов измерений дала следующие значения оценок ошибки радиодальномера: $\overline{x}_1 = -0{,}03$ км, $s_1^2 = 0{,}0324$ км^2. После юстировки устройства произведено еще 18 независимых измерений и получены такие значения оценок: $\overline{x}_2 = 0{,}05$ км, $s_2^2 = 0{,}0225$ км^2. Можно ли считать, что юстировка не повлияла на систематическую ошибку радиодальномера? Принять $\alpha = 0{,}10$.

19.236. При исследовании стабилизатора напряжения самолета на стенде проведено 9 независимых испытаний и получена оценка дисперсии выходного напряжения, равная $0{,}08$ В^2. В полете проведено еще 15 испытаний, в результате которых оценка дисперсии выходного напряжения оказалась равной $0{,}13$ В^2. Есть ли основания полагать, что факторы, воздействующие на стабилизатор в полете, оказывают существенное влияние на его точность? Принять $\alpha = 0{,}10$.

19.237 (сравнение дисперсий нескольких генеральных совокупностей). В таблице 4.3 приведены результаты измерений производительности 6 агрегатов и оценки дисперсий s_i^2, $i = 1, 2, \ldots, 6$, этих измерений. Используя эти данные, проверить гипотезу о равенстве дисперсий σ_i^2. Принять $\alpha = 0{,}10$.

Таблица 4.3

№ измерения	Агрегаты					
	1	2	3	4	5	6
1	14,0	14,1	14,0	14,5	12,5	14,0
2	14,5	10,1	12,3	14,2	12,3	14,0
3	13,7	14,7	12,8	15,0	11,5	13,5
4	12,7	13,7	11,0	14,7	12,9	14,7
5	14,1	14,0	13,1	13,5	12,8	13,6
s_i^2	0,46	3,37	1,22	0,33	0,31	0,22

◁ Для проверки гипотезы $H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_6^2$ воспользуемся критерием Бартлетта (таблица 4.1, нижняя строка). Предварительно вычислим

$$s^2 = \frac{4 \cdot (0{,}46 + 3{,}37 + 1{,}22 + 0{,}33 + 0{,}31 + 0{,}22)}{4 \cdot 6} \approx 0{,}98,$$

$$c \approx 1 + \frac{1}{3 \cdot (6-1)} \left[6 \cdot \frac{1}{4} - \frac{1}{6 \cdot 4} \right] \approx 1{,}097.$$

Выборочное значение $z_в$ статистики критерия равно:

$$z_в \approx \frac{1}{1{,}097}\,[6 \cdot 4 \ln 0{,}98 - 4\,(\ln 0{,}46 + \ln 3{,}37 + \ln 1{,}22 + \\ + \ln 0{,}33 + \ln 0{,}31 + \ln 0{,}22)] \approx 11{,}07.$$

По таблице П5 находим $\chi^2_{0,90}(5) = 9{,}24$. Так как выборочное значение статистики критерия превосходит квантиль $\chi^2_{0,90}(5)$, гипотеза о равенстве дисперсий отклоняется. $\rhd$

19.238. Проверить гипотезу о равенстве дисперсий трех совокупностей, используя следующие результаты наблюдений:

№ выборки	№ наблюдения					$\sum x_i$	$\sum x_i^2$
	1	2	3	4	5		
1	6	5	12	9	10	42	386
2	14	11	5	6	–	36	378
3	12	4	7	–	–	23	209

Принять $\alpha = 0{,}05$.

19.239. Проверить гипотезу о равенстве дисперсий по приведенным ниже данным.

а) Выборочные дисперсии, вычисленные по результатам трех серий независимых измерений концентрации (в процентах):

$$s_1^2 = 11{,}2, \quad s_2^2 = 15{,}8, \quad s_3^2 = 10{,}1, \quad n_1 = 10, \quad n_2 = 7, \quad n_3 = 12;$$

принять $\alpha = 0{,}10$.

б) Выборочные дисперсии величины контролируемого размера (мкм), полученные по результатам выборок из продукции четырех станков, производящих одни и те же детали:

$$s_1^2 = 1{,}16, \quad s_2^2 = 4{,}33, \quad s_3^2 = 2{,}17, \quad s_4^2 = 6{,}41,$$
$$n_1 = 5, \quad n_2 = 9, \quad n_3 = 6, \quad n_4 = 8;$$

принять $\alpha = 0{,}05$.

Пусть проверяется гипотеза $H_0:\ \theta = \theta_0$, а $V_к$ — критическая область критерия с заданным уровнем значимости α. *Функцией мощности* $\mathcal{M}\,(V_к,\,\theta)$ критерия называется вероятность отклонения гипотезы H_0 как функция параметра θ, т.е.

$$\mathcal{M}\,(V_к,\,\theta) = \mathrm{P}\,[Z \in V_к]. \qquad (3)$$

Вероятность отклонения гипотезы H_0 при конкретном значении параметра θ называется *мощностью критерия*. Очевидно, $\mathcal{M}(V_к, \theta_0) = \alpha$. Если альтернативная гипотеза H_1 простая, причем $H_1: \theta = \theta_1$, то мощность критерия равна $1-\beta$, т.е.

$$\mathcal{M}(V_к, \theta_1) = 1 - \beta.$$

Обычно строят *график функции мощности*, вычисляя мощность критерия при нескольких значениях параметра θ.

Пример 5. Построить графики функции мощности критерия значимости в примере 2, если используется:

а) выборка объема $n = 25$;

б) выборка объема $n = 100$.

◁ а) Вычислим мощность критерия при нескольких значениях параметра m. Используя данные примера 2 ($V_к: \overline{X} < 9{,}44$, $\sigma^2 = 4$) и формулу (3), при $n = 25$ имеем

$$\mathcal{M}(V_к, m) = \mathrm{P}\left[\overline{X} < 9{,}44\,\big|_m\right] = \Phi\left(\frac{9{,}44 - m}{\sqrt{4/25}}\right).$$

Используя таблицу П1, получим следующие значения мощности критерия:

m	$z = \dfrac{9{,}44 - m}{\sqrt{4/25}}$	$\mathcal{M}(V_к, m)$
8,50	2,350	0,991
8,75	1,725	0,958
9,00	1,100	0,864
9,25	0,475	0,682
9,50	−0,150	0,440
9,75	−0,775	0,219
10,00	−1,400	0,081

б) Аналогично, при $n = 100$ получим такие значения мощности критерия:

Графики функций мощности приведены на рис. 36. ▷

19.240. По данным задачи 19.207 вычислить мощность критерия для значений $m = 40{,}0$; $40{,}1$; $40{,}2$; …; $40{,}5$ при условии, что решение принимается по выборке объема $n = 36$ и по выборке объема $n = 100$. Используя эти результаты, построить график функции мощности.

19.241. Найти мощность критерия, используемого в задаче 19.212, если альтернативная гипотеза предполагает среднее время безотказной работы 950 часов. Какого объема должна быть выборка, чтобы в этих условиях вероятность ошибки второго рода не превышала 0,1?

m	$z = \dfrac{9{,}44 - m}{\sqrt{4/100}}$	$\mathcal{M}(V_к, m)$
8,50	4,70	1,000
8,75	3,45	1,000
9,00	2,20	0,986
9,25	0,95	0,829
9,50	−0,30	0,302
9,75	−1,55	0,061
10,00	−2,80	0,003

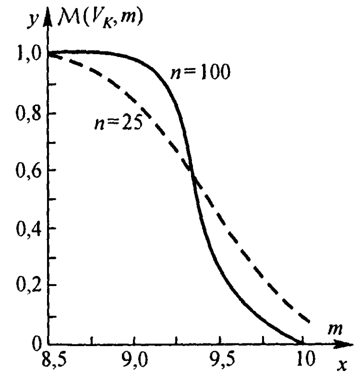


Рис. 36

19.242. Вычислить мощность критерия в задаче 19.214, считая, что дисперсия известна и равна $\sigma^2 = 1$, для нескольких альтернативных гипотез $H_1 :\ d = d_1$ при $d_1 = 10{,}1;\ 10{,}2;\ \ldots;\ 10{,}5$, если: а) объем выборки $n = 100$; б) объем выборки $n = 16$. Используя полученные результаты, построить график функции мощности критерия и найти вероятности ошибок второго рода при альтернативной гипотезе H_1: $d_1 = 10{,}25$.

19.243. Вычислить мощность критерия в задаче 19.215 а), если альтернативная гипотеза предполагает номинал партии равным 9,5 кОм. Какой объем выборки необходимо взять, чтобы ошибка второго рода не превосходила 0,01?

19.244*. Построить графики функции мощности критерия для задачи 19.215 а) при $n = 36$ и $n = 100$.

2. Проверка гипотез о параметре p биномиального распределения. При статистическом анализе данных, связанных с повторными независимыми испытаниями (схемой Бернулли), обычно рассматривают два вида задач: сравнение вероятности «успеха» p в одном испытании с заданным значением p_0 и сравнение вероятностей «успеха» в двух сериях испытаний.

В первом случае проверяется гипотеза $H_0 :\ p = p_0$. Пусть в n испытаниях по схеме Бернулли «успех» произошел x раз. В качестве статистики критерия выбирают относительную частоту $h = x/n$. При больших значениях n ($n > 50$) и при выполнении условий $nh > 5$,

$n(1-h)>5$ распределение случайной величины h с достаточной для практических расчетов точностью аппроксимируется нормальным распределением $N(p, \sqrt{p(1-p)/n})$. Отсюда следует, что если гипотеза H_0 верна, то статистика

$$Z = \frac{h - p_0}{\sqrt{p_0(1-p_0)/n}} \tag{4}$$

имеет распределение, близкое к нормальному распределению $N(0, 1)$. Критическая область критерия при уровне значимости α определяется неравенствами

$z_в > u_{1-\alpha}$ при альтернативной гипотезе $H_1^{(1)}: \ p > p_0$,

$z_в < u_\alpha$ при альтернативной гипотезе $H_1^{(2)}: \ p < p_0$,

$|z_в| > u_{1-\alpha/2}$ при альтернативной гипотезе $H_1^{(3)}: \ p \neq p_0$.

Для проверки гипотезы $H_0: \ p = p_0$ также можно использовать доверительные интервалы для параметра p (§ 3, п. 2). При этом гипотеза H_0 принимается на уровне значимости α, если соответствующий односторонний или двусторонний доверительный интервал накрывает значение p_0; в противном случае гипотеза H_0 отклоняется.

Пример 6. Предполагается, что большая партия деталей содержит 15 % брака. Для проверки из партии случайным образом отобрано 100 деталей, среди которых оказалось 10 бракованных. Считая, что число бракованных деталей в партии имеет биномиальное распределение, и используя двусторонний критерий при $\alpha = 0{,}05$, проверить предположение о том, что в партии содержится 15 % бракованных деталей.

◁ Проверяется гипотеза $H_0: \ p = 0{,}15$ при альтернативной гипотезе $H_1: \ p \neq 0{,}15$. Значение $h = \dfrac{10}{100} = 0{,}1$. Так как $n > 50$, $nh = 10$ и $n(1-h) = 9$, то для проверки гипотезы H_0 можно использовать статистику (4). Выборочное значение этой статистики

$$z_в = \frac{0{,}1 - 0{,}15}{\sqrt{\dfrac{0{,}15 \cdot 0{,}85}{100}}} \approx -1{,}4.$$

По таблице П1 находим $u_{0{,}975} = 1{,}96$. Значение $|z_в|$ лежит в области принятия гипотезы H_0, следовательно, предположение о том, что в партии содержится 15 % брака, согласуется с результатами наблюдений. Этот же результат получим, используя двусторонний доверительный интервал $(0{,}041; 0{,}159)$ для p при доверительной вероятности 0,95 (см. пример 4 § 3). Так как этот доверительный интервал накрывает значение $p = 0{,}15$, гипотеза H_0 принимается. ▷

19.245. Количество бракованных деталей в партии не должно превышать 5 %. В результате контроля 100 деталей из этой партии обнаружено 6 бракованных. Можно ли считать, что процент брака превосходит допустимый при $\alpha = 0{,}01$?

19.246. При 600 подбрасываниях игральной кости шестерка появилась 75 раз.

а) Можно ли утверждать, что кость симметрична и однородна? Принять $\alpha = 0{,}05$.

б) Верна ли гипотеза о том, что вероятность появления шестерки меньше, чем $1/6$, если $\alpha = 0{,}01$?

19.247. В урне содержатся неразличимые на ощупь черные и белые шары. Предполагается, что число черных шаров равно числу белых. Эта гипотеза принимается, если при извлечении 50 шаров (с возвращением) число черных шаров будет в пределах от 20 до 30.

а) Какова вероятность ошибки первого рода?

б) Найти вероятность ошибки второго рода, если альтернативная гипотеза утверждает, что вероятность появления черного шара равна $1/3$.

19.248. При исследовании 50 корпусов микросхем, случайным образом выбранных из большой партии этих изделий, оказалось, что шесть из них не имеют необходимой прочности. Согласуются ли эти данные с утверждением о том, что данная партия содержит более, чем 90 % прочных корпусов. Принять $\alpha = 0{,}05$.

19.249. Из суточной продукции цеха случайным образом отобрано и проверено 200 приборов. При этом 16 приборов признаны негодными к эксплуатации. Можно ли считать, что годная продукция цеха составляет 90 %, если $\alpha = 0{,}10$?

Для проверки гипотезы** H_0: $p_1 = p_2$ **о равенстве параметров двух биномиально распределенных совокупностей проводятся две серии испытаний. Пусть некоторое событие A в серии из n_1 испытаний появилось n_{11} раз, а в серии из n_2 испытаний — n_{21} раз. Проверяется гипотеза о равенстве вероятностей появления события A в обеих сериях испытаний. Представим результаты испытаний в обеих сериях в виде таблицы.

Серия	Событие		Сумма
	A	$\overline{A}$	
1	n_{11}	n_{12}	$n_{1\cdot}$
2	n_{21}	n_{22}	$n_{2\cdot}$
Сумма	$n_{\cdot 1}$	$n_{\cdot 2}$	$n_{\cdot\cdot} = n$

Обозначим $h_1 = \dfrac{n_{11}}{n_{1\cdot}}$, $h_2 = \dfrac{n_{21}}{n_{2\cdot}}$, $h = \dfrac{n_{11} + n_{21}}{n_{1\cdot} + n_{2\cdot}} = \dfrac{n_{\cdot 1}}{n}$.

При больших значениях n и при условии, что наименьшая из величин $\frac{n_{\cdot i} n_{j\cdot}}{n}$, $i, j = 1, 2$, будет больше 5, в качестве статистики критерия для проверки гипотезы $H_0\colon\ p_1 = p_2$ используют статистику

$$Z = \frac{h_1 - h_2}{\sqrt{\tilde{\sigma}^2_{h_1 - h_2}}}, \tag{5}$$

где $\tilde{\sigma}^2_{h_1 - h_2}$ — оценка дисперсии разности случайных величин h_1 и h_2, вычисляем по формуле

$$\tilde{\sigma}^2_{h_1 - h_2} = h\,(1 - h)\left(\frac{1}{n_{1\cdot}} + \frac{1}{n_{2\cdot}}\right). \tag{6}$$

Если гипотеза H_0 верна, то распределение статистики (5) близко к нормальному распределению $N(0, 1)$. Критическая область критерия при уровне значимости α определяется неравенствами

$z_в > u_{1-\alpha}$ при альтернативной гипотезе $H_1^{(1)}\colon\ p_1 > p_2$,

$z_в < u_\alpha$ при альтернативной гипотезе $H_1^{(2)}\colon\ p_1 < p_2$,

$z_в > u_{1-\alpha/2}$ при альтернативной гипотезе $H_1^{(3)}\colon\ p_1 \neq p_2$.

В случае, когда результаты наблюдений таковы, что условие $\frac{n_{\cdot i} n_{j\cdot}}{n} > 5$, $i, j = 1, 2$, не удовлетворяется, для проверки гипотезы H_0 следует использовать критерий χ^2 (см. § 6, п. 3).

Пример 7. Ниже приведены результаты выборочного обследования двух партий изделий:

№ партии	Число изделий		Сумма
	Бракованные	Небракованные	
1	8	92	100
2	13	287	300
Сумма	21	379	400

Можно ли считать, что доля брака в обеих партиях одна и та же, если уровень значимости $\alpha = 0{,}05$?

$\lhd$ Проверяется гипотеза $H_0\colon\ p_1 = p_2$ при альтернативной гипотезе $H_1\colon\ p_1 \neq p_2$. Условие $\frac{n_{\cdot i} n_{j\cdot}}{n} > 5$, $i, j = 1, 2$, выполняется, следовательно, для проверки гипотезы H_0 можно использовать статистику (5). По результатам обследования определим

$$h_1 = \frac{8}{100} = 0{,}08, \quad h_2 = \frac{13}{300} \approx 0{,}043, \quad h = \frac{21}{400} \approx 0{,}052.$$

Предварительно по формуле (6) найдем оценку дисперсии:

$$\tilde{\sigma}^2_{h_1-h_2} = 0{,}052 \cdot (1 - 0{,}052) \cdot \left(\frac{1}{100} + \frac{1}{300}\right) \approx 6{,}57 \cdot 10^{-4}.$$

Выборочное значение $z_в$ статистики критерия по формуле (5)

$$\frac{0{,}08 - 0{,}043}{\sqrt{6{,}57 \cdot 10^{-4}}} \approx 1{,}44.$$

Так как $u_{0,975} = 1{,}96$, то выборочное значение статистики критерия принадлежит области принятия гипотезы H_0; поэтому следует считать, что доля брака в обеих партиях одна и та же. ▷

19.250. Два пресса штампуют детали одного наименования. Из партии деталей, изготовленных первым прессом, проверено 1000 деталей, из которых 25 оказались негодными. Из 800 деталей, изготовленных вторым прессом, негодными оказались 36 деталей. Согласуются ли эти результаты с предположением о равенстве доли брака в продукции двух прессов при $\alpha = 0{,}10$?

19.251. Предполагается, что применение новой технологии в производстве микросхем приведет к увеличению выхода годной продукции. Результаты контроля двух партий продукции, изготовленных по старой и новой технологии, приведены ниже:

Изделия	Технология	
	Старая	Новая
Годные	140	185
Негодные	10	15
Всего	150	200

Подтверждают ли эти результаты предположение об увеличении выхода годной продукции? Принять $\alpha = 0{,}01$.

16.252. Для изучения эффективности профилактического лекарства против аллергии обследовались две группы людей, предрасположенных к этому заболеванию. Результаты обследования следующие:

Принимали лекарство		Не принимали лекарства	
Заболели	Не заболели	Заболели	Не заболели
3	172	32	168

Показывают ли эти результаты эффективность лекарства, если уровень значимости $\alpha = 0{,}05$?

19.253. В 105 опытах событие A произошло 42 раза. Повторная серия опытов состояла из 195 опытов, причем событие произошло 65 раз. Можно ли считать, что вероятность события A в обеих сериях одна и та же, если исходы опытов независимы. Принять $\alpha = 0{,}01$.

3. Проверка гипотез о коэффициенте корреляции ρ. Пусть r — выборочный коэффициент корреляции, вычисленный по выборке объема n из генеральной совокупности, имеющей двумерное нормальное распределение.

Для проверки гипотезы $H_0: \ \rho = \rho_0$, где ρ_0 — заданное значение, используют статистику

$$Z = \frac{\operatorname{Arth} r - \operatorname{Arth} \rho_0}{1/\sqrt{n-3}}, \tag{7}$$

где $\operatorname{Arth} r = \dfrac{1}{2} \ln \dfrac{1+r}{1-r}$.

Если гипотеза H_0 верна, то статистика (7) имеет распределение, близкое к нормальному $N(0, 1)$.

Критическая область критерия при уровне значимости α определяется неравенствами

$z_в > u_{1-\alpha}$ при альтернативной гипотезе $H_1^{(1)}: \ \rho > \rho_0$,

$z_в < u_\alpha$ при альтернативной гипотезе $H_1^{(2)}: \ \rho < \rho_0$,

$|z_в| > u_{1-\alpha/2}$ при альтернативной гипотезе $H_1^{(3)}: \ \rho \neq \rho_0$.

В случае, когда нужно определить значимость выборочного значения коэффициента корреляции r, т. е. проверить гипотезу $H_0: \rho = 0$, можно использовать другой критерий, статистикой которого является r. На уровне значимости α критическая область этого критерия определяется неравенствами

$$r > \frac{t_{1-\alpha}(n-2)}{\sqrt{n-2+t_{1-\alpha}^2(n-2)}}$$

при альтернативной гипотезе $H_1^{(1)}$: $\rho > 0$;

$$r < \frac{t_{\alpha}(n-2)}{\sqrt{n-2+t_{\alpha}^2(n-2)}}$$

при альтернативной гипотезе $H_1^{(2)}$: $\rho < 0$;

$$|r| > \frac{t_{1-\alpha/2}(n-2)}{\sqrt{n-2+t_{1-\alpha/2}^2(n-2)}}$$

при альтернативной гипотезе $H_1^{(3)}: \ \rho \neq 0$.

Пример 8. Из генеральной совокупности, имеющей двумерное нормальное распределение, получена выборка объема $n = 67$. Выборочный коэффициент корреляции оказался равным $r = -0{,}159$. Можно ли считать, что наблюдаемые переменные отрицательно коррелированы, если уровень значимости $\alpha = 0{,}05$?

$\lhd$ Проверим гипотезу $H_0: \ \rho = 0$ при альтернативной гипотезе $H_1: \ \rho < 0$. Вычислим выборочное значение статистики критерия (7). Значение $\operatorname{Arth} r$ находим по таблице П8. Имеем

$$z_в = \frac{\operatorname{Arth}(-0{,}159) - \operatorname{Arth} 0}{1/\sqrt{67-3}} \approx -1{,}28.$$

Так как $u_{0,05} = -1{,}645$, то выборочное значение статистики критерия принадлежит области принятия гипотезы H_0, следовательно, наблюдаемые переменные не коррелированы.

Такой же результат получим, воспользовавшись критерием, статистикой которого является r. Найдем границу критической области при альтернативной гипотезе $H_1: \ \rho < 0$. Определим квантили $t_{0,05}(65) = -t_{0,95}(65) \approx -1{,}67$ (таблица П6). Вычислим границу критической области:

$$\frac{t_\alpha(n-2)}{\sqrt{n-2+t_\alpha^2(n-2)}} = \frac{-1{,}67}{\sqrt{67-2+(-1{,}67)^2}} \approx -0{,}203.$$

Так как выборочное значение $r = -0{,}159$ статистики принадлежит области принятия гипотезы H_0, то гипотеза H_0 принимается; следует считать, что наблюдаемые переменные не коррелированы. $\rhd$

Пусть r_1 и r_2 — выборочные коэффициенты корреляции, вычисленные по выборкам объема n_1 и n_2 из генеральных совокупностей, имеющих двумерное нормальное распределение. Для проверки гипотезы $H_0: \ \rho_1 = \rho_2$ используют статистику

$$Z = \frac{\operatorname{Arth} r_1 - \operatorname{Arth} r_2}{\sqrt{1/(n_1-3) + 1/(n_2-3)}}. \tag{8}$$

При условии, что гипотеза H_0 верна, статистика (8) имеет распределение, близкое к нормальному распределению $N(0, 1)$. Критическая область критерия при уровне значимости α определяется неравенствами

$z_в > u_{1-\alpha}$ при альтернативной гипотезе $H_1^{(1)}: \ \rho_1 > \rho_2$,

$z_в < u_\alpha$ при альтернативной гипотезе $H_1^{(2)}: \ \rho_1 < \rho_2$,

$|z_в| > u_{1-\alpha/2}$ при альтернативной гипотезе $H_1^{(3)}: \ \rho_1 \neq \rho_2$.

Пример 9. Сравнить коэффициенты корреляции двух нормально распределенных генеральных совокупностей по следующим выборочным данным: $r_1 = 0{,}77$, $n_1 = 28$, $r_2 = 0{,}604$, $n_2 = 33$. Принять $\alpha = 0{,}10$.

◁ Имеем H_0: $\rho_1 = \rho_2$; H_1: $\rho_1 < \rho_2$. Вычислим выборочное значение статистики критерия (8):

$$\frac{\operatorname{Arth} 0{,}77 - \operatorname{Arth} 0{,}604}{\sqrt{\dfrac{1}{28-3} + \dfrac{1}{33-3}}} \approx 1{,}85.$$

Так как $u_{0,95} \approx 1{,}645$, то выборочное значение статистики критерия принадлежит критической области; коэффициенты корреляций генеральных совокупностей следует считать различными. ▷

В задачах 19.254–19.260 предполагается, что выборки получены из генеральных совокупностей, имеющих двумерное нормальное распределение.

19.254. Выборочный коэффициент корреляции r, вычисленный по выборке объема $n = 39$, равен 0,25. Проверить значимость этого результата при альтернативных гипотезах: а) H_1: $\rho \neq 0$, б) H_1: $\rho > 0$. Принять $\alpha = 0{,}05$.

19.255. Проверить значимость коэффициента корреляции по следующим данным:

а) $r = -0{,}41$, $n = 52$, $\alpha = 0{,}1$, альтернативная гипотеза H_1: $\rho < 0$;

б) $r = 0{,}15$, $n = 39$, $\alpha = 0{,}01$, альтернативная гипотеза H_1: $\rho \neq 0$;

в) $r = -0{,}32$, $n = 103$, $\alpha = 0{,}05$, альтернативная гипотеза H_1: $\rho \neq 0$.

19.256. По выборке объема $n = 28$ вычислен коэффициент корреляции $r = 0{,}88$. Согласуются ли следующие гипотезы относительно коэффициента корреляции генеральной совокупности ρ с результатами наблюдений: а) $\rho > 0{,}90$; б) $\rho < 0{,}6$; в) $\rho \neq 0{,}96$? Принять $\alpha = 0{,}05$.

19.257. По двум выборкам объемов $n_1 = 28$ и $n_2 = 39$ для наблюдений над двумя определенными переменными некоторого процесса вычислены оценки коэффициента корреляции, равные $r_1 = 0{,}71$ и $r_2 = 0{,}85$ соответственно.

а) Можно ли считать, что оценки коэффициентов корреляции, вычисленные по двум выборкам, действительно различны?

б) Для каких значений r_2 можно считать, что разница оценок коэффициентов корреляции r_1 и r_2 незначима? Принять $\alpha = 0{,}01$.

Решить задачу 19.257 для следующих данных:

19.258. $n_1 = 124$, $r_1 = -0{,}87$, $n_2 = 147$, $r_2 = -0{,}65$, $\alpha = 0{,}10$.

19.259. $n_1 = 12$, $r_1 = 0{,}42$, $n_2 = 19$, $r_2 = 0{,}36$, $\alpha = 0{,}05$.

19.260. $n_1 = 82$, $r_1 = 0{,}95$, $n_2 = 67$, $r_2 = 0{,}82$, $\alpha = 0{,}01$.

4. Определение наилучшей критической области для проверки простых гипотез. Очевидно, на множестве значений статистики критерия можно выбрать сколько угодно критических областей V_κ для заданного уровня значимости α, однако соответствующие им критерии будут иметь, вообще говоря, различные вероятности ошибок второго рода (см. задачу 19.209). *Наилучшей критической областью (НКО)* называют критическую область, которая при заданном уровне значимости α обеспечивает минимальную вероятность ошибки второго рода. Критерий, использующий НКО, имеет максимальную мощность.

При проверке простой гипотезы H_0 против простой альтернативы H_1 НКО определяется *леммой Неймана–Пирсона: НКО критерия заданного уровня значимости α состоит из точек выборочного пространства (выборок объема n), для которых удовлетворяется неравенство*

$$\frac{L(x_1, x_2, \ldots, x_n/H_0)}{L(x_1, x_2, \ldots, x_n/H_1)} < c_\alpha, \tag{9}$$

где c_α — константа, зависящая от заданного уровня значимости, x_1, $x_2, \ldots, x_n$ — элементы выборки, а $L(x_1, x_2, \ldots, x_n/H_i)$ — функция правдоподобия, вычисленная при условии, что верна гипотеза H_i, $i = 0, 1$.

Для рассмотренных в пп. 1–4 критериев значимости НКО размещаются на «хвостах» распределений статистик критериев.

Пример 10. Пусть случайная величина X имеет нормальное распределение $N(m, \sigma)$ с известной дисперсией σ^2.

а) Найти НКО для проверки гипотезы $H_0: m = m_0$ против простой альтернативной гипотезы $H_1: m = m_1$, где $m_1 > m_0$.

б) Найти функцию мощности критерия и вычислить ее значение при значениях $m_1 = 1$ и $m_1 = 5$, если $m_0 = 0$, объем выборки $n = 25$, дисперсия генеральной совокупности $\sigma^2 = 25$. Принять уровень значимости $\alpha = 0{,}05$.

$\triangleleft$ а) Запишем отношение функций правдоподобия:

$$\frac{L(x_1, x_2, \ldots, x_n/H_0)}{L(x_1, x_2, \ldots, x_n/H_1)} = \frac{\left(\dfrac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\dfrac{\sum(x_i - m_0)^2}{2\sigma^2}\right\}}{\left(\dfrac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\dfrac{\sum(x_i - m_1)^2}{2\sigma^2}\right\}} =$$

$$= a\exp\left\{-\frac{\overline{x}(m_1 - m_0)}{\sigma^2/n}\right\},$$

где $a = \exp\left\{-\dfrac{m_0^2 - m_1^2}{2\sigma^2/n}\right\}$, $\overline{x} = \dfrac{1}{n}\sum x_i$. По лемме Неймана–Пирсона НКО содержит только те точки выборочного пространства, для которых удовлетворяется неравенство (9):

$$a\exp\left\{-\frac{\overline{x}(m_1 - m_0)}{\sigma^2/n}\right\} < c_\alpha,$$

причем по условию задачи $m_1 - m_0 > 0$. Так как отношение правдоподобия является убывающей функцией аргумента $\overline{x}$, условие леммы удовлетворяется при $\overline{x} > \overline{x}_\kappa$, где граница критической области $\overline{x}_\kappa$ находится по заданному уровню значимости α из соотношения

$$\mathrm{P}\,[\overline{X} > \overline{x}_\kappa / H_0] = \alpha.$$

При условии, что справедлива гипотеза H_0, $\overline{X}$ имеет нормальное распределение $N(m_0, \sigma/\sqrt{n})$, следовательно,

$$\mathrm{P}\,[\overline{X} < \overline{x}_\kappa / H_0] = \Phi\left(\frac{\overline{x}_\kappa - m_0}{\sigma/\sqrt{n}}\right) = 1 - \alpha.$$

Отсюда следует, что

$$\frac{\overline{x}_\kappa - m_0}{\sigma/\sqrt{n}} = u_{1-\alpha}.$$

Таким образом, граница $\overline{x}_\kappa$ критической области равна $m_0 + u_{1-\alpha}\dfrac{\sigma}{\sqrt{n}}$, а НКО V_κ имеет вид $\overline{x} > m_0 + u_{1-\alpha}\dfrac{\sigma}{\sqrt{n}}$.

б) Найдем функцию мощности полученного критерия:

$$\mathcal{M}\,(V_\kappa, m) = \mathrm{P}\left[\overline{X} > \overline{x}_\kappa \Big|_m\right] = 1 - \Phi\left(\frac{\overline{x}_\kappa - m}{\sigma/\sqrt{n}}\right) = \\ = \Phi\left(u_\alpha + \sqrt{n}\,\frac{m - m_0}{\sigma}\right).$$

При значениях $m_1 = 1$, $m_0 = 0$, $\alpha = 0{,}05$, $n = 25$, $\sigma^2 = 25$ мощность критерия

$$\mathcal{M}\,(V_\kappa, 1) = \Phi\left(-1{,}645 + \sqrt{25}\cdot\frac{1-0}{\sqrt{25}}\right) = 0{,}259.$$

Если $m_1 = 5$, то мощность критерия

$$\mathcal{M}\,(V_\kappa, 5) = \Phi\left(-1{,}645 + \sqrt{25}\cdot\frac{5-0}{\sqrt{25}}\right) = 0{,}999. \ \triangleright$$

19.261*. Найти наилучшую критическую область для проверки гипотезы H_0 о распределении случайной величины X с плотностью распределения

$$f_X(x) = \begin{cases} 1/2, & |\,x\,| < 1, \\ 0, & |\,x\,| \geqslant 1 \end{cases}$$

против альтернативной гипотезы H_1, предполагающей нормальное распределение случайной величины X при $m = 0$ и $\sigma^2 = 1$, если для проверки гипотезы используются результаты одного наблюдения. Вычислить мощность полученного критерия. Принять $\alpha = 0{,}05$.

19.262. Решить задачу 19.261, поменяв ролями нулевую и альтернативную гипотезы.

19.263. Случайная величина X имеет показательное распределение $\mathrm{Ex}(\theta)$. Определить наилучшую критическую область для проверки гипотезы $H_0: \ \theta = 1$ против альтернативной гипотезы $H_1: \ \theta = \theta_1$, где $\theta_1 > 1$, по результатам одного наблюдения. Найти функцию мощности критерия и вычислить мощность критерия при $\theta_1 = 2$. Принять $\alpha = 0{,}1$.

19.264. Случайная величина имеет распределение Пуассона с параметром λ. Используя выборку наблюдений объема n, определить наилучшую критическую область для проверки гипотезы $H_0: \ \lambda = \lambda_0$ против альтернативной гипотезы $H_1: \ \lambda = \lambda_1$. Рассмотреть случаи: а) $\lambda_0 < \lambda_1$; б) $\lambda_0 > \lambda_1$.

19.265. Пусть случайная величина X — число «успехов» в независимых испытаниях, а p — вероятность «успеха» в каждом испытании. Определить наилучшую критическую область для проверки гипотезы $H_0: \ p = p_0$ против альтернативной гипотезы $H_1: \ p = p_1$.

19.266. Случайная величина X имеет нормальное распределение $N(0, \sigma)$. Определить наилучшую критическую область для проверки гипотезы $H_0: \ \sigma^2 = \sigma_0^2$ против альтернативной гипотезы $H_1: \sigma^2 = \sigma_1^2$, используя выборку n наблюдений этой случайной величины.

19.267. Случайная величина X имеет нормальное распределение $N(m, 1)$. Проверяется гипотеза $H_0: \ m = 0$ против альтернативной гипотезы $H_1: \ m = 1$ на уровне значимости $\alpha = 0{,}05$. Сколько наблюдений необходимо, чтобы мощность критерия была не меньше 0,90?

19.268. Пусть Θ — некоторое подмножество множества значений параметра θ. Критерий для проверки простой гипотезы $H_0: \ \theta = \theta_0$ против сложной альтернативной гипотезы $H_1: \ \theta \in \Theta$ называется *равномерно наиболее мощным*, если для любого $\theta \in \Theta$ он имеет наибольшую мощность. Является ли критерий, полученный в примере 10, равномерно наиболее мощным при проверке простой гипотезы $H_0: \ m = m_0$ против сложной альтернативной гипотезы H_1, если а) $H_1: \ m > m_0$; б) $H_1: \ m \neq m_0$?

Критерий для проверки простой гипотезы $H_0: \ \theta = \theta_0$ против сложной альтернативной гипотезы $H_1: \ \theta \in \Theta$ называется *несмещенным*, если для любого $\theta \in \Theta$ функция мощности этого критерия удовлетворяет условию: $\mathcal{M}_0(V_\kappa, \theta) \geqslant \alpha$, где α — уровень значимости критерия.

19.269. Показать, что критерий, полученный в примере 10 для проверки простой гипотезы H_0: $m = m_0$ против сложной гипотезы H_1: $m > m_0$, является несмещенным, а при сложной альтернативе $H_1: m \neq m_0$ — смещенным.

19.270. Какой несмещенный критерий можно предложить для проверки гипотезы $H_0 : m = m_0$ против альтернативной гипотезы $H_1 : m \neq m_0$?

§ 5. Однофакторный дисперсионный анализ

Пусть результаты наблюдений составляют l независимых выборок (групп), полученных из l нормально распределенных генеральных совокупностей, которые имеют, вообще говоря, различные средние m_1, $m_2, \ldots, m_l$ и равные дисперсии σ^2. Проверяется гипотеза о равенстве средних $H_0 : m_1 = m_2 = \cdots = m_l$. На практике такая задача возникает при исследовании влияния, которое оказывает изменение некоторого фактора на измеряемую величину. Например, если измерения проводятся на l различных приборах, то можно исследовать влияние фактора «прибор» на результаты измерений. В данном случае нас интересует вопрос, имеют ли различные приборы одну и ту же систематическую ошибку (гипотеза H_0). При $l = 2$ для проверки гипотезы H_0 используются известные критерии значимости (см. § 4, таблица 4.2). Если $l > 2$, то для проверки гипотезы о равенстве l средних применяют *однофакторный дисперсионный анализ*, суть которого состоит в следующем.

Пусть x_{ik} обозначает i-й элемент k-й выборки, $i = 1, 2, \ldots, n_k$; $k = 1, 2, \ldots, l$; $\overline{x}_k$ — выборочное среднее k-й выборки, т.е.

$$\overline{x}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} x_{ik} = \frac{1}{n_k} x_{\cdot k};$$

$\overline{x}$ — общее выборочное среднее, т.е.

$$\overline{x} = \frac{1}{n} \sum_{k=1}^{l} \sum_{i=1}^{n_k} x_{ik} = \frac{1}{n} x_{\cdot\cdot},$$

где n — общее число наблюдений, $n = \sum\limits_{k=1}^{l} n_k$.

Общая сумма квадратов отклонений наблюдений от общего среднего $\overline{x}$ может быть представлена так:

$$\sum_{k=1}^{l} \sum_{i=1}^{n_k} (x_{ik} - \overline{x})^2 = \sum_{k=1}^{l} n_k (\overline{x}_k - \overline{x})^2 + \sum_{k=1}^{l} \sum_{i=1}^{n_k} (x_{ik} - \overline{x}_k)^2. \tag{1}$$

Это — основное тождество дисперсионного анализа. Запишем его в виде

$$Q = Q_1 + Q_2, \tag{2}$$

где Q — общая сумма квадратов отклонений наблюдений от общего среднего, Q_1 — сумма квадратов отклонений выборочных средних $\overline{x}_k$ от общего среднего $\overline{x}$ (между группами), Q_2 — сумма квадратов отклонений наблюдений от выборочных средних групп (внутри групп).

Тождество (1) легко проверяется, если воспользоваться очевидным равенством

$$(x_{ik} - \overline{x}) = [(\overline{x}_k - \overline{x}) + (x_{ik} - \overline{x}_k)]$$

возвести обе его части в квадрат, просуммировать по i и k и учесть, что

$$\sum_{k=1}^{l} \sum_{i=1}^{n_k} (x_{ik} - \overline{x}_k)(\overline{x}_k - \overline{x}) = 0$$

в силу определения средних $\overline{x}_k$ и $\overline{x}$.

Если верна гипотеза $H_0: m_1 = m_2 = \dots = m_l$, то (доказательство см., например, [10, с. 284–287]) статистики Q_1/σ^2 и Q_2/σ^2 независимы и имеют распределение χ^2 с $l-1$ и $n-l$ степенями свободы. Следовательно, статистики $S_1^2 = \dfrac{Q_1}{l-1}$ и $S_2^2 = \dfrac{Q_2}{n-l}$ являются несмещенными оценками неизвестной дисперсии σ^2. Оценка S_1^2 характеризует рассеяние групповых средних, а оценка S_2^2 — рассеяние внутри групп, которое обусловлено случайными вариациями результатов наблюдений. Значительно превышение величины S_1^2 над значением величины S_2^2 можно объяснить различием средних в группах. Отношение этих оценок имеет распределение Фишера с $l-1$ и $n-l$ степенями свободы, т.е.

$$\frac{S_1^2}{S_2^2} = \frac{Q_1/(l-1)}{Q_2/(n-l)} = F\,(l-1,\ n-l). \tag{3}$$

Статистика (3) используется для проверки гипотезы $H_0: m_1 = m_2 = \dots$ $\dots = m_l = m$. Гипотеза H_0 не противоречит результатам наблюдений, если выборочное значение $F_в$ статистики (3) меньше квантили $F_{1-\alpha}(l-1,\ n-l)$, т. е. если $F_в < F_{1-\alpha}(l-1,\ n-l)$. В этом случае $\overline{x}$ и $\dfrac{Q_2}{n-l}$ являются несмещенными оценками параметров m и σ^2. Если $F_в \geqslant F_{1-\alpha}(l-1,\ n-l)$, то гипотеза H_0 отклоняется и следует считать, что среди средних $m_1, m_2, \dots, m_l$ имеется хотя бы два не равных друг другу.

Пример 1. Три группы водителей обучались по различным методикам. После окончания срока обучения был произведен тестовый контроль над случайно отобранными водителями из каждой группы. Получены следующие результаты:

На уровне значимости $\alpha = 0{,}05$ проверить гипотезу об отсутствии влияния различных методик обучения на результаты тестового контроля водителей. Предполагается, что выборки получены из независимых нормально распределенных совокупностей с одной и той же дисперсией.

Номер группы, k	Число ошибок, допущенных водителями, x_{ik}							Сумма $x_{\cdot k}$	Число контролируемых водителей, n_k
1	1	3	2	1	0	2	1	10	7
2	2	3	2	1	4	–	–	12	5
3	4	5	3	–	–	–	–	12	3

◁ Очевидно, задача заключается в проверке гипотезы $H_0 : m_1 = m_2 = m_3$, где m_k — математическое ожидание числа ошибок для водителей k-й группы. В нашем случае $l = 3$, $n = 15$.

Вычисления удобно проводить в такой последовательности:

$$x_{\cdot\cdot} = \sum_{k=1}^{l} \sum_{i=1}^{n_k} x_{ik} = 10 + 12 + 12 = 34,$$

$$\sum_{k=1}^{l} \sum_{i=1}^{n_k} x_{ik}^2 = 104.$$

Далее из (1) и (2) получаем

$$Q = \sum_{k=1}^{l} \sum_{i=1}^{n_k} x_{ik}^2 - \frac{1}{n} x_{\cdot\cdot}^2 = 104 - \frac{1}{15} \cdot 34^2 \approx 26{,}93,$$

$$Q_1 = \sum_{k=1}^{l} \frac{1}{n_k} x_{\cdot k}^2 - \frac{1}{n} x_{\cdot\cdot}^2 = 91{,}086 - \frac{1}{15} \cdot 34^2 \approx 14{,}02,$$

$$Q_2 = Q - Q_1 = 26{,}93 - 14{,}02 = 12{,}91.$$

Вычисляем выборочное значение статистики (3):

$$F_в = \frac{Q_1/(l-1)}{Q_2/(n-l)} = \frac{14{,}02/2}{12{,}91/12} \approx 6{,}52.$$

Из таблицы П7 находим $F_{0,95}(2{,}12) = 3{,}89$. Так как $F_в = 6{,}52 > 3{,}89$, то гипотеза H_0 о равенстве средних отклоняется: исследуемые методики обучения водителей дают значимо различные результаты тестового контроля. ▷

Линейные контрасты. Если гипотеза H_0 о равенстве средних отклоняется, то требуется определить, какие именно группы имеют значимое различие средних. Для этих целей используется *метод линейных* *контрастов.* Линейный контраст Lk определяется как линейная комбинация:

$$\mathrm{Lk} = \sum_{k=1}^{l} c_k m_k,$$

где c_k, $k = 1, 2, \ldots, l$,— константы, однозначно определяемые из формулировки проверяемых гипотез, причем $\sum\limits_{k=1}^{l} c_k = 0$. Оценка Lk равна $\widetilde{\mathrm{Lk}} = \sum\limits_{k=1}^{l} c_k \overline{x}_k$, а оценка дисперсии Lk равна

$$s_{\mathrm{Lk}}^2 = \mathrm{D}\,[\mathrm{Lk}] = \tilde{\sigma}^2 \sum_{k=1}^{l} \frac{c_k^2}{n_k} = \frac{Q_2}{n-l} \sum_{k=1}^{l} \frac{c_k^2}{n_k}.$$

Границы доверительного интервала для Lk имеют вид

$$\widetilde{\mathrm{Lk}} \pm s_{\mathrm{Lk}} \sqrt{(l-1) \cdot F_{1-\alpha}(l-1,\ n-1)}. \tag{4}$$

Пример 2. В условиях примера 1 при двусторонних альтернативных гипотезах проверить гипотезы $H_0^{(1)}:\ m_1 = m_2$, $H_0^{(2)}:\ m_1 = m_3$; $H_0^{(3)}:\ m_2 = m_3$; $H_0^{(4)}:\ \dfrac{1}{2}(m_1 + m_3) = m_2$.

◁ В соответствии с проверяемыми гипотезами $H_0^{(i)}$, $i = 1, 2, 3, 4$, определяем линейные контрасты

$$\begin{array}{llll}
\mathrm{Lk}_1 = m_1 - m_2; & c_1 = 1, & c_2 = -1, & c_3 = 0; \\
\mathrm{Lk}_2 = m_1 - m_3; & c_1 = 1, & c_2 = 0, & c_3 = -1; \\
\mathrm{Lk}_3 = m_2 - m_3; & c_1 = 0, & c_2 = 1, & c_3 = -1; \\
\mathrm{Lk}_4 = \dfrac{1}{2}(m_1 + m_2) - m_3; & c_1 = \dfrac{1}{2}, & c_2 = \dfrac{1}{2}, & c_3 = -1.
\end{array}$$

Найдем границы доверительных интервалов для линейных контрастов Lk_i, $i = 1, 2, 3, 4$.

Предварительно вычислим оценки линейных контрастов и их дисперсий. Выборочные средние $\overline{x}_1 = 1{,}43$, $\overline{x}_2 = 2{,}4$, $\overline{x}_3 = 4$. Оценка дисперсии

$$\tilde{\sigma}^2 = \frac{Q_2}{n-l} = \frac{12{,}91}{15-3} \approx 1{,}08.$$

Оценки контрастов и их дисперсий

$$\widetilde{\mathrm{Lk}}_1 = 1{,}43 - 2{,}4 = -0{,}97, \quad s^2_{\mathrm{Lk}_1} = 1{,}08\left(\frac{1}{7}+\frac{1}{5}\right) \approx 0{,}37;$$

$$\widetilde{\mathrm{Lk}}_2 = 1{,}43 - 4 = -2{,}57, \quad s^2_{\mathrm{Lk}_2} = 1{,}08\left(\frac{1}{7}+\frac{1}{3}\right) \approx 0{,}51;$$

$$\widetilde{\mathrm{Lk}}_3 = 2{,}4 - 4 = -1{,}60, \quad s^2_{\mathrm{Lk}_3} = 1{,}08\left(\frac{1}{5}+\frac{1}{3}\right) \approx 0{,}58;$$

$$\widetilde{\mathrm{Lk}}_4 = \frac{1}{2}(1{,}43+2{,}4) - 4 = -2{,}08,$$

$$s^2_{\mathrm{Lk}_4} = 1{,}08\left(\frac{(1/2)^2}{7}+\frac{(1/2)^2}{5}+\frac{1}{3}\right) \approx 0{,}45.$$

При $\alpha = 0{,}05$ по таблице П7 находим $F_{1-\alpha}(l-1,\, n-l) = F_{0{,}95}(2{,}12) \approx$ $\approx 3{,}89$. Чтобы определить доверительные интервалы для линейных контрастов, предварительно вычислим

$$\sqrt{(l-1)\,F_{1-\alpha}(l-1,\; n-l)} = \sqrt{(3-1)\cdot 3{,}89} \approx 2{,}79.$$

Таким образом, доверительные границы для контрастов Lk_i, $i = 1, 2, 3, 4$, по формуле (4) равны соответственно $-0{,}97 \pm 1{,}67$; $-2{,}57 \pm 2{,}00$; $-1{,}60 \pm 2{,}12$; $-2{,}08 \pm 1{,}87$. Так как нулевое значение накрывается доверительными интервалами для Lk_1 и Lk_3, то гипотезы $H_0^{(1)}$ и $H_0^{(3)}$ принимаются, гипотезы $H_0^{(2)}$ и $H_0^{(4)}$ отклоняются. Таким образом, значимо различны средние первой и третьей группы, а также среднее арифметическое средних для первых двух групп и среднее третьей группы. ▷

В задачах 19.271–19.278 предполагается, что выборки получены из нормально распределенных генеральных совокупностей с равными дисперсиями. В каждой задаче требуется проверить гипотезу H_0 о равенстве средних. Если гипотеза H_0 принимается, то найти несмещенные оценки среднего и дисперсии. В случае, если гипотеза H_0 отклоняется, провести попарное сравнение средних, используя метод линейных контрастов.

19.271.

выборка 1	6	5	12	9	10
выборка 2	14	11	5	6	–
выборка 3	12	4	7	–	–

$$\alpha = 0{,}05.$$

19.272.

выборка 1	4	2	3	4	5	3
выборка 2	6	5	4	7	6	8
выборка 3	8	9	10	7	8	6

$$\alpha = 0{,}10.$$

19.273.

выборка 1	8	11	8	9	-	-
выборка 2	9	10	7	11	8	10
выборка 3	16	9	12	14	-	-
выборка 4	9	8	-	-	-	-

$$\alpha = 0{,}05.$$

19.274. В трех магазинах, продающих товары одного вида, данные товарооборота за 8 месяцев работы (в тыс. руб.) составили следующую сводку:

Магазин	Месяц								$\sum_{i=1}^{8} x_{ik}$
	1	2	3	4	5	6	7	8	
I	19	23	26	18	20	20	18	35	179
II	20	20	32	27	40	24	22	18	203
III	16	15	18	26	19	17	19	18	148

$$\sum_{k=1}^{3}\sum_{i=1}^{8} x_{ik}^2 = 12\,592. \text{ Принять } \alpha = 0{,}10.$$

19.275. Ниже приводятся данные о содержании иммуноглобулина Ig A в сыворотке крови (в мг %) у больных пяти возрастных групп:

Возрастная группа	Содержание Ig A (мг %)											Сумма
1	83	85	–	–	–	–	–	–	–	–	–	168
2	84	85	85	86	86	87	–	–	–	–	–	513
3	86	87	87	87	88	88	88	88	88	89	90	966
4	89	90	90	91	–	–	–	–	–	–	–	360
5	90	92	–	–	–	–	–	–	–	–	–	182

$$\sum\sum x_{ik} = 2189, \quad \sum\sum x_{ik}^2 = 191\,791. \text{ Принять } \alpha = 0{,}01.$$

19.276. На химическом заводе разработаны два новых варианта технологического процесса. Чтобы оценить, как изменится дневная производительность при переходе на работу по новым вариантам технологического процесса, завод в течение 10 дней работает по каждому варианту, включая существующий вариант. Дневная производительность завода (в условных единицах) приводится в таблице:

День работы	Суточная производительность		
	Существующая схема	Вариант 1	Вариант 2
1	46	74	52
2	48	82	63
3	73	64	72
4	52	72	64
5	72	84	48
6	44	68	70
7	66	76	78
8	46	88	68
9	60	70	70
10	48	60	54
Сумма	555	738	639

$\sum\sum x_{ik} = 1932$, $\sum\sum x_{ik}^2 = 128\,810$. Принять $\alpha = 0{,}10$.

19.277. Из большой группы полевых транзисторов с недельным интервалом были получены три выборки. Ниже приводятся результаты измерения емкости затвор — сток у этих транзисторов (в пикофарадах):

№ выборки	Емкость (пФ)							
1	2,8	3,2	2,9	3,5	3,3	3,7	3,9	3,1
2	3,1	3,2	3,3	3,4	3,7	3,4	3,0	3,1
3	3,6	2,8	3,0	3,2	3,0	3,7	3,2	3,2
1	3,2	3,1	3,4	3,0	3,6	3,1	3,2	3,2
2	2,9	3,5	3,2	3,2	—	—	—	—
3	3,6	3,4	3,1	3,2	—	—	—	—

$$\sum x_{i1} = 52{,}2,\ \sum x_{i2} = \sum x_{i3} = 39,\ \sum\sum x_{ik} = 130{,}2,$$

$$\sum\sum x_{ik}^2 = 426{,}48,\ n = 40.\text{ Принять } \alpha = 0{,}10.$$

19.278. Время химической реакции при различном содержании катализатора распределилось следующим образом (в секундах):

Содержание катализатора	№ эксперимента												Сумма
	1	2	3	4	5	6	7	8	9	10	11	12	
5%	5,9	6,0	7,0	6,5	5,5	7,0	8,1	7,5	6,2	6,4	7,1	6,9	80,1
10%	4,0	5,1	6,2	5,3	4,5	4,4	5,3	5,4	5,6	5,2	—	—	51,0
15%	8,2	6,8	8,0	7,5	7,0	7,2	7,9	8,1	8,5	7,8	8,1	—	85,1

$$\sum\sum x_{ik}^2 = 1465{,}68. \text{ Принять } \alpha = 0{,}10.$$

19.279*. Доказать основное тождество (1) дисперсионного анализа.

19.280. Показать, что если гипотеза о равенстве средних верна, то общее среднее $\overline{x}$ и статистика $\dfrac{Q_2}{n-l}$ являются несмещенными оценками математического ожидания и дисперсии генеральной совокупности.

§ 6. Критерий χ^2 и его применение

1. Проверка гипотезы о виде распределения генеральной совокупности. Пусть $x_1, x_2, \ldots, x_n$ — выборка наблюдений случайной величины X. Проверяется гипотеза H_0, утверждающая, что X имеет функцию распределения $F_X(x)$.

Проверка гипотезы H_0 при помощи критерия χ^2 осуществляется по следующей схеме. По выборке наблюдений находят оценки неизвестных параметров предполагаемого закона распределения случайной величины X. Далее, область возможных значений случайной величины X разбивается на r множеств $\Delta_1, \Delta_2, \ldots, \Delta_r$, например, r интервалов в случае, когда X — непрерывная случайная величина, или r групп, состоящих из отдельных значений, для дискретной случайной величины X.

Пусть n_k — число элементов выборки, принадлежащих множеству Δ_k, $k = 1, 2, \ldots, r$. Очевидно, что $\sum\limits_{k=1}^{r} n_k = n$. Используя предполагаемый закон распределения случайной величины X, находят вероятности p_k

того, что значение X принадлежит множеству Δ_k, т.е. $p_k = \mathrm{P}[X \in \Delta_k]$, $k = 1, 2 \ldots, r$. Очевидно, что $\sum\limits_{k=1}^{r} p_k = 1$. Полученные результаты можно представить в виде следующей таблицы:

	Число наблюдений в интервале Δ_i				Всего
	Δ_1	Δ_2	$\ldots$	Δ_r	
Наблюдаемое	n_1	n_2	$\ldots$	n_r	n
Ожидаемое	np_1	np_2	$\ldots$	np_r	n

Выборочное значение статистики критерия χ^2 вычисляется по формуле

$$\chi_в^2 = \sum_{k=1}^{r} \frac{(n_k - np_k)^2}{np_k}.$$

Гипотеза H_0 согласуется с результатами наблюдений на уровне значимости α, если

$$\chi_в^2 < \chi_{1-\alpha}^2\,(r - l - 1),$$

где $\chi_{1-\alpha}^2\,(r-l-1)$ — квантиль порядка $1-\alpha$ распределения χ^2 с $r-l-1$ степенями свободы, а l — число неизвестных параметров распределения, оцениваемых по выборке; если же $\chi_в^2 \geqslant \chi_{1-\alpha}^2\,(r - l - 1)$, то гипотеза H_0 отклоняется.

Замечание. Критерий χ^2 использует тот факт, что случайная величина $\dfrac{n_k - np_k}{\sqrt{np_k}}$, $k = 1, 2, \ldots, r$, имеет распределение, близкое к нормальному $N(0, 1)$. Чтобы это утверждение было достаточно точным необходимо, чтобы для всех интервалов выполнялось условие $np_k \geqslant 5$. Если в некоторых интервалах это условие не выполняется, то их следует объединить с соседними.

Пример 1. Проверка гипотезы о распределении по закону Пуассона. В первых двух столбцах таблицы 6.1 приведены данные об отказах аппаратуры за 10 000 часов работы. Общее число обследованных экземпляров аппаратуры $n = 757$, при этом наблюдался

$$0 \cdot 427 + 1 \cdot 235 + 2 \cdot 72 + 3 \cdot 21 + 4 \cdot 1 + 5 \cdot 1 = 451 \text{ отказ.}$$

Проверить гипотезу о том, что число отказов имеет распределение Пуассона:

$$p_k = \mathrm{P}[X = k] = \frac{\lambda^k}{k!}e^{-\lambda}, \quad k = 0, 1, \ldots, \text{ при } \alpha = 0{,}01.$$

Таблица 6.1

Число отказов, k	Количество случаев, в которых наблюдалось k отказов, n_k	$p_k = \dfrac{0{,}6^k}{k!} e^{-0{,}6}$	Ожидаемое число случаев с k отказами, np_k
0	427	0,54881	416
1	235	0,32929	249
2	72	0,09879	75
3	21	0,01976	15
4	1	0,00296	2
5	1	0,00036	0
$\geqslant 6$	0	0,00004	0
Сумма	757	–	–

◁ Оценка параметра λ равна среднему числу отказов: $\tilde{\lambda} = 451/757 \approx \approx 0{,}6$. По таблице П3 с $\lambda = 0{,}6$ находим вероятности p_k и ожидаемое число случаев с k отказами (третий и четвертый столбцы таблицы 6.1).

Для $k = 4$, 5 и 6 значения $np_k < 5$, поэтому объединяем эти строки со строкой для $k = 3$. В результате получим значения, приведенные в таблице 6.2.

Таблица 6.2

k	n_k	np_k	$\dfrac{(n_k - np_k)^2}{np_k}$
0	427	416	0,291
1	235	249	0,787
2	72	75	0,120
$\geqslant 3$	23	17	2,118
–	–	–	$\chi^2_в = 3{,}316$

Так как по выборке оценивался один параметр λ, то $l = 1$, число степеней свободы равно $4 - 1 - 1 = 2$. По таблице П5 находим $\chi^2_{0,99}(2) = = 9{,}21$, следовательно, гипотеза о распределении числа отказов по закону Пуассона принимается. ▷

Пример 2. Проверить гипотезу о нормальном распределении выборки из примера 2 § 1. Принять $\alpha = 0{,}1$.

◁ Объем выборки $n = 55$. Для проверки гипотезы о нормальном распределении нужно найти оценки математического ожидания и дисперсии. Имеем

$$\tilde{m} = \overline{x} = \frac{1}{n} \sum_{i=1}^{55} x_i \approx 17{,}84,$$

$$\tilde{\sigma}^2 = s^2 = \frac{1}{n-1} \sum_{i=1}^{55} (x_i - \overline{x})^2 \approx 8{,}53.$$

Воспользуемся результатами группировки выборки в примере 2 § 1 (см. таблицу 1.1), расширив первый и последний интервалы. Результаты группировки приведены во втором и третьем столбцах таблицы 6.3.

Таблица 6.3

Номер интервала k	Границы интервала Δk	Наблюдаемая частота n_k	Вероятность попадания в интервал Δ_k, p_k	Ожидаемая частота np_k	np_k	$n_k - np_k$	$\dfrac{(n_k - np_k)^2}{np_k}$
1	$-\infty - 12$	2	0,0228	1,254	5,274	0,725	0,010
2	$12 - 14$	4	0,0731	4,020			
3	$14 - 16$	8	0,1686	9,273	9,273	-1,273	0,175
4	$16 - 18$	12	0,2576	14,168	14,168	-2,168	0,332
5	$18 - 20$	16	0,2484	13,662	13,662	-2,338	0,400
6	$20 - 22$	10	0,1519	8,354	12,633	0,366	0,011
7	$-22 - +\infty$	3	0,0778	4,279			
	Сумма	55	1,0001	55	55	–	0,928

В четвертом столбце таблицы 6.3 приведены вероятности p_k, вычисляемые по формуле

$$p_k = \mathrm{P}\left[X \in \Delta_k\right] = \Phi\left(\frac{b_k - \overline{x}}{s}\right) - \Phi\left(\frac{a_k - \overline{x}}{s}\right), \quad k = 1, 2, \ldots, 7,$$

где a_k и b_k — соответственно нижняя и верхняя границы интервалов, а значения функции $\Phi(x)$ берутся из таблицы П1. В пятом столбце приводятся ожидаемые частоты np_k, а в шестом — значения np_k после объединения первых двух и последних двух интервалов.

Так как после объединения осталось $r = 5$ интервалов, а по выборке определены оценки двух параметров, т.е. $l = 2$, то число степеней свободы равно $5 - 2 - 1 = 2$. По таблице П5 находим $\chi^2_{0,90}(2) = 4{,}61$. Выборочное значение статистики критерия равно $\chi^2_в = 0{,}928$, следовательно, гипотеза о нормальном распределении выборки принимается. ▷

19.281.** При 50 подбрасываниях монеты герб появился 20 раз. Можно ли считать монету симметричной? Принять $\alpha = 0{,}10$.

19.282. При 120 бросаниях игральной кости шестерка выпала 40 раз. Согласуется ли этот результат с утверждением, что кость правильная? Принять $\alpha = 0{,}05$.

19.283. Решить задачи 19.281, 19.282, используя методы проверки гипотез из § 4, п. 3.

19.284. Число выпадений герба при 20 подбрасываниях двух монет распределились следующим образом:

Количество гербов	0	1	2
Число подбрасываний	4	8	8

Согласуются ли эти результаты с предположениями о симметричности монет и независимости результатов подбрасываний? Принять $\alpha = 0{,}05$.

19.285. Ниже приводятся данные о фактических объемах сбыта (в условных единицах) в пяти районах:

Район	1	2	3	4	5
Фактический объем сбыта	110	130	70	90	100

Согласуются ли эти результаты с предположением о том, что сбыт продукции в этих районах должен быть одинаковым? Принять $\alpha = 0{,}01$.

19.286. На экзамене студент отвечает только на один вопрос по одной из трех частей курса. Анализ вопросов, заданных 60 студентам, показал, что 23 студента получили вопросы из первой, 15 — из второй и 22 — из третьей части курса.

Можно ли считать, что студент, идущий на экзамен, с равной вероятностью получит вопрос по любой из трех частей курса? Принять $\alpha = 0{,}10$.

19.287. Метод получения случайных чисел был применен 250 раз. при этом получены следующие результаты:

Цифра	0	1	2	3	4	5	6	7	8	9
Частота появления	27	18	23	31	21	23	28	25	22	32

Можно ли считать, что примененный метод действительно дает случайные числа? Принять $\alpha = 0{,}10$.

19.288. В цехе с 10 станками ежедневно регистрировалось число вышедших из строя станков. Всего было проведено 200 наблюдений, результаты которых приведены ниже:

Число выбывших станков	0	1	2	3	4	5	6	7	8	9	10
Число зарегистрированных случаев	41	62	45	22	16	8	4	2	0	0	0

Проверить гипотезу H_0 о том, что число выбывших из строя станков имеет распределение Пуассона. Принять $\alpha = 0{,}05$.

19.289. Во время второй мировой войны на Лондон упало 537 самолетов-снарядов. Вся территория Лондона была разделена на 576 участков площадью по 0,25 км^2. Ниже приведены числа участков n_k, на которые упало k снарядов:

k	0	1	2	3	4	5 и больше
n_k	229	211	93	35	7	1

Согласуются ли эти данные с гипотезой о том, что число снарядов, упавших на каждый из участков, имеет распределение Пуассона? Принять $\alpha = 0{,}05$.

19.290. Ниже приводятся данные о числе деталей, поступающих на конвейер в течение 600 двухминутных интервалов:

Число деталей	0	1	2	3	4	5	6
Число интервалов	400	167	29	3	0	0	1

Используя критерий χ^2, проверить гипотезу H_0 о пуассоновском распределении числа деталей при $\alpha = 0{,}05$.

19.291. При испытании радиоэлектронной аппаратуры фиксировалось число отказов. Результаты 59 испытаний приводятся ниже:

Число отказов	0	1	2	3
Число испытаний	42	10	4	3

Проверить гипотезу H_0 о том, что число отказов имеет распределение Пуассона, при $\alpha = 0{,}10$.

В задачах 19.292–19.296 для приведенных группированных выборок, приняв 10%-ный уровень значимости, проверить гипотезу H_0 о том, что они получены из нормально распределенной генеральной совокупности.

19.292. 200 отклонений размера вала от номинального значения (мкм):

Середина интервала	−0,14	−0,12	−0,10	−0,08	−0,06	−0,04	−0,02
Частота	3	8	11	20	27	36	29

Середина интервала	0,00	0,02	0,04	0,06	0,08	0,10	0,12
Частота	18	17	17	8	4	1	1

19.293. 150 отклонений диаметров цапф передней оси от номинального размера (мкм):

Середина интервала	26	29	32	35	38	41	44	47	50	53
Частота	1	4	13	23	22	29	29	16	11	2

19.294. Величина контрольного размера 68 деталей, изготовленных на одном станке (мм):

Границы интервала	2,9–3,9	3,9–4,9	4,9–5,9	5,9–6,9	6,9–7,9
Частота	5	15	23	19	6

19.295. Входное сопротивление 130 электронных ламп (Ом):

Границы интервала	3,0–3,6	3,6–4,2	4,2–4,8	4,8–5,4	5,4–6,0	6,0–6,6	6,6–7,2
Частота	2	3	35	43	22	15	5

19.296. Рост 1004 девушек в возрасте 16 лет (см):

Границы интервала	134–137	137–140	140–143	143–146	146–149	149–152	152–155
Частота	1	4	16	53	121	197	229

Границы интервала	155–158	158–161	161–164	164–167	167–170	170–173
Частота	186	121	53	17	5	1

2. Проверка гипотезы о независимости двух случайных величин. Предположим, что проведено n экспериментов, результаты которых являются значениями дискретных случайных величин X и Y, которые принимают соответственно значения $x_1, x_2, \ldots, x_k$ и $y_1, y_2, \ldots, y_l$. Обозначим n_{ij} число экспериментов, в которых $X = x_i$ и $Y = y_i$, $i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, l$. Если X и Y — непрерывные случайные величины, то область значений каждой из них разбивается на конечное число интервалов. В этом случае n_{ij} — число экспериментов, в которых случайная величина X попала в i-й интервал, а случайная величина Y — в j-й интервал. Результаты n экспериментов можно представить в виде *таблицы сопряженности признаков* размера $k \times l$ (таблица 6.4).

Таблица 6.4

X	Y				
	y_1	y_2	$\ldots$	y_l	$\sum_{j=1}^{l} n_{ij} = n_{i\cdot}$
x_1	n_{11}	n_{12}	$\ldots$	n_{1l}	$n_{1\cdot}$
x_2	n_{21}	n_{22}	$\ldots$	n_{2l}	$n_{2\cdot}$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
x_k	n_{k1}	n_{k2}	$\ldots$	n_{kl}	$n_{k\cdot}$
$\sum_{i=1}^{k} n_{ij} = n_{\cdot j}$	$n_{\cdot 1}$	$n_{\cdot 2}$	$\ldots$	$n_{\cdot l}$	$n_{\cdot\cdot} = n$

Проверяется гипотеза H_0, утверждающая, что случайные величины X и Y независимы. Если гипотеза H_0 верна, то по определению

$$\mathrm{P}\left[X = x_i;\ Y = y_i\right] = \mathrm{P}\left[X = x_i\right]\mathrm{P}\left[Y = y_j\right] = p_i q_j.$$

Пусть $\tilde{p}_i = \dfrac{n_{i\cdot}}{n}$ и $\tilde{q}_j = \dfrac{n_{\cdot j}}{n}$ — оценки вероятностей p_i и q_j.

Если гипотеза H_0 верна, то ожидаемое число экспериментов $\tilde{n}_{ij}$, в которых случайная величина X попала в i-й интервал, а случайная величина Y в j-й интервал, равно

$$\tilde{n}_{ij} = n\tilde{p}_i\tilde{q}_j = \frac{n_{i\cdot}n_{\cdot j}}{n}.$$

Для проверки гипотезы H_0 по критерию χ^2 используют следующую статистику:

$$\chi^2 = \sum_{i=1}^{k}\sum_{j=1}^{l}\frac{(n_{ij} - \tilde{n}_{ij})^2}{\tilde{n}_{ij}}. \tag{1}$$

При условии, что гипотеза H_0 верна, а все ожидаемые частоты $\tilde{n}_{ij} \geqslant 4$, $i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, l$, статистика (1) имеет распределение χ^2 с $(k-1)(l-1)$ степенями свободы.

Гипотеза H_0 о независимости случайных величин X и Y принимается на уровне значимости α, если выборочное значение статистики (1) меньше квантили $\chi^2_{1-\alpha}\left((k-1)(l-1)\right)$, т.е. если

$$\chi^2_в < \chi^2_{1-\alpha}\left((k-1)(l-1)\right).$$

Если $\chi^2_в \geqslant \chi^2_{1-\alpha}\left((k-1)(l-1)\right)$, то гипотеза H_0 отклоняется.

Для вычисления выборочного значения статистики (1) критерия удобно использовать формулу

$$\chi^2_в = n\left(\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{n_{ij}^2}{n_{i\cdot}n_{\cdot j}} - 1\right) = n\left(\sum_{i=1}^{n}\frac{1}{n_{i\cdot}}\left(\sum_{j=1}^{l}\frac{n_{ij}^2}{n_{\cdot j}}\right) - 1\right). \qquad (2)$$

Замечания. 1. Если ожидаемые частоты $\tilde{n}_{ij}$ для некоторых клеток таблицы 6.4 не удовлетворяют условию $\tilde{n}_{ij} \geqslant 4$, то соответствующие строки и столбцы должны быть объединены с соседними строками и столбцами.

2. Если $(k-1)(l-1) \geqslant 8$ и $n \geqslant 40$, то минимальное допустимое значение ожидаемых частот может быть равным единице.

Случайные величины X и Y можно рассматривать как два признака, по которым классифицируется выборка объема n; независимость X и Y соответствует независимости этих признаков.

Во многих случаях требуется проверить гипотезу об однородности нескольких выборок или, иными словами, гипотезу о том, что эти выборки получены из одной генеральной совокупности. Если проверяется однородность k различных выборок с объемами $n_1, n_2, \ldots, n_k$ и эти выборки могут быть записаны в виде таблицы сопряженности признаков размера $k \times l$ (см. табл. 6.4), то для проверки используется тот же критерий, что и для проверки независимости двух признаков.

Пример 3. Комплектующие изделия одного наименования поступают с трех предприятий A, B и C. Результаты проверки изделий следующие:

Результаты проверки	Поставщики			Всего
	A	B	C	
Годные	29	38	53	120
Негодные	1	2	7	10
Всего	30	40	60	130

Можно ли считать, что качество изделий не зависит от поставщика? Принять $\alpha = 0{,}10$.

$\lhd$ Проверяется гипотеза о независимости двух признаков: качества изделия и места его изготовления. По формуле (2) находим

$$\chi^2_в = $$
$$= 130\left(\frac{29^2}{30\cdot 120} + \frac{38^2}{40\cdot 120} + \frac{53^2}{60\cdot 120} + \frac{1^2}{30\cdot 10} + \frac{2^2}{40\cdot 10} + \frac{7^2}{60\cdot 10} - 1\right) \approx 2{,}546;$$

число степеней свободы $(2-1)(3-1) = 2$. Так как по таблице П5 $\chi^2_{0,90}(2) \approx 4{,}605$, следует считать, что качество изделий не зависит от поставщика.

Заметим, что утверждение о том, что качество изделий не зависит от поставщика, можно трактовать как проверку гипотезы об однородности трех выборок изделий объемом 30, 40 и 60, полученных соответственно от поставщиков A, B и C. $\triangleright$

19.297. Утверждается, что результат действия лекарства зависит от способа его применения. Проверить это утверждение при $\alpha = 0{,}05$ по следующим данным:

Результат	Способ применения		
	A	B	C
Неблагоприятный	11	17	16
Благоприятный	20	23	19

19.298. Отношение зрителей к включению одной из телепередач в программу выразилось следующими данными:

	Положительное	Безразличное	Отрицательное
Мужчины	14	24	2
Женщины	29	36	15

Можно ли считать, что отношение к включению данной передачи в программу не зависит от пола зрителя? Принять $\alpha = 0{,}10$.

19.299. Изменение производительности труда на предприятии при проведении мероприятий A, B и C выражается следующими данными:

Производительность	Мероприятие		
	A	B	C
Увеличилась	14	47	16
Не изменилась	22	37	7
Уменьшилась	20	25	2

Можно ли считать, что проведение этих мероприятий не влияет на производительность труда? Принять $\alpha = 0{,}10$.

19.300. Ниже приводятся результаты опроса 100 студентов первых трех курсов на вопрос «считаете ли Вы, что курение мешает учебе?»

Ответ	Курс		
	I	II	III
Нет	15	10	–
Не знаю	8	5	7
Да	–	30	25

Подтверждают ли эти данные предположение о том, что отношение к курению студентов разных курсов различно? Принять $\alpha = 0{,}01$.

19.301. Для определения зависимости цвета волос жителей от их местожительства были обследованы три группы людей из районов A, B и C. Свидетельствуют ли приводимые ниже результаты обследования о зависимости цвета волос жителей от их местожительства? Принять $\alpha = 0{,}05$.

Район	Цвет волос		
	Рыжий	Светлый	Темный
A	2	9	9
B	3	6	21
C	15	15	20

19.302. Содержание никотина (в мг) для двух марок сигарет характеризуется следующими данными:

Марка A	24	26	25	22
Марка B	27	28	25	29

Указывают ли эти результаты на различие в содержании никотина в сигаретах этих марок? Принять $\alpha = 0{,}10$.

3. Проверка гипотезы о равенстве параметров двух биномиальных распределений. Предположим, что независимо проведены две серии, содержащие n_1 и n_2 испытаний соответственно. В первой серии событие A произошло n_{11} раз, а во второй n_{21} раз. Требуется проверить гипотезу H_0 о том, что вероятность появления события A в обеих сериях одна и та же, т.е. $H_0 : p_1 = p_2$. Результаты обеих серий можно представить в виде таблицы сопряженности признаков размера 2×2:

Серия	Событие		Сумма
	A	$\overline{A}$	
1	n_{11}	n_{12}	$n_{1\cdot}$
2	n_{21}	n_{22}	$n_{2\cdot}$
Сумма	$n_{\cdot 1}$	$n_{\cdot 2}$	$n_{\cdot\cdot} = n$

Гипотеза $H_0 : p_1 = p_2$ эквивалентна гипотезе о том, что обе выборки получены из одной генеральной совокупности, т.е. однородны, и проверяется по критерию χ^2 (см. п. 2 настоящего параграфа).

В этом случае для вычисления выборочного значения статистики (1) критерия удобно использовать формулу

$$\chi^2_в = \frac{n\,(n_{11}n_{22} - n_{12}n_{21})^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}. \qquad (3)$$

Гипотеза H_0 принимается на уровне значимости α, если

$$\chi^2_в < \chi^2_{1-\alpha}(1),$$

если $\chi^2_в \geqslant \chi^2_{1-\alpha}(1)$, то гипотеза H_0 отклоняется.

Замечание. Критерий χ^2 можно использовать при условии, что все ожидаемые частоты $\tilde{n}_{ij} = \dfrac{n_{i\cdot}n_{\cdot j}}{n} > 3$, $i, j = 1, 2$ и $n > 20$. Для малых n при вычислении $\chi^2_в$ по формуле (3) нужно n заменить на $n-1$; при этом должно быть $n_{1\cdot} > 5$ и $n_{2\cdot} > \dfrac{n_{1\cdot}}{3}$.

Пример 4. Проверить гипотезу H_0 о равенстве вероятностей появления события A в двух сериях опытов по результатам таблицы 6.5. Принять $\alpha = 0{,}05$.

Таблица 6.5

Серия	Событие		Сумма
	A	$\bar{A}$	
1	3	10	13
2	6	17	23
Сумма	9	27	36

◁ Минимальное значение ожидаемой частоты

$$\tilde{n}_{11} = \frac{n_{1\cdot}n_{\cdot 1}}{n} = \frac{13\cdot 9}{36} = 3{,}25 > 3, \quad n = 36,$$

следовательно, для проверки гипотезы H_0 можно использовать критерий χ^2. По формуле (3) выборочное значение $\chi^2_в$ равно $\dfrac{36\cdot(3\cdot 17 - 6\cdot 10)^2}{13\cdot 23\cdot 9\cdot 27} \approx 0{,}04$. Число степеней свободы $(2-1)\cdot(2-1) = 1$. По таблице П5 находим $\chi^2_{0,95}(1) \approx 3{,}84$; так как $\chi^2_в < \chi^2_{0,95}(1)$, то гипотеза H_0 принимается. ▷

19.303. Исследуются два производственных процесса изготовления поршневых колец. Используя критерий χ^2, проверьте гипотезу о равенстве процента брака в обоих процессах по следующим данным при $\alpha = 0{,}01$:

Кольца	Процесс	
	1	2
Годные	195	149
Бракованные	5	2

19.304. В течение месяца завод поставил предприятию 200 корпусов, из которых 3 оказались дефектными. В следующий месяц было поставлено 850 корпусов, из которых 7 оказались дефектными. Изменилась ли доля дефектных корпусов в поставках завода? Принять $\alpha = 0{,}01$.

19.305. 1000 человек классифицировали по признаку дальтонизма. По приведенным ниже данным проверить, есть ли зависимость между наличием дальтонизма и полом человека, при $\alpha = 0{,}05$.

	Мужчины	Женщины
Дальтоники	38	6
Не дальтоники	442	514

19.306. Во время эпидемии гриппа изучалась эффективность прививок против этого заболевания. Получены следующие результаты:

После прививки		Без прививки	
Заболели	Не заболели	Заболели	Не заболели
4 чел.	192 чел.	34 чел.	111 чел.

Указывают ли эти результаты на эффективность прививок? Принять $\alpha = 0{,}01$.

§ 7. Элементы регрессионного анализа и метод наименьших квадратов

1. Линейная регрессия. В регрессионном анализе изучается связь между зависимой переменной Y и одной или несколькими независимыми переменными. Пусть переменная Y зависит от одной переменной x. При этом предполагается, что переменная x принимает заданные (фиксированные) значения, а зависимая переменная Y имеет случайный разброс из-за ошибок измерения, влияния неучтенных факторов или других причин. Каждому значению переменной x соответствует некоторое вероятностное распределение случайной величины Y. Предположим, что случайная величина Y «в среднем» линейно зависит от значений переменной x. Это означает, что условное математическое ожидание случайной величины Y при заданном значении переменной x имеет вид

$$\mathrm{M}[Y/x] = \beta_0 + \beta_1 x. \tag{1}$$

Функция переменной x, определяемая правой частью формулы (1), называется *линейной регрессией Y на x*, а параметры β_0 и β_1 — *параметрами линейной регрессии*. На практике параметры линейной регрессии (1) неизвестны, и их оценки определяют по результатам наблюдений переменных Y и x.

Пусть проведено n независимых наблюдений случайной величины Y при значениях переменной $x = x_1,\ x_2, \ldots, x_n$, при этом измерения величины Y дали следующие результаты: $y_1,\ y_2, \ldots, y_n$. Так как эти значения имеют «разброс» относительно линейной регрессии (1), то связь между переменными Y и x можно записать в виде *линейной (по параметрам β_0 и β_1) регрессионной модели:*

$$Y = \beta_0 + \beta_1 x + \varepsilon, \tag{2}$$

где ε — случайная ошибка наблюдений, причем предполагается $\mathrm{M}[\varepsilon] = 0$, $\mathrm{D}[\varepsilon] = \sigma^2$. Значение дисперсии ошибок наблюдений σ^2 неизвестно, и оценка ее определяется по результатам наблюдений.

Задача линейного регрессионного анализа состоит в том, чтобы по результатам наблюдений $(x_i,\ y_i)$, $i = 1, 2, \ldots, n$:

а) получить наилучшие точечные и интервальные оценки неизвестных параметров β_0, β_1 и σ^2 модели (2);

б) проверить статистические гипотезы о параметрах модели;

в) проверить, достаточно ли хорошо модель согласуется с результатами наблюдений (адекватность модели результатам наблюдений).

В соответствии с моделью (2) результаты наблюдений зависимой переменной Y : $y_1, y_2, \ldots, y_n$ являются реализациями случайных величин

$$\beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, \ldots, n, \tag{3}$$

обозначаемых Y_i.

Задача линейного регрессионного анализа решается в предположении, что случайные ошибки наблюдений ε_i и ε_j не коррелированы при $i \neq j$, $i, j = 1, 2, \ldots, n$ имеют математические ожидания, равные нулю, и одну и ту же дисперсию, равную σ^2, т.е.

$$\mathrm{M}[\varepsilon_i] = 0,$$
$$K_{\varepsilon_i \varepsilon_j} = \begin{cases} 0, & i \neq j, \\ \sigma^2, & i = j, \quad i, j = 1, 2, \ldots, n. \end{cases} \tag{4}$$

При статистическом анализе регрессионной модели (2) предполагается также, что случайные ошибки наблюдений ε_i, $i = 1, 2, \ldots, n$, имеют нормальное распределение, т.е.

$$\varepsilon_i \sim N(0, \sigma), \quad i = 1, 2, \ldots, n. \tag{5}$$

В этом случае ошибки наблюдений ε_i также являются независимыми случайными величинами.

Для нахождения оценок параметров модели (2) по результатам наблюдений используется *метод наименьших квадратов* (МНК). По этому методу в качестве оценок параметров выбирают такие значения $\tilde{\beta}_0$ и $\tilde{\beta}_1$, которые минимизируют сумму квадратов отклонений наблюдаемых значений случайных величин Y_i, $i = 1, 2, \ldots, n$ от их математических ожиданий, т.е. сумму

$$Q(\beta_0, \beta_1) = \sum_{i=1}^{n} [y_i - (\beta_0 + \beta_1 x_i)]^2. \tag{6}$$

Из необходимых условий минимума функции (6)

$$\frac{\partial Q}{\partial \beta_0} = 0, \quad \frac{\partial Q}{\partial \beta_1} = 0$$

получим, что МНК-оценки параметров линейной регрессии имеют вид:

$$\tilde{\beta}_1 = \frac{\sum x_i y_i - \frac{1}{n}\left(\sum x_i\right)\left(\sum y_i\right)}{\sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2} = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sum (x_i - \overline{x})^2} = \frac{Q_{xy}}{Q_x}, \tag{7}$$

$$\tilde{\beta}_0 = \overline{y} - \tilde{\beta}_1 \overline{x}, \tag{8}$$

где

$$\overline{x} = \frac{1}{n}\sum x_i, \qquad \overline{y} = \frac{1}{n}\sum y_i,$$

$$Q_{xy} = \sum (x_i - \overline{x})(y_i - \overline{y}), \quad Q_x = \sum (x_i - \overline{x})^2.$$

Оценки параметров линейной регрессии, получаемые по методу наименьших квадратов, при любом законе распределения ошибок наблюдений ε_i, $i = 1, 2, \ldots, n$, при условиях (4) имеют следующие свойства:

1. Они являются *линейными функциями результатов наблюдений* y_i, $i = 1, 2, \ldots, n$, и *несмещенными оценками* параметров, т.е. $\mathrm{M}[\tilde{\beta}_j] = \beta_j$, $j = 0, 1$.

2. Они имеют *минимальные дисперсии* в классе несмещенных оценок, являющихся линейными функциями результатов наблюдений (теорема Гаусса–Маркова).

Если ошибки наблюдений ε_i не коррелированы и имеют нормальное распределение, т.е. $\varepsilon_i \sim N(0, \sigma)$ (предположение (5)), то в дополнение к свойствам 1 и 2 выполняется свойство

3. *МНК-оценки совпадают с оценками, вычисляемыми по методу максимального правдоподобия.*

Функция

$$y = \tilde{\beta}_0 + \tilde{\beta}_1 x \tag{9}$$

определяет выборочную (эмпирическую) регрессию Y на x (см. также § 1, п. 3). Последняя является оценкой предполагаемой (теоретической) линейной регрессии (1) по результатам наблюдений. Разности между наблюдаемыми значениями переменной Y при $x = x_i$, $i = 1, 2, \ldots, n$, и расчетными значениями $\tilde{y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 x_i$ называются остатками и обозначаются e_i:

$$e_i = y_i - \tilde{y}_i, \quad i = 1, 2, \ldots, n. \tag{10}$$

19.307. Найти оценки параметров линейной регрессии по выборке (9; 6), (10; 4), (12; 7), (5; 3). Нанести прямую регрессию на диаграмму рассеяния и показать отрезки, соответствующие разностям $y_i - \overline{y}$, $\tilde{y}_i - \overline{y}$ и $y_i - \tilde{y}_i$, $i = 1, 2, 3, 4$.

19.308. Показать, что сумма остатков $\sum e_i$ равна нулю.

19.309. Для представления некоторых данных предполагается использовать модель $y = \beta_0 + \beta_1 x$, где значение β_1 известно. Найти оценку параметра β_0.

19.310. В модели $y = \beta_0 + \beta_1 x$ параметр β_0 известен. Найти оценку параметра β_1.

19.311. Показать, что точка $(\overline{x}, \overline{y})$ лежит на прямой $y = \tilde{\beta}_0 + \tilde{\beta}_1 x$.

В задачах 19.312–19.315 исследуются статистические свойства МНК-оценок параметров линейной регрессии. Оценки $\tilde{\beta}_0$ и $\tilde{\beta}_1$ (см. (7), (8)) являются линейными функциями случайных величин $Y_i = \beta_0 + \beta_1 x + \varepsilon_i$, $i = 1, 2, \ldots, n$, причем ε_i удовлетворяют предположениям (4).

19.312*. Показать, что МНК-оценки параметров линейной регрессии являются несмещенными оценками этих параметров.

19.313*. Показать, что дисперсии оценок $\tilde{\beta}_1$ и $\tilde{\beta}_0$ равны соответственно

$$\mathrm{D}[\tilde{\beta}_1] = \frac{\sigma^2}{Q_x}, \quad \mathrm{D}[\tilde{\beta}_0] = \frac{\sigma^2 \sum x_i^2}{n Q_x}.$$

19.314*. Показать, что коэффициент ковариации $K_{\tilde{\beta}_0 \tilde{\beta}_1}$ равен

$$K_{\tilde{\beta}_0 \tilde{\beta}_1} = -\frac{\overline{x}\sigma^2}{Q_x}.$$

19.315*. Пусть независимая переменная x принимает значение x_0. Вычислить математическое ожидание случайной величины $Y_0 = \tilde{\beta}_0 + \tilde{\beta}_1 x_0$. Показать, что

$$\mathrm{D}[Y_0] = \frac{\sigma^2}{n} + \frac{(x_0 - \overline{x})^2 \sigma^2}{Q_x}.$$

19.316. Показать, что выборочную регрессию $y = \tilde\beta_0 + \tilde\beta_1 x$ можно записать в виде $y = \overline{y} - r_{xy}\dfrac{s_x}{s_Y}(x - \overline{x})$, где s_x и s_Y — оценки средних квадратичных отклонений переменных x и Y по результатам наблюдений, а r_{xy} — выборочный коэффициент корреляции (см. задачу 19.62).

19.317*. При нескольких выбранных значениях x изменена величина Y. Можно ли полученные данные использовать для оценки параметров модели

$$x = \beta_0 + \beta_1 y\,?$$

19.318. Найти МНК-оценки параметров модели

$$y = \beta_0 + \beta_1(x - \overline{x}).$$

Показать, что полученные оценки являются несмещенными.

19.319. Показать, что МНК-оценки параметров модели

$$y = \beta_0 + \beta_1(x - \overline{x})$$

являются некоррелированными и имеют дисперсии

$$\mathrm{D}[\tilde\beta_0] = \frac{\sigma^2}{n}, \quad \mathrm{D}[\tilde\beta_1] = \frac{\sigma^2}{Q_x}.$$

19.320. Пусть $\theta_1, \theta_2, \ldots, \theta_n$ — результаты измерений величины Θ. Предположим, что ошибки измерений ε_i не коррелированы и имеют равные дисперсии. Используя метод наименьших квадратов, найти оценку Θ и несмещенную оценку дисперсии ошибок наблюдений.

Качество аппроксимации результатов наблюдений (x_i, y_i), $i = 1, 2, \ldots, n$, выборочной регрессии (9) определяется величиной *остаточной дисперсии*, вычисляемой по формуле:

$$s^2 = \frac{\sum e_i^2}{n-2} = \frac{1}{n-2}\sum\left[y_i - (\tilde\beta_0 + \tilde\beta_1 x_i)\right]^2 = \frac{Q_e}{n-2}. \qquad (11)$$

Величина Q_e, определяемая выражением

$$Q_e = \sum e_i^2 = \sum (y_i - \tilde y_i)^2, \qquad (12)$$

называется *остаточной суммой квадратов*.

Если модель согласуется с результатами наблюдений (адекватна результатам наблюдений, о проверке адекватности см. ниже), то остаточная дисперсия является несмещенной оценкой дисперсии ошибок наблюдений σ^2, т.е. $\mathrm{M}[s^2] = \sigma^2$. *Всюду в дальнейшем будем предполагать, что ошибки наблюдений ε_i, $i = 1, 2, \ldots, n$, имеют нормальное* *распределение*: $\varepsilon_i \sim N(0, \sigma)$ *и независимы* (предположение (5)). Это предположение в силу (3) эквивалентно тому, что результаты наблюдений y_i, $i = 1, 2, \ldots, n$, являются реализациями независимых нормально распределенных случайных величин Y_i:

$$Y_i \sim N(\beta_0 + \beta_1 x_i, \sigma), \quad i = 1, 2, \ldots, n.$$

В этом случае можно показать (см., например, [10], с. 313–315), что статистика Q_e/σ^2 имеет распределение χ^2 с $n-2$ степенями свободы, т.е.

$$\frac{Q_e}{\sigma^2} = \chi^2(n-2), \tag{13}$$

и эта статистика распределена независимо от распределения оценок $\tilde{\beta}_0$ и $\tilde{\beta}_1$. Используя это утверждение, можно построить доверительные интервалы для параметров линейной регрессии (см. задачи 19.322, 19.323) и проверить гипотезы о параметрах.

В практических вычислениях остаточную сумму квадратов получают из тождества

$$\sum (y_i - \overline{y})^2 = \sum (\tilde{y}_i - \overline{y})^2 + \sum (y_i - \tilde{y}_i)^2, \tag{14}$$

которое записывается в виде

$$Q_y = Q_R + Q_e,$$

где

$$Q_y = \sum (y_i - \overline{y})^2 = \sum y_i^2 - n\overline{y}^2,$$

$$Q_R = \sum (\tilde{y}_i - \overline{y})^2 = \tilde{\beta}_1 Q_{xy} = \tilde{\beta}_1^2 Q_x = \frac{Q_{xy}^2}{Q_x}. \tag{15}$$

Величина Q_R называется *суммой квадратов, обусловленной регрессией.*

Тождество (14) и формулы (15) легко доказываются (см. задачи 19.311, 19.312).

Линейная регрессионная модель (2) называется *незначимой*, если параметр $\beta_1 = 0$. Для проверки гипотезы $H_0: \ \beta_1 = 0$ используют либо доверительный интервал для параметра β_1 (см. задачу 19.322), либо статистику

$$F = \frac{Q_R(n-2)}{Q_e} = \frac{\tilde{\beta}_1^2 Q_x}{s^2}. \tag{16}$$

Если гипотеза $H_0: \ \beta_1 = 0$ верна, то статистика (16) имеет распределение Фишера с 1 и $n-2$ степенями свободы (см. задачу 19.313).

В случае, когда гипотеза $H_0: \ \beta_1 = 0$ отклоняется, говорят, что регрессионная модель (2) *статистически значима*. Из этого не следует, конечно, что модель хорошо согласуется с результатами наблюдений, т.е. адекватна им.

Полезной характеристикой линейной регрессии является *коэффициент детерминации* R^2, вычисляемый по формуле

$$R^2 = \frac{Q_R}{Q_y} = 1 - \frac{Q_e}{Q_y}. \tag{17}$$

Коэффициент детерминации R^2 равен той доле разброса результатов наблюдений (x_i, y_i), $i = 1, 2, \ldots, n$, относительно горизонтальной прямой $y = \overline{y}$, которая объясняется выборочной регрессией (9). Величина $R = +\sqrt{R^2}$ является оценкой коэффициента корреляции между результатами наблюдений y_i и вычисленными значениями $\tilde{y}_i$, предсказываемыми регрессией (9), т.е.

$$R = \tilde{\rho}_{Y\hat{Y}} = r_{y\tilde{y}}.$$

В случае линейной регрессии Y на x (одной независимой переменной x) между коэффициентом R и выборочным коэффициентом корреляции r_{xy} имеется следующее соотношение:

$$r_{xy} = (\text{знак } \tilde{\beta}_1)\, R.$$

Пример 1. На химическом производстве в ходе пяти рабочих смен получены следующие данные о зависимости выхода продуктов Y (кг/ч) от температуры реакции x (°C):

x	51	32	80	73	64	45	83	44	93
Y	52,7	15,2	89,5	94,8	76	39,3	114,8	36,5	137,4

x	28	35	40	29	53	58	65	74
Y	5,3	20,7	21,7	9,2	55,4	64,3	79,1	101

Найти оценки параметров линейной регрессии Y на x, оценку дисперсии ошибок наблюдений, доверительные интервалы для параметров и среднего значения Y при $x = x_0$, коэффициент детерминации и выборочный коэффициент корреляции r_{xy}. Проверить значимость линейной регрессии. Принять $\alpha = 0{,}05$.

$\triangleleft$ По данным измерений имеем

$$n = 17, \quad \sum x_i = 948, \quad \sum x_i^2 = 59\,422, \quad \sum y_i = 1012{,}39$$

$$\sum y_i^2 = 85\,389{,}73, \quad \overline{x} = 55{,}765, \quad \sum x_i y_i = 69\,189{,}4, \quad \overline{y} = 59{,}58.$$

Для контроля вычислим

$$\sum (x_i + y_i)^2 = 283\,190{,}5,$$

$$\sum x_i^2 + 2\sum x_i y_i + \sum y_i^2 = 59\,422 + 2 \cdot 69\,189{,}4 + 85\,389{,}73 = 283\,190{,}53.$$

Следовательно, вычисления проведены верно. Предварительно найдем

$$Q_x = 59\,442 - \frac{948^2}{17} \approx 6557{,}0588,$$

$$Q_y = 85\,389{,}73 - \frac{1012{,}9^2}{17} \approx 25\,038{,}76,$$

$$Q_{xy} = 69\,189{,}4 - \frac{948 \cdot 1012{,}9}{17} \approx 12\,705{,}33.$$

По формулам (7) и (8) найдем оценки коэффициентов регрессии:

$$\tilde{\beta}_1 = \frac{12\,705{,}33}{6557{,}0588} \approx 1{,}94,$$

$$\tilde{\beta}_0 = 59{,}58 - 1{,}94 \cdot 55{,}765 \approx -48{,}47.$$

Таким образом, выборочная линейная регрессия (9) имеет вид

$$y = -48{,}47 + 1{,}94x.$$

По формуле (15) находим сумму квадратов, обусловленную регрессией Q_R:

$$Q_R = \tilde{\beta}_1 Q_{xy} = 1{,}94 \cdot 12\,705{,}33 \approx 24\,618{,}57.$$

Используя тождество (14), находим остаточную сумму квадратов Q_e:

$$Q_e = Q_y - Q_R = 25\,038{,}76 - 24\,618{,}57 \approx 420{,}19.$$

Оценка дисперсии ошибок наблюдений по формуле (11) равна

$$s^2 = \frac{Q_e}{n-2} = \frac{420{,}19}{17-2} \approx 28{,}01, \quad s \approx 5{,}29.$$

Значение квантили $t_{1-\alpha/2}(n-2) = t_{0,975}(15) = 2{,}131$ (таблица П6). Границы доверительных интервалов (см. задачи 19.322, 19.323) равны: для коэффициента β_0

$$-48{,}47 \pm 2{,}131 \cdot 5{,}29 \cdot \sqrt{\frac{59\,422}{17 \cdot 6557{,}059}} \approx -48{,}47 \pm 8{,}23;$$

для коэффициента β_1

$$1{,}94 \pm 2{,}131 \cdot 5{,}29 \cdot \sqrt{\frac{1}{6557{,}059}} \approx 1{,}94 \pm 0{,}14.$$

Границы доверительного интервала для значения Y_0, соответствующего заданному значению переменной $x = x_0$,

$$\tilde{y}_0 \pm 2{,}131 \cdot 5{,}29 \cdot \sqrt{\frac{1}{17} + \frac{(x_0 - 55{,}765)^2}{6557{,}059}} =$$

$$= \tilde{y}_0 \pm 11{,}27 \sqrt{0{,}059 + \frac{(x_0 - 55{,}765)^2}{6557{,}059}}.$$

Коэффициент детерминации R^2 по формуле (17) равен

$$1 - \frac{420{,}19}{25\,038{,}76} \approx 0{,}983.$$

Этот результат означает, что полученное уравнение регрессии на 98,4 % объясняет общий разброс результатов наблюдений относительно горизонтальной прямой $y = 59{,}547$. Выборочный коэффициент корреляции

$$r_{xy} = +\sqrt{0{,}983} \approx 0{,}991.$$

Проверим значимость линейной регрессии. Гипотеза $H_0 : \beta_1 = 0$ отклоняется на уровне значимости $\alpha = 0{,}05$, так как доверительный интервал для β_1, приведенный выше, не накрывает нуль с доверительной вероятностью 0,95.

Этот же результат получим, используя для проверки гипотезы H_0 статистику (16). Выборочное значение этой статистики

$$F_в = \frac{(n-2) \cdot Q_R}{Q_e} = \frac{15 \cdot 24\,618{,}57}{420{,}19} \approx 878{,}82.$$

Так как квантиль распределения Фишера $F_{1-\alpha}(1, n-2) = F_{0,95}(1{,}15) = 4{,}54$ (таблица П7), что меньше выборочного значения статистики $F_в$, гипотеза H_0 отклоняется на уровне значимости $\alpha = 0{,}05$. Таким образом, линейная регрессия Y на x статистически значима. ▷

19.321*. Какое распределение имеют оценки параметров линейной регрессии $\tilde{\beta}_0$ и $\tilde{\beta}_1$?

19.322*. Показать, что границы доверительных интервалов для параметров линейной регрессии имеют вид

$$\tilde{\beta}_0 \pm t_{1-\alpha/2}\,(n-2)\, s \sqrt{\frac{\sum x_i^2}{nQ_x}},$$

$$\tilde{\beta}_1 \pm t_{1-\alpha/2}\,(n-2)\, s \sqrt{\frac{1}{Q_x}}.$$

Здесь $t_p(n-2)$ — квантиль распределения Стьюдента с $n-2$ степенями свободы порядка p, а $s = \sqrt{s^2}$, где s^2 — остаточная дисперсия.

19.323*. Показать, что границы доверительного интервала для среднего значения Y_0, соответствующего заданному значению $x = x_0$, определяются формулой

$$\tilde{y}_0 \pm t_{1-\alpha/2}(n-2)\, s \sqrt{\frac{1}{n} + \frac{(x_0 - \overline{x})^2}{Q_x}}.$$

19.324*. Показать, что доверительный интервал для дисперсии ошибок наблюдений σ^2 имеет вид

$$\frac{(n-2)\, s^2}{\chi^2_{1-\alpha/2}(n-2)} < \sigma^2 < \frac{(n-2)\, s^2}{\chi^2_{\alpha/2}(n-2)},$$

где $\chi^2_p(n-2)$ — квантиль распределения χ^2 с $n-2$ степенями свободы, а s^2 — остаточная дисперсия.

19.325*. Найти оценки параметров линейной регрессии (1), используя метод максимального правдоподобия.

19.326*. Доказать тождество $Q_y = Q_R + Q_e$.

19.327*. Вывести формулы (15) для вычисления суммы квадратов, обусловленной регрессией Q_R.

19.328*. Показать, что если верна гипотеза $H_0: \ \beta_1 = 0$, то статистика $\dfrac{(n-2)\, Q_R}{Q_e} = \dfrac{\tilde{\beta}_1^2 Q_x}{s^2}$ имеет распределение Фишера с 1 и $n-2$ степенями свободы.

В задачах 19.329–19.332 по заданным выборкам найти оценки для параметров линейной регрессии Y на x, проверить значимость линейной регрессии и вычислить коэффициент корреляции. Найти границы доверительных интервалов для параметров линейной модели и для среднего значения Y при $x = x_0$.

Предполагается, что ошибки наблюдений независимы и имеют нормальное распределение $N(0, \sigma)$. Уровень значимости α задается. При вычислении следует использовать значения сумм переменных, их квадратов и попарных произведений.

19.329.

x	2	3	8	10	14	15
y	14,39	9,45	7,05	5,32	16,94	1,97

x	4	12	3	7	6
y	8,75	3,41	13,37	8,22	9,39

$$\sum x_i = 86, \quad \sum y_i = 98{,}26, \quad \sum x_i^2 = 868,$$

$$\sum y_i^2 = 1087{,}91, \quad \sum x_i y_i = 682{,}25, \quad \alpha = 0{,}05.$$

19.330.

x	2,7	4,6	6,3	7,8	9,2	10,6	12,0	13,4	14,7
y	17,0	16,2	13,3	13,0	9,7	9,9	6,2	5,8	5,7

$$\sum x_i = 81{,}3, \quad \sum y_i = 96{,}8, \quad \sum x_i^2 = 865{,}63,$$

$$\sum y_i^2 = 1194, \quad \sum x_i y_i = 735{,}7, \quad \alpha = 0{,}10.$$

19.331.

x	7,9	11,6	12,8	14,9	16,3	18,6	20,3	21,9	23,6
y	13,0	22,8	24,8	28,6	31,6	38,7	40,0	44,9	43,0

$$\sum x_i = 147{,}9, \quad \sum y_i = 287{,}4, \quad \sum x_i^2 = 2643{,}13,$$

$$\sum y_i^2 = 10\,083{,}1, \quad \sum x_i y_i = 5155{,}77, \quad \alpha = 0{,}01.$$

19.332.

x	1	2	3	4	5	6	7	8	9
y	0,21	0,32	0,58	1,02	1,76	2,68	3,75	5,07	6,62

x	10	11	12	13	14	15	16	17
y	8,32	10,21	12,33	14,58	17,07	19,53	22,72	29,05

$$\sum x_i = 153, \quad \sum y_i = 155{,}82, \quad \sum x_i^2 = 1785,$$

$$\sum y_i^2 = 2666{,}37, \quad \sum x_i y_i = 2080{,}25, \quad \alpha = 0{,}05.$$

Линейная регрессионная модель называется *адекватной*, если предсказанные по ней значения переменной Y согласуются с результатами наблюдений. Грубая оценка адекватности модели может быть проведена непосредственно по графику остатков, т.е. разностей между наблюдаемыми значениями y_i и вычисленными значениями $\tilde{y}_i$, $i = 1, 2, \ldots, n$. Если модель адекватна, то остатки e_i являются реализациями случайных ошибок наблюдений ε_i, $i = 1, 2, \ldots, n$, которые в силу предположений (5) должны быть независимыми нормально распределенными случайными величинами с нулевыми средними и одинаковыми дисперсиями σ^2. Проверка выполнения этих предположений различными статистическими методами и лежит в основе оценки адекватности по графику остатков ([12, гл. 3]).

Если для каждого или некоторых значений переменной x имеется несколько повторных наблюдений случайной величины Y, то для проверки адекватности модели можно использовать следующую процедуру.

Пусть повторные наблюдения получены при различных значениях $x_1, x_2, \ldots, x_m$ переменной x, причем при $x = x_i$ произведено n_i наблюдений Y, где $\sum\limits_{i=1}^{m} n_i = n$ — объем всей выборки наблюдений. Обозначим y_{ij}, $j = 1, 2, \ldots, n_i$, результаты повторных наблюдений Y при $x = x_i$, $i = 1, 2, \ldots, m$.

Если модель адекватна наблюдаемым данным, то средние n_i наблюдений, т.е. $\overline{y}_i = \dfrac{1}{n_i}\sum\limits_{j=1}^{n_i} y_{ij}$, $i = 1, 2, \ldots, m$, должны быть близки к вычисленным значениям $\tilde{y}_i$. Следовательно, сумма квадратов $Q_n = \sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n_i}(\overline{y}_i - \tilde{y}_i)^2 = \sum\limits_{i=1}^{m} n_i(\overline{y}_i - \tilde{y}_i)^2$ является *мерой неадекватности модели*. Остаточная сумма квадратов Q_e может быть разбита на две суммы:

$$Q_e = Q_n + Q_p, \tag{18}$$

где Q_p — *сумма квадратов чистой ошибки*:

$$Q_p = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_i)^2. \tag{19}$$

Тождество (18) легко доказывается. Для этого обе части равенства

$$(y_{ij} - \tilde{y}_i) = (y_{ij} - \overline{y}_i) - (\tilde{y}_i - \overline{y}_i)$$

нужно возвести в квадрат и просуммировать по i и j.

Как и в случае однофакторного дисперсионного анализа (§ 5), можно показать, что если линейная регрессия адекватна данным, то статистики Q_n/σ^2 и Q_p/σ^2 независимы и имеют распределение χ^2 с $m-2$ и $n-m$ степенями свободы, следовательно отношение

$$\frac{Q_n/(m-2)}{Q_p/(n-m)} = F\,(m-2,\ n-m) \tag{20}$$

имеет распределение Фишера.

Статистика (20) используется для проверки адекватности линейной регрессии. Если выборочное значение статистики (20) удовлетворяет условию

$$F_в < F_{1-\alpha}(m-2, n-m),$$

то гипотеза об адекватности линейной регрессии результатам наблюдений принимается и остаточную дисперсию $s^2 = \dfrac{Q_e}{n-2}$ можно использовать в качестве оценки дисперсии σ^2, найти доверительные интервалы для параметров линейной регрессии и проверить гипотезы о параметрах.

В противном случае нужно попытаться использовать другую модель, например параболическую регрессию.

П р и м е р 2. Найти оценки параметров линейной регрессии по следующим данным:

x	1	1	2	2	3	3	2,7	2,7	4,3	4,3	4,3	5,0	5,0
y	0,5	0,1	0,5	1,2	1,2	1,7	0,9	2,2	1,1	1,7	2,5	2,0	2,2

Проверить адекватность регрессии этим данным. Принять $\alpha = 0{,}05$.

◁ Выпишем результаты повторных наблюдений:

x_i	1	2	3	2,7	4,3	5,0
y_{ij}	0,5; 0,1	0,5; 1,2	1,2; 1,7	0,9; 2,2	1,1; 1,7; 2,5	2,0; 2,2
n_i	2	2	2	2	3	2

Имеем: $n = 13$, $m = 6$. Находим вспомогательные суммы:

$$\sum_{i=1}^{m}\sum_{j=1}^{n_j} y_{ij} = 17{,}8, \quad \sum\sum y_{ij}^2 = 31{,}12,$$

$$\sum n_i x_i = 40{,}3, \quad \sum n_i x_i^2 = 148{,}05, \quad \sum x_i y_i = 64{,}08.$$

Отсюда

$$Q_{xy} = 64{,}86 - \frac{40{,}3 \cdot 17{,}8}{13} = 9{,}68,$$

$$Q_x = 148{,}05 - \frac{40{,}3^2}{13} = 23{,}12,$$

$$Q_y = 31{,}12 - \frac{17{,}8^2}{13} = 6{,}74.$$

По формулам (7), (8) находим

$$\tilde{\beta}_1 = \frac{9{,}68}{23{,}12} \approx 0{,}419, \quad \tilde{\beta}_0 = \frac{17{,}8}{13} - 0{,}419 \cdot \frac{40{,}13}{13} \approx 0{,}070.$$

По формулам (15) находим сумму квадратов, обусловленную регрессией:

$$Q_R = 0{,}419^2 \cdot 23{,}12 \approx 4{,}059.$$

Остаточная сумма квадратов определяется из тождества (14):

$$Q_e = 6{,}74 - 4{,}059 = 2{,}681.$$

Для парных повторных наблюдений при вычислении Q_p удобно пользоваться соотношением

$$\sum_{j=1}^{2}(y_{ij} - \overline{y}_i)^2 = \frac{1}{2}\,(y_{i1} - y_{i2})^2.$$

Используя это соотношение, а также то, что при $x = 4{,}3$ имеем $\sum_{j=1}^{3}(y_{ij} - \overline{y}_j)^2 \approx 0{,}99$, по формуле (19) получим

$$Q_p = \frac{1}{2}\cdot 0{,}4^2 + \frac{1}{2}\cdot 0{,}7^2 + \frac{1}{2}\cdot 0{,}5^2 + \frac{1}{2}\cdot 1{,}3^2 + 0{,}99 + \frac{1}{2}\cdot 0{,}2^2 \approx 2{,}294.$$

Используя тождество (18) находим сумму квадратов, обусловленную неадекватностью:

$$Q_n = 2{,}681 - 2{,}294 = 0{,}387.$$

Выборочное значение статистики (20) равно

$$F_в = \frac{0{,}387/(6-2)}{2{,}294/(13-6)} \approx 0{,}296.$$

Так как квантиль $F_{0,95}(4{,}7) = 4{,}14$ (таблица П7), то линейная регрессия адекватна результатам наблюдений. ▷

В задачах 19.333, 19.334 проверить адекватность линейной регрессии. Построить график остатков. Принять $\alpha = 0{,}05$.

19.333.

x	10	10	10	10	10	20	20	20	20	35	35	35
Y	5	6	5	6	7	12	13	14	13	17	19	16

x	35	35	40	40	40	40	40	60	60	60	60	60
Y	15	15	18	20	21	18	20	17	19	16	14	16

19.334. Выборка задана в виде таблицы частот:

y	x					
	1	2	3	4	5	6
1	2	1	—	—	—	—
2	1	2	—	—	—	—
3	—	3	1	—	—	—
4	—	1	3	1	—	—
5	—	—	2	2	2	1
6	—	—	—	1	1	1

2. Линейная регрессионная модель общего вида (криволинейная регрессия). В общем случае, если регрессия Y на x отличается от линейной, рассматривают *линейную (по параметрам) регрессионную модель* вида

$$\mathrm{M}\,[Y/x] = \beta_0 + \beta_1 a_1(x) + \dots + \beta_{k-1} a_{k-1}(x), \tag{21}$$

где $a_1(x), \dots, a_{k-1}(x)$ — известные функции, а $\beta_0, \beta_1, \dots, \beta_{k-1}$ — неизвестные параметры.

Пусть (x_i, y_i), $i = 1, 2, \dots, n$ — результаты наблюдений переменных x и Y. С учетом случайных флуктуаций переменной Y результаты наблюдений $y_1, y_2, \dots, y_n$ являются реализациями случайных величин

$$Y_i = \beta_0 + \beta_1 a_1(x) + \dots + \beta_{k-1} a_{k-1}(x) + \varepsilon_i,$$

где ε_i — случайные ошибки наблюдений, распределение которых удовлетворяет условиям

$$\mathrm{M}[\varepsilon_i] = 0,$$
$$K_{\varepsilon_i \varepsilon_j} = \begin{cases} 0, & i \neq j, \\ \sigma^2, & i = j, \quad i, j = 1, 2, \ldots, n. \end{cases}$$

Как и в п. 1, при статистическом анализе регрессионной модели (21) предполагается, что случайные ошибки наблюдений ε_i имеют нормальное распределение, т.е.

$$\varepsilon_i \sim N(0, \sigma), \quad i = 1, 2, \ldots, n,$$

и, следовательно, являются независимыми случайными величинами.

Методы, используемые для решения задачи регрессионного анализа в п. 1, легко обобщаются на случай линейной регрессионной модели (21). Для нахождения оценок параметров β_j, $j = 0, 1, \ldots, k-1$, по результатам наблюдений используется метод наименьших квадратов. При этом МНК-оценки параметров модели (20) имеют те же свойства, что и МНК-оценки параметров линейной регрессии.

По методу наименьших квадратов в качестве оценок параметров β_0, $\beta_1, \ldots, \beta_{k-1}$ (МНК-оценок) принимаются значения $\tilde{\beta}_0, \tilde{\beta}_1, \ldots, \tilde{\beta}_{k-1}$, дающие минимум функции

$$Q(\beta_0, \ldots, \beta_{k-1}) = \sum_{i=1}^{n} (y_i - \tilde{y}_i)^2 = \\ = \sum_{i=1}^{n} [y_i - (\beta_0 + \beta_1 a_1(x_i) + \cdots + \beta_{k-1} a_{k-1}(x_i))]^2. \quad (22)$$

Из необходимых условий минимума функции $Q(\beta_0, \beta_1, \ldots, \beta_{k-1})$ в (22) следует, что оценки $\tilde{\beta}_0, \tilde{\beta}_1, \ldots, \tilde{\beta}_{k-1}$ являются решениями линейной алгебраической системы k уравнений

$$\beta_0 n + \beta_1 \sum a_1(x_i) + \cdots + \beta_{k-1} \sum a_{k-1}(x_i) = \sum y_i,$$
$$\beta_0 \sum a_1(x_i) + \beta_1 \sum a_1(x_i)\, a_1(x_i) + \ldots \\ \cdots + \beta_{k-1} \sum a_{k-1}(x_i)\, a_1(x_i) = \sum y_i a_1(x_i), \quad (23)$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\beta_0 \sum a_{k-1}(x_i) + \beta_1 \sum a_1(x_i)\, a_{k-1}(x_i) + \ldots \\ \cdots + \beta_{k-1} \sum a_{k-1}(x_i)\, a_{k-1}(x_i) = \sum y_i a_{k-1}(x_i),$$

называемой *нормальной системой*.

С использованием следующих матричных обозначений:

$$\boldsymbol{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \text{— вектор наблюдений,}$$

$$A = \begin{pmatrix} 1 & a_1(x_1) & \ldots & a_{k-1}(x_1) \\ 1 & a_1(x_2) & \ldots & a_{k-1}(x_2) \\ \ldots & \ldots & \ldots & \ldots \\ 1 & a_1(x_n) & \ldots & a_{k-1}(x_n) \end{pmatrix} \text{— регрессионная матрица размера } n \times k;$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{k-1} \end{pmatrix} \text{— вектор параметров,}$$

система (23) принимает вид

$$(A^\top A)\boldsymbol{\beta} = A^\top \boldsymbol{Y},$$

где $A^\top A = (a_{ij})$, $i, j = 0, 1, 2, \ldots, k-1$ — квадратная матрица k-го порядка. При условии, что $A^\top A$ — невырожденная матрица, решение системы (23) можно записать в виде

$$\tilde{\boldsymbol{\beta}} = (A^\top A)^{-1} A^\top \boldsymbol{Y}, \tag{24}$$

где $\tilde{\boldsymbol{\beta}} = \begin{pmatrix} \tilde{\beta}_0 \\ \tilde{\beta}_1 \\ \vdots \\ \tilde{\beta}_{k-1} \end{pmatrix}$ — вектор МНК-оценок параметров модели (21).

Пример 3. Измерение температуры корпуса работающего агрегата, производимое с интервалом 5 минут, дало следующие результаты:

t, мин	5	10	15	20	25
T, $^\circ$C	59,3	59,8	60,1	64,9	70,2

Считая, что зависимость межу этими переменными имеет вид $T = a + bt + ct^2$, найти оценки параметров a, b и c по методу наименьших квадратов.

◁ Предварительно преобразуем исходные данные по формулам

$$x = \frac{t-15}{5}, \quad y = 10\,(T-60)$$

и вычислим оценки параметров линейной модели

$$y = \beta_0 + \beta_1 x + \beta_2 x^2.$$

Так как в этом случае $a_1(x) = x$, $a_2(x) = x^2$, то система нормальных уравнений (23) имеет вид

$$\begin{aligned} &\beta_0 n + \beta_1 \sum x_i + \beta_2 \sum x_i^2 = \sum y_i, \\ &\beta_0 \sum x_i + \beta_1 \sum x_i^2 + \beta_2 \sum x_i^3 = \sum x_i y_i, \\ &\beta_0 \sum x_i^2 + \beta_1 \sum x_i^3 + \beta_2 \sum x_i^4 = \sum x_i^2 y_i. \end{aligned} \tag{25}$$

Для вычисления коэффициентов системы (25) составим таблицу 7.1.

Таблица 7.1

t	x	T	y	xy	x^2	yx^2	x^3	x^4
5	-2	59,3	-7	14	4	-28	-8	16
10	-1	59,8	-2	2	1	-2	-1	1
15	0	60,1	1	0	0	0	0	0
20	1	64,9	49	49	1	49	1	1
25	2	70,2	102	204	4	408	8	16
$\sum$	0	–	143	269	10	427	0	34

Система нормальных уравнений (25) такова:

$$\begin{aligned} 5\beta_0 + \qquad\quad 10\beta_2 &= 143, \\ 10\beta_1 \qquad\quad &= 269, \\ 10\beta_0 + \qquad\quad 34\beta_2 &= 427. \end{aligned}$$

Решая эту систему, получим

$$\tilde\beta_0 \approx 8{,}457, \quad \tilde\beta_1 = 26{,}9, \quad \tilde\beta_2 \approx 10{,}07;$$

таким образом, зависимость между y и x имеет вид

$$y = 8{,}457 + 26{,}9x + 10{,}07x^2.$$

Переход к исходным переменным дает

$$(T-60)\cdot 10 = 8{,}457 + 26{,}9\,\frac{t-15}{5} + 10{,}07\left(\frac{t-15}{5}\right)^2,$$

откуда получаем окончательно

$$T = 61{,}84 - 0{,}67t + 0{,}04t^2. \quad \triangleright$$

Считая, что зависимость между переменными x и $\overline{Y}$ имеет вид $y = \beta_0 + \beta_1 x + \beta_2 x^2$, в задачах 19.335–19.338 найти оценки параметров по следующим выборкам:

19.335.

x	0	2	4	6	8	10
y	5	−1	−0,5	1,5	4,5	8,5

19.336.

x	0,07	0,31	0,61	0,99	1,29	1,78	2,09
y	1,34	1,08	0,94	1,06	1,25	2,01	2,60

19.337.

x	26	30	34	38	42	46	50
y	3,94	4,60	5,67	6,93	8,25	7,73	10,55

19.338.

x	−2	−1	0	1	2
y	4,8	0,4	-3,4	0,8	3,2

В задачах 19.339–19.341 найти оценки параметров β_0 и β_1, считая, что зависимость между переменными x и Y имеет вид

$$y = \beta_0 + \frac{\beta_1}{x}.$$

19.339.

x	2	4	6	12
y	8	5,25	3,50	3,25

19.340.

x	5,67	4,45	3,84	3,74	3,73	2,18
y	6,8	8,5	10,5	10,2	6,8	11,8

19.341.

x	1	2	3	4	5	6	7	8	9	10
y	16,50	13,75	13,31	12,50	13,52	12,75	12,30	12,83	12,28	12,34

Пусть $(x_i,\ y_i)$, $i = 1, 2, \ldots, n$ — результаты наблюдений двух переменных x и Y. Записать матрицу A для следующих линейных моделей (задачи 19.342–19.344):

19.342. $y = \beta_0 + \beta_1 x + \beta_2 x^2$.

19.343. $y = \beta_0 + \beta_1 \sin \omega x + \beta_2 \cos \omega x$, где ω — заданная константа.

19.344. $y = \beta_0 + \beta_1 e^x$.

Как и в случае линейной регрессии, качество аппроксимации результатов наблюдений (x_i, y_i), $i = 1, 2, \ldots, n$, регрессионной моделью (21) определяется остаточной дисперсией

$$s^2 = \frac{Q_e}{n-k}, \tag{26}$$

где Q_e — остаточная сумма квадратов, равная

$$Q_e = \sum (y_i - \tilde{y}_i)^2 = \sum [y_i - \tilde{\beta}_0 - \tilde{\beta}_1 a_1(x_i) - \cdots - \tilde{\beta}_{k-1} a_{k-1}(x_i)]^2.$$

В практических вычислениях остаточную сумму квадратов вычисляют из тождества

$$Q_y = Q_R + Q_e$$

(ср. с (14)).

Величина Q_R, называемая суммой квадратов, обусловленной регрессией, вычисляется по формуле

$$Q_R = \beta^\top A^\top Y - n\overline{y}^2. \tag{27}$$

Если модель (21) *адекватна результатам наблюдений, то остаточная дисперсия является несмещенной оценкой дисперсии ошибок наблюдений* σ^2, *т.е.* $M[s^2] = \sigma^2$, *причем статистика* Q_e/σ^2 *имеет распределение* χ^2 *с* $n-k$ *степенями свободы. В этом случае можно проверить гипотезы о параметрах модели и найти доверительные интервалы для этих параметров.*

Для проверки гипотезы $H_0: \ \beta_1 = 0, \ \beta_2 = 0, \ldots, \beta_{k-1} = 0$ используют статистику

$$F = \frac{Q_R/(k-1)}{Q_e/(n-k)} = \frac{Q_R}{(k-1)\, s^2}. \tag{28}$$

Если гипотеза H_0 верна (в этом случае говорят, что модель (21) статистически незначима), то статистика (28) имеет распределение Фишера с $k-1$ и $n-k$ степенями свободы.

Оценка ковариационной матрицы МНК-оценок параметров

$$K = s^2 (A^\top A)^{-1}. \tag{29}$$

Границы *доверительных интервалов для параметров* вычисляются по формуле

$$\tilde{\beta}_j \pm t_{1-\alpha/2}(n-k)\, s\sqrt{a_{jj}}, \quad j = 0, 1, \ldots, k-1, \tag{30}$$

где a_{jj} — диагональный элемент матрицы $(A^\top A)^{-1}$, а α — заданный уровень значимости; *доверительный интервал для дисперсии ошибок наблюдений* определяется соотношением

$$\frac{(n-k)\,s^2}{\chi^2_{1-\alpha/2}(n-k)} < \sigma^2 < \frac{(n-k)\,s^2}{\chi^2_{\alpha/2}(n-k)}. \tag{31}$$

Пусть при различных значениях $x_1, x_2, \ldots, x_m$ переменной X получены повторные наблюдения переменной Y, причем при $x = x_i$ произведено n_i наблюдений Y : $y_{i_1}, y_{i_2}, \ldots, y_{i_{n_i}}$ и $\sum\limits_{i=1}^{m} n_i = n$, где n — объем всей выборки. В этом случае можно проверить адекватность модели результатам наблюдений. Для этого используется статистика (сравните с (20)):

$$\frac{Q_n/(m-k)}{Q_p/(n-m)} = F\,(m-k,\; n-m),$$

где

$$Q_p = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_i)^2, \quad \overline{y}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} y_{ij}, \quad i = 1, 2, \ldots, m,$$

а

$$Q_n = Q_e - Q_p.$$

Пример 4. Найти оценки параметров модели

$$y = \beta_0 + \beta_1 x + \beta_2 x^2$$

по следующим данным:

x	-3	-2	-1	0	1	2	3
y	-10	0	4	5	4	2	-2

Проверить значимость модели. Определить коэффициент детерминации, вычислить оценки дисперсии ошибок наблюдений и ковариационной матрицы МНК-оценок параметров модели. Найти доверительные интервалы для параметров модели и дисперсии ошибок наблюдений. Принять $\alpha = 0{,}05$.

◁ Для вычисления оценок параметров составим таблицу 7.2. Система нормальных уравнений (23) имеет вид

$$\begin{aligned} 7\beta_0 + \qquad\qquad 28\beta_2 &= 3,\\ 28\beta_1 \qquad\qquad &= 28,\\ 28\beta_0 + \qquad\qquad 196\beta_2 &= -92. \end{aligned} \tag{32}$$

Решая эту систему, получим

$$\tilde{\beta}_0 \approx 5{,}38, \quad \tilde{\beta}_1 = 1, \quad \tilde{\beta}_2 \approx -1{,}24.$$

Таблица 7.2

x	y	xy	x^2	x^2y	x^3	x^4
−3	−10	30	9	−90	−27	81
−2	0	0	4	0	−8	16
−1	4	−4	1	4	−1	1
0	5	0	0	0	0	0
1	4	4	1	4	1	1
2	2	4	4	8	8	16
3	−2	−6	9	−18	27	81
0	3	28	28	−92	0	196

Для проверки значимости модели последовательно находим $\overline{y} = 3/7 \approx$ $\approx 0{,}429$, $\sum y_i^2 = 165$, $A^\top \boldsymbol{Y} = (3;\, 28;\, -92)^\top$ — правая часть системы нормальных уравнений (32), $\tilde{\boldsymbol{\beta}}^\top A^\top \boldsymbol{Y} \approx 158{,}22$. Далее вычисляем

$$Q_R = \tilde{\boldsymbol{\beta}}^\top A^\top Y - n\overline{y}^2 = 158{,}22 - 7 \cdot 0{,}429^2 \approx 156{,}93,$$

$$Q_y = \sum y_i^2 - n\overline{y}^2 = 165 - 7 \cdot 0{,}429^2 \approx 163{,}71,$$

$$Q_e = Q_y - Q_R = 163{,}71 - 156{,}93 = 6{,}78.$$

Выборочное значение статистики (28) равно

$$F_в = \frac{156{,}93/(3-1)}{6{,}78/(7-3)} \approx 46{,}29.$$

Так как $F_{0,95}(2,\, 4) = 6{,}94$ (таблица П7), то гипотеза о незначимости модели отклоняется.

Коэффициент детерминации по формуле (17)

$$R^2 = \frac{156{,}93}{163{,}71} \approx 0{,}958.$$

Оценка дисперсии ошибок наблюдений определяется по формуле (26):

$$s^2 = \frac{6{,}78}{7-3} = 1{,}695, \quad s = 1{,}302.$$

Оценка ковариационной матрицы по формуле (29)

$$K = 1{,}695 \begin{pmatrix} 7 & 0 & 28 \\ 0 & 28 & 0 \\ 28 & 0 & 196 \end{pmatrix}^{-1} \approx$$

$$\approx 1{,}695 \begin{pmatrix} 0{,}333 & 0 & -0{,}048 \\ 0 & 0{,}036 & 0 \\ -0{,}048 & 0 & 0{,}012 \end{pmatrix} = \begin{pmatrix} 0{,}565 & 0 & -0{,}081 \\ 0 & 0{,}061 & 0 \\ -0{,}081 & 0 & 0{,}02 \end{pmatrix}.$$

Определим доверительные интервалы для параметров модели. По таблице П6 находим $t_{0,975}(4) = 2{,}776$. По формуле (30) границы доверительных интервалов

для $\beta_0: \ 5{,}38 \pm 2{,}776 \cdot 1{,}302 \cdot \sqrt{0{,}333}$;

для $\beta_1: \ 1 \pm 2{,}776 \cdot 1{,}302 \cdot \sqrt{0{,}036}$;

для $\beta_2: \ -1{,}24 \pm 2{,}776 \cdot 1{,}302 \cdot \sqrt{0{,}012}$

или

$$\beta_0 \in (3{,}29;\ 7{,}47), \quad \beta_1 \in (0{,}31;\ 1{,}69), \quad \beta_2 \in (-1{,}64;\ -0{,}84).$$

Найдем доверительный интервал для дисперсии ошибок наблюдений. По таблице П5 находим $\chi^2_{0,975}(4) = 11{,}1$, $\chi^2_{0,025}(4) = 0{,}48$. По формуле (31) доверительный интервал для σ^2 имеет вид

$$\left(\frac{(7-3)\cdot 1{,}695}{11{,}1};\ \frac{(7-3)\cdot 1{,}695}{0{,}48}\right) \quad \text{или} \quad (0{,}611;\ 14{,}008).$$

Большой разброс границ доверительных интервалов объясняется тем, что оценки параметров и дисперсии ошибок наблюдений определены по малому числу наблюдений ($n = 7$). ▷

19.345. Предполагается, что зависимость между переменными Y и x достаточно точно описывается функцией $y = \beta_0 + \beta_1 x + \beta_2 x^2$. Найти оценки параметров β_0, β_1 и β_2, а также оценку ковариационной матрицы этих оценок по следующей выборке:

x	-3	-2	-1	0	1	2	3
y	-5	-2	$-1{,}5$	-1	0	3	14

В задачах 19.346–19.348 по выборкам наблюдений требуется:

а) найти оценки параметров модели $y = \beta_0 + \beta_1 x + \beta_2 x^2$;

б) проверить значимость модели;

в) найти оценки дисперсии ошибок наблюдений и ковариационной матрицы;

г) определить доверительные интервалы для параметров в дисперсии ошибок наблюдений при заданном уровне значимости α.

Предполагается, что ошибки наблюдений не коррелированы и имеют нормальное распределение $N(0, \sigma)$.

19.346.

x	-3	-2	-1	0	1	2	3
y	6	0	-1	-1	1	5	12

$$\alpha = 0{,}10.$$

19.347.

x	-2	-1	0	1	2
y	3	0	3	6	9

$$\alpha = 0{,}05.$$

19.348.

x	-3	-2	-1	0	-1	2	3
y	-6	-4	-2	-1	1	1	0

$$\alpha = 0{,}10.$$

В задачах 19.349–19.352 предполагается, что результаты наблюдений достаточно точно описываются многочленом второго порядка. Найти МНК-оценки параметров этой модели и проверить адекватность модели результатам наблюдений. Можно ли использовать для представления данных линейную регрессию?

19.349.

x	0	0	0	1	2	2	3	3	4	4
y	22,8	21,9	22,1	24,5	26,0	26,1	26,8	27,3	28,2	28,5

x	5	6	6	6	7	8	8	9	10
y	28,9	30,0	30,3	29,8	30,4	31,4	31,5	31,8	33,1

$$\alpha = 0{,}5.$$

19.350.

x	1,0	5,5	3,0	1,0	2,5	5,5	4,0	3,0	5,5	2,0
y	0,2	2,3	0,6	0,3	0,4	2,2	1,1	0,5	2,4	0,3

x	2,5	4,5	5,0	1,5	3,5	6,0	6,5	4,0	3,5	2,0
y	0,5	1,2	1,7	0,3	0,8	2,7	3,3	1,0	0,7	0,4

$$\alpha = 0{,}10.$$

19.351. Использовать данные из задачи 19.333. Принять $\alpha = 0{,}05$. Построить график остатков.

19.352. Использовать данные из задачи 19.334. Принять $\alpha = 0{,}05$.

В задачах 19.353–19.356 найти оценки для параметров модели $y = \beta_0 + \beta_1 x + \beta_2 x^2$.

19.353.

x	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0
y	0,4	0,3	1,0	1,7	2,1	3,4	4,1	5,8

x	4,5	5,0	5,5	6,0	6,5	7,0	7,5	8,0
y	7,7	9,4	11,4	13,6	15,6	18,6	21,2	24,1

19.354.

x	0,4	0,8	1,2	1,6	2	2,4	2,8	3,2	3,6
y	0,43	0,94	1,91	3,01	4	4,56	6,45	8,59	11,15

x	4,0	4,4	4,8	5,2	5,6	6,0	6,4	6,8
y	13,88	16,93	20,47	24,15	28,29	32,61	37,41	42,39

19.355.

x	0,00	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45
y	25	26	4	7	6	13	30	26	32	40

x	0,50	0,55	0,60	0,65	0,70	0,75	0,80	0,85	0,90	0,95
y	32	21	11	5	16	3	21	22	19	32

19.356.

x	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0	5,5
y	0,22	0,23	0,31	0,43	0,56	0,82	1,06	1,25	1,72	2,28

x	6,0	6,5	7,0	7,5	8,0	8,5	9,0	9,5	10,0
y	2,67	3,26	3,72	4,32	5,11	5,98	6,64	7,02	8,32

В задачах 19.357, 19.358 найти оценки для параметров модели $y = \beta_0 + \beta_1 \ln x$.

19.357.

x	1	2	3	4	5	6
y	2,11	2,45	2,61	2,73	2,75	2,81

x	7	8	9	10	11	12
y	2,87	2,91	2,96	3,03	3,05	3,12

19.358.

x	0,3	0,6	0,9	1,2	1,5	1,8	2,1
y	4,39	4,75	4,98	5,11	5,12	5,18	5,28

x	2,4	2,7	3,0	3,3	3,6	3,9	4,2
y	5,36	5,45	5,52	5,53	5,57	5,63	5,64

В задачах 19.359–19.360 найти оценки для параметров модели $y = \beta_0 + \beta_1 e^{0,1x}$.

19.359.

x	1	2	3	4	5	6
y	0,10	0,21	0,43	0,51	0,62	0,81

x	7	8	9	10	11	12
y	1,01	1,23	1,47	1,53	1,75	2,25

19.360.

x	1	1,5	2,0	2,5	3,0	3,5	4,0
y	4,11	4,16	4,23	4,29	4,36	4,42	4,53

x	4,5	5,0	5,5	6,0	6,5	7,0
y	4,57	4,63	4,75	4,87	4,88	5,01

В задачах 19.361, 19.362 найти оценки для параметров модели $y = \beta_0 + \beta_1 \sin x + \beta_2 \cos x$.

19.361.

x	0,5	1	1,5	2,0	2,5	3,0
y	2,47	2,86	3,01	2,91	2,55	2,11

x	3,5	4,0	4,5	5,0	5,5	6,0
y	2,61	1,25	0,97	1,03	1,34	1,70

19.362.

x	8,0	8,5	9,0	9,5	10,0	10,5	11,0
y	2,83	2,31	1,95	2,05	2,21	2,58	3,10

x	11,5	12,0	12,5	13,0	13,5	14,0
y	3,25	3,75	4,15	3,78	3,43	3,07

3. Использование ортогональных систем функций. Система функций $\psi_0(x)$, $\psi_1(x)$, ..., $\psi_{k-1}(x)$ называется *ортогональной на множестве* x_1, x_2, ..., x_n, если

$$\sum_{i=1}^{n} \psi_m(x_i)\,\psi_l(x_i) = 0 \quad \text{при} \quad m \neq l;\ m,\, l = 0,\, 1,\, \ldots,\, k-1.$$

Пусть (x_i, y_i), $i = 1, 2, \ldots, n$ — результаты наблюдений переменных x и Y. Оценки параметров линейной модели

$$y = \beta_0\psi_0(x) + \beta_1\psi_1(x) + \cdots + \beta_{k-1}\psi_{k-1}(x),$$

определяемые по методу наименьших квадратов, при использовании ортогональной системы функций $\psi_j(x)$, $j = 0, 1, 2, \ldots, k-1$, вычисляются по формулам

$$\tilde{\beta}_j = \frac{\sum\limits_{i=1}^{n} y_i\psi_j(x_i)}{\sum\limits_{i=1}^{n} \psi_j^2(x_i)}, \qquad j = 0,\, 1,\, \ldots,\, k-1.$$

В случае, когда для представления данных используется *полиномиальная модель*, а значения x_1, x_2, ..., x_n переменной x одинаково отстоят одно от другого с шагом b, применяют *ортогональные на этом множестве многочлены Чебышева*:

$$\zeta_0(x) \equiv 1, \quad \zeta_1(x) = x - \frac{n+1}{2},$$

многочлены более высокого порядка определяются по рекуррентной формуле

$$\zeta_{l+1}(x) = \zeta_1(x)\,\zeta_l(x) - \frac{l^2(n^2 - l^2)}{4\,(4l^2 - l)}\,\zeta_{l-1}(x).$$

Конкретные вычисления удобно проводить, используя табулированные значения ортогональных многочленов $P_k(i) = \lambda_k\zeta_k(i)$, где i принимает значения 1, 2, ..., n. В таблице П9 приведены значения $P_k(i)$ при $k = 1, 2, \ldots, 5$, и $n = 8, 10, 12, 13$. Там же приведены значения $\sum\limits_{i=1}^{n} P_k^2(i)$, необходимые для определения оценок параметров и остаточной суммы квадратов, а также значения коэффициентов λ_k, которые нужны для записи модели в исходных переменных.

При вычислениях с использованием таблицы значений ортогональных многочленов модель имеет следующий вид:

$$y_i = \beta_0 P_0(i) + \beta_1 P_1(i) + \cdots + \beta_{k-1} P_{k-1}(i), \quad i = 1, 2, \ldots, n. \tag{33}$$

Для определения оценок параметров модели (33) предварительно преобразуют x_i по формуле

$$i = \frac{x_i - x_1}{b} + 1, \quad i = 1, 2, \ldots, n. \tag{34}$$

Оценки параметров вычисляем последовательно по формулам

$$\tilde{\beta}_0 = \frac{\sum y_i}{n} = \overline{y}, \quad \tilde{\beta}_l = \frac{\sum\limits_{i=1}^{n} y_i P_l(i)}{\sum\limits_{i=1}^{n} P_l^2(i)}, \quad l = 1, 2, \ldots, k-1, \tag{35}$$

сразу же определяя значения остаточной суммы квадратов из соотношения

$$Q_e(\tilde{\beta}_0, \tilde{\beta}_1, \ldots, \tilde{\beta}_l) = \\ = Q_y - \tilde{\beta}_1^2 \sum_{i=1}^{n} P_1^2(i) - \tilde{\beta}_2^2 \sum_{i=1}^{n} P_2^2(i) - \cdots - \tilde{\beta}_l^2 \sum_{i=1}^{n} P_l^2(i). \tag{36}$$

Наилучший порядок модели определяется следующим образом. На каждом шаге находится оценка дисперсии ошибок наблюдений:

$$s^2(l) = \frac{Q_e(\tilde{\beta}_0, \tilde{\beta}_1, \ldots, \tilde{\beta}_l)}{n - (l+1)}, \quad l = 1, 2, \ldots \tag{37}$$

Это значение сравнивается со значением $s^2(l-1)$ оценки дисперсии ошибок наблюдений на предыдущем шаге. Если $s^2(l-1) > s^2(l)$, причем $\dfrac{s^2(l-1)}{s^2(l)} > F_{1-\alpha}(n-l,\, n-l-1)$, то различие дисперсий $s^2(l-1)$ и $s^2(l)$ ошибок наблюдений значимо и порядок модели следует увеличить на единицу. В случае, если различие дисперсий $s^2(l-1)$ и $s^2(l)$ ошибок наблюдений незначимо, нужно остановиться на модели $(l-1)$-го порядка. Проверку значимости с помощью предыдущего соотношения можно производить, если ошибки наблюдения ε_i не коррелированы, имеют равные дисперсии и нормальное распределение.

Пример 5. Подобрать порядок и найти оценки параметров полиномиальной модели (33) для представления данных

x	1,2	1,4	1,6	1,8	2,0	2,2	2,4	2,6	2,8	3,0
y	0,11	0,26	0,16	0,28	1,27	2,32	2,68	3,85	4,63	7,2

Предполагается, что ошибки наблюдений не коррелированы, имеют равные дисперсии и нормальное распределение. Принять $\alpha = 0{,}05$.

◁ Для вычислений используем модель (33). Значения x_i преобразуем по формуле (34): $i = \dfrac{x_i - x_1}{0{,}2} + 1$, $i = 1, 2, \ldots, 10$. В таблице 7.6

Таблица 7.6

i	y_i	$P_1(i)$	$P_2(i)$	$P_3(i)$
1	0,11	−9	6	−42
2	0,26	−7	2	+14
3	0,16	−5	−1	+35
4	0,28	−3	−3	+31
5	1,27	−1	−4	+12
6	2,32	+1	−4	−12
7	2,68	+3	−3	−31
8	3,85	+5	−1	−35
9	4,63	+7	2	−14
10	7,20	+9	6	+42
$\sum\limits_{i=1}^{n} P_k^2(i)$	–	330	132	8580
λ_k	–	2	1/2	5/3

выписаны значения переменных i и y и первых трех ортогональных многочленов при $n = 10$ (см. таблицу П9).

Предварительно вычисляем $\overline{y} = 2{,}276$, $Q_y = 50{,}659$. Используя значения $P_1(i)$ в третьем столбце таблицы 7.6, получим $\sum\limits_{i=1}^{n} y_i P_1(i) = 121{,}1$.

Следовательно, по формуле (35) $\tilde{\beta}_1 = \dfrac{121{,}1}{330} = 0{,}367$, а значение остаточной суммы квадратов и оценка дисперсии по формулам (36) и (37) таковы:

$$Q_e(\tilde{\beta}_0, \tilde{\beta}_1) = 50{,}659 - 0{,}367^2 \cdot 330 = 6{,}219;$$

$$s^2(1) = \frac{6{,}219}{10-2} = 0{,}777.$$

Далее, используя значения $P_2(i)$ в четвертом столбце, вычислим $\sum_{i=1}^{n} y_i P_2(i) = 26{,}39$. Находим

$$\tilde{\beta}_2 = \frac{26{,}39}{132} \approx 0{,}200; \quad Q_e(\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\beta}_2) = 0{,}853, \quad s^2(2) = \frac{0{,}853}{10-3} = 0{,}122.$$

Так как $s^2(1) > s^2(2)$, причем $\dfrac{s^2(1)}{s^2(2)} = \dfrac{0{,}777}{0{,}122} = 6{,}372$, что больше, чем $F_{0,95}(8,\ 7) = 3{,}726$, увеличим степень многочлена на единицу. По значениям $P_3(i)$ в пятом столбце таблицы 7.6 находим $\sum_{i=1}^{n} y_i P_3(i) = 20{,}45$, следовательно,

$$\tilde{\beta}_3 = \frac{20{,}45}{8580} = 0{,}0024, \ Q_e(\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\beta}_2, \tilde{\beta}_3) = 0{,}804, \ s^2(3) = \frac{0{,}804}{10-4} = 0{,}134.$$

Так как $\dfrac{s^2(3)}{s^2(2)} = \dfrac{0{,}134}{0{,}122} = 1{,}099$ [5]), что меньше, чем $F_{0,95}(6,\ 7) = 5{,}119$, различие оценок дисперсии $s^2(2)$ и $s^2(3)$ незначимо и вычисления прекращаются. Таким образом, для представления данных получена модель 2-го порядка:

$$\tilde{y}_i = 2{,}276 + 0{,}367 P_1(i) + 0{,}2 P_2(i), \quad i = 1,\ 2,\ 3,\ \ldots,\ 10. \quad \triangleright$$

По выборкам, приведенным в задачах 19.363–19.365, найти оценки параметров и порядок полиномиальной модели. Предполагается, что ошибки наблюдений не коррелированы и имеют распределение $N(0,\ \sigma)$; уровень значимости α задается.

19.363.

x	60	70	80	90	100	110	120	130	140	150
y	0	4	18	13	18	24	29	35	42	50

$\alpha = 0{,}10.$

19.364.

x	0,25	0,50	0,75	1,00	1,25	1,50	1,75	2,00
y	0	0,06	0,18	0,40	0,39	0,67	0,54	1,40

$\alpha = 0{,}05.$

[5]) Здесь, как обычно при сравнении дисперсий, в числитель ставится бóльшая дисперсия.

19.365.

x	0,25	0,50	0,75	1,00	1,25	1,50
y	5,01	5,02	5,32	6,13	7,41	9,98

x	1,75	2,00	2,25	2,50	2,75	3,00
y	14,30	20,95	30,10	44,03	60,50	75,20

$$\alpha = 0{,}10.$$

19.366. Пусть $\psi_j(x)$, $j = 0, 1, \ldots, k-1$ — ортогональная система функций на множестве значений $x_1, x_2, \ldots, x_n$. Показать, что оценки параметров модели

$$y = \beta_0\psi_0(x) + \beta_1\psi_1(x) + \cdots + \beta_{k-1}\psi_{k-1}(x),$$

вычисляемые по результатам наблюдений (x_i, y_i), $i = 1, 2, \ldots, n$, не коррелированы. Определить дисперсии этих оценок.

19.367. Показать, что система функций $\psi_0(x) \equiv 1$, $\psi_1(x) = x - \overline{x}$ ортогональна на множестве значений $x_1, x_2, \ldots, x_n$.

19.368*. Система функций 1, $\cos x$, $\sin x$, $\cos 2x$, $\sin 2x$, ... , $\cos mx$, $\sin mx$ ортогональна на множестве значений $x_k = 2\pi k/n$, где $k = 0, 1, 2, \ldots, n-1$, $n \geqslant 2m+1$. Вычислить МНК-оценки параметров модели

$$y = \alpha_0 + \alpha_1\cos x + \beta_1\sin x + \cdots + \alpha_m\cos mx + \beta_m\sin mx$$

по результатам наблюдений (x_k, y_k) и найти дисперсии этих оценок.

Используя результаты задачи 19.368, найти оценки параметров модели указанного порядка m и вычислить дисперсии этих оценок по выборкам в задачах 19.369 и 19.370.

19.369.

x	0°	30°	60°	90°	120°	150°	180°
y	4,01	3,95	3,45	2,87	3,23	3,83	4,12

x	220°	240°	270°	300°	330°	360°
y	2,87	1,76	1,15	1,51	4,11	4,17

$$m = 1.$$

19.370.

x	0°	30°	60°	90°	120°	150°	180°
y	0,97	−0,65	−0,35	2,06	3,80	2,80	−0,98

x	210°	240°	270°	300°	330°	360°
y	−4,39	−4,71	−2,16	2,32	1,05	0,81

$$m = 1.$$

4. Некоторые нелинейные задачи, сводящиеся к линейным моделям. Во многих практически важных задачах зависимость между переменными Y и x нелинейна по параметрам. Однако часто можно найти преобразование переменных, которое приводит к линейной модели. Как правило, вычисление оценок параметров для линейной модели существенно упрощается. Тем не менее следует иметь в виду, что при вычислении оценок параметров по методу наименьших квадратов в этом случае минимизируется сумма квадратов отклонений преобразованных, а не исходных данных. Очевидно, что свойства полученных оценок и возможность их дальнейшего статистического анализа зависят от того, удовлетворяются ли условия, сформулированные в начале этого параграфа, для преобразованных переменных.

Пример 6. Зависимость между переменными z и x имеет вид $z = \dfrac{1}{\beta_0 + \beta_1 x}$. Преобразовать эту модель в линейную по параметрам.

◁ Запишем зависимость между z и x в таком виде:

$$\frac{1}{z} = \beta_0 + \beta_1 x.$$

Обозначим $\dfrac{1}{z} = y$; тогда получим $y = \beta_0 + \beta_1 x$. Полученная модель линейна по параметрам β_0 и β_1. ▷

В задачах 19.371–19.376 преобразовать каждую нелинейную модель в линейную по параметрам.

19.371. $y = e^{-\beta_1 x + \beta_0}$. **19.372.** $\dfrac{1}{y} = \beta_0 + \dfrac{1}{\beta_1 x}$.

19.373. $y = \beta_0 \ln x + \beta_1$. **19.374.** $y = \sin\beta_0 + \beta_1 \sin x$.

19.375. $y = \beta_0 x^{\beta_1}$. **19.376.** $y = \dfrac{x}{\beta_0 + \beta_1 x}$.

Для нахождения оценок параметров в задачах 19.377–19.381 использовать метод наименьших квадратов, предварительно выполнив преобразование нелинейной модели в линейную по параметрам.

19.377. Барометрическое давление связано с высотой следующим соотношением:

$$\frac{p}{p_0} = \exp\left\{-\frac{k}{T}\,z\right\},$$

где p — барометрическое давление на высоте z, T — температура, а p_0 и k — параметры. Найти оценки параметров k/T и p_0 по результатам 6 наблюдений, проведенных при приблизительно постоянной температуре:

z, м	100	1100	1200	1400	1500	1600
ρ, мм рт. ст.	640	595	504	363	310	267

19.378. Для исследования зависимости давления p насыщенного пара (Н/см^2) от удельного объема v (м^3/кг) составлена таблица опытных данных:

v	3,334	1,630	0,866	0,423	0,265	0,170	0,115
p	0,482	1,034	2,027	4,247	7,164	11,480	17,60

Считая, что функциональная зависимость между этими переменными имеет вид

$$p = \beta v^{\alpha},$$

найти оценки параметров α и β.

19.379. Функциональная зависимость удельного сопротивления кристаллического кварца ρ (Ом · см) от абсолютной температуры T (К) имеет вид

$$\rho = 10^{(a/T)+b}.$$

Используя опытные данные

ρ	$5\cdot10^{16}$	$4\cdot10^{15}$	$3\cdot10^{14}$	$2\cdot10^{13}$	$2\cdot10^{12}$	$1{,}5\cdot10^{11}$	$1\cdot10^{10}$
T	335	365	400	445	500	570	670

найти оценки параметров a и b.

19.380. Исследование зависимости продолжительности t решения систем линейных уравнений от порядка системы n дало следующие результаты:

n	2	3	4	5	6	7	8	9	10
t, мин	12	35	75	130	210	315	445	600	800

Предполагая, что $t = An^{\gamma}$, найти оценки параметров A и γ.

19.381. Получена выборка наблюдений переменных x и y:

x	1	2	3	4	5	6	7
y	62,1	87,2	109,3	127,3	134,7	136,2	136,9

Для представления этих данных предлагается использовать модель

$$y = \frac{x}{\beta_0 + \beta_1 x}.$$

Найти оценки параметров β_0 и β_1.

5. Множественная линейная регрессия (случай двух независимых переменных). Предположим, что зависимость между переменными имеет вид

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2, \tag{38}$$

где переменные x_1 и x_2 принимают заданные фиксированные значения, причем между переменными x_1 и x_2 нет линейной зависимости. Результаты наблюдения (x_{1i}, x_{2i}, y_i), $i = 1, 2, \ldots, n$, представляются в виде

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i.$$

Оценки параметров модели (38) могут быть найдены по формуле $\tilde{\boldsymbol{\beta}} = (A^\top A)^{-1} A^\top \boldsymbol{Y}$, однако более удобно находить оценки для модели

$$y - \overline{y} = \beta_1(x_1 - \overline{x}_1) + \beta_2(x_2 - \overline{x}_2).$$

Последовательность вычислений в этом случае следующая. Сначала для каждой переменной находят суммы наблюдаемых значений, суммы квадратов, а также суммы попарных произведений. Затем вычисляют Q_y, Q_{x_1}, Q_{x_2}, Q_{x_1y},Q_{x_2y}, $Q_{x_1x_2}$, используя формулы

$$Q_y = \sum y_i^2 - \frac{\left(\sum y_i\right)^2}{n}, \quad Q_{x_j} = \sum x_{ji}^2 - \frac{\left(\sum x_{ji}^2\right)^2}{n}.$$

$$Q_{x_jy} = \sum x_{ji}y_i - \frac{\left(\sum x_{ji}\right)\left(\sum y_i\right)}{n},$$

$$Q_{x_1x_2} = \sum x_{1i}x_{2i} - \frac{\left(\sum x_{1i}\right)\left(\sum x_{2i}\right)}{n}, \quad j = 1, 2.$$

Далее вычисляют матрицы

$$A^\top A = \begin{pmatrix} Q_{x_1} & Q_{x_1x_2} \\ Q_{x_1x_2} & Q_{x_2} \end{pmatrix}, \quad (A^\top A)^{-1} = \begin{pmatrix} Q_{x_2} & -Q_{x_1x_2} \\ -Q_{x_1x_2} & Q_{x_1} \end{pmatrix} \frac{1}{|A^\top A|},$$

$$A^\top \boldsymbol{Y} = \begin{pmatrix} Q_{x_1y} \\ Q_{x_2y} \end{pmatrix}.$$

Чтобы исключить потерю точности при использовании элементов матрицы $(A^\top A)^{-1}$, деление этой матрицы на определитель $|A^\top A|$ следует выполнять в последнюю очередь. Вектор оценок определяется по формуле

$$\tilde{\boldsymbol{\beta}} = \begin{pmatrix} \tilde{\beta}_1 \\ \tilde{\beta}_2 \end{pmatrix} = (A^\top A)^{-1} A^\top \boldsymbol{Y} = \begin{pmatrix} Q_{x_2} & -Q_{x_1 x_2} \\ -Q_{x_1 x_2} & Q_{x_1} \end{pmatrix} \begin{pmatrix} Q_{x_1 y} \\ Q_{x_2 y} \end{pmatrix} \frac{1}{|A^\top A|}.$$

Оценка параметра β_0 вычисляется по формуле

$$\tilde{\beta}_0 = \overline{y} - \tilde{\beta}_1 \overline{x}_1 - \tilde{\beta}_2 \overline{x}_2. \tag{39}$$

Остаточная сумма квадратов Q_e вычисляется по формуле

$$Q_e = Q_y - \tilde{\boldsymbol{\beta}}^\top A^\top \boldsymbol{Y}.$$

Предположим, что ошибки наблюдений ε_i независимы, имеют равные дисперсии и нормально распределены. В этом случае можно проверить гипотезу H_0 : $\beta_1 = \beta_2 = 0$. Эта гипотеза позволяет установить, находятся ли переменные x_1 и x_2 во взаимосвязи с Y. Статистикой критерия для проверки гипотезы H_0 является отношение

$$F = \frac{\tilde{\boldsymbol{\beta}}^\top A^\top \boldsymbol{Y}/2}{Q_e/(n-3)}. \tag{40}$$

Если выборочное значение этой статистики $F_в > F_{1-\alpha}(2,\, n-3)$, то гипотеза H_0 отклоняется; в противном случае следует считать, что взаимосвязи Y с переменными x_1 и x_2 нет.

Границы доверительных интервалов для параметров β_1 и β_2 определяются по формуле

$$\tilde{\beta}_j \pm t_{1-a/2}(n-3)\, s \sqrt{a_{jj}}\,, \quad j = 1,\, 2, \tag{41}$$

где a_{jj} — диагональный элемент матрицы $(A^\top A)^{-1}$, $s = \sqrt{\dfrac{Q_e}{n-3}}$.

При использовании модели (38) для представления данных необходимо решить вопрос о целесообразности включения переменной x_1 или x_2 в модель. Для этого проверяются гипотезы $H_0^{(j)}$: $\beta_j = 0$, $j = 1,\, 2$. Очевидно, эти гипотезы могут быть проверены непосредственно по доверительным интервалам для параметров β_1 и β_2: если доверительный интервал для β_j, $j = 1,\, 2$, накрывает нуль, то гипотеза $H_0^{(j)}$: $\beta_j = 0$ принимается. В противном случае $H_0^{(j)}$ отклоняется.

Коэффициент множественной корреляции, характеризующий отклонение результатов наблюдений от плоскости регрессии $y = \tilde{\beta}_0 + \tilde{\beta}_1 x_1 + \tilde{\beta}_2 x_2$, определяется по формуле

$$R = \sqrt{\frac{\tilde{\boldsymbol{\beta}}^\top A^\top \boldsymbol{Y}}{Q_y}}\,. \tag{42}$$

Пример 7. Температура объекта Y зависит от процентного содержания x_1 компоненты A в теплоносителе и температуры окружающей среды x_2. Ниже приведены результаты 11 замеров этих данных.

y, °C	6	8	1	0	5	3	2	−4	10	−3	5
x_1, %	1	4	9	11	3	8	5	10	2	7	6
x_2, °C	8	2	−8	−10	6	−6	0	−12	4	−2	−4

Используя эту выборку, выполнить следующие задания:

а) найти оценки и доверительные интервалы параметров модели

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2.$$

б) проверить взаимосвязь Y с переменными x_1 и x_2;

в) проверить гипотезы $H_0^{(j)}: \beta_j = 0,\ j = 1, 2$;

г) вычислить коэффициент множественной корреляции.

Предполагается,что ошибки наблюдений не коррелированы, имеют равные дисперсии и распределены по нормальному закону. Принять уровень значимости $\alpha = 0{,}10$.

◁ а) Предварительно вычислим

$$\sum x_{1i} = 66, \quad \sum x_{2i} = -22, \quad \sum y_i = 33, \quad \sum x_{1i}^2 = 506,$$

$$\sum x_{2i}^2 = 484, \quad \sum y_i^2 = 289, \quad \sum x_{2i}x_{1i} = -346,$$

$$\sum y_i x_{1i} = 85, \quad \sum y_i x_{2i} = 142, \quad \overline{y} = 3, \quad \overline{x}_1 = 6, \quad \overline{x}_2 = -2.$$

Далее находим

$$Q_{x_1} = 506 - \frac{66^2}{11} = 110, \qquad Q_{x_2} = 484 - \frac{(-22)^2}{11} = 440,$$

$$Q_y = 289 - \frac{33^2}{11} = 190, \qquad Q_{x_1x_2} = -346 - \frac{66\cdot(-22)}{11} = -214,$$

$$Q_{x_1y} = 85 - \frac{33\cdot 66}{11} = -113, \quad Q_{x_2y} = 142 - \frac{33\cdot(-22)}{11} = 208.$$

Матрица $A^\top A = \begin{pmatrix} 110 & -214 \\ -214 & 440 \end{pmatrix}$, $|A^\top A| = 2604$. Найдем матрицу $(A^\top A)^{-1}$; имеем

$$(A^\top A)^{-1} = \frac{1}{2604}\begin{pmatrix} 440 & 214 \\ 214 & 110 \end{pmatrix} \approx \begin{pmatrix} 0{,}169 & 0{,}082 \\ 0{,}082 & 0{,}042 \end{pmatrix}.$$

Так как $A^{\top}Y = \begin{pmatrix} -113 \\ 208 \end{pmatrix}$, вектор оценок равен

$$\tilde{\beta} = \begin{pmatrix} \tilde{\beta}_1 \\ \tilde{\beta}_2 \end{pmatrix} \approx \begin{pmatrix} 0{,}169 & 0{,}082 \\ 0{,}082 & 0{,}042 \end{pmatrix} \cdot \begin{pmatrix} -113 \\ 208 \end{pmatrix} \approx \begin{pmatrix} -2{,}041 \\ -0{,}530 \end{pmatrix}.$$

Найдем оценку параметра β_0. По формуле (38) имеем

$$\tilde{\beta}_0 = 3 - (-2{,}041) \cdot 6 - (-0{,}530) \cdot (-2) = 14{,}186.$$

Таким образом, уравнение плоскости регрессии имеет вид

$$y = 14{,}186 - 2{,}041x_1 - 0{,}53x_2.$$

Определим доверительные интервалы для параметров β_1 и β_2. Для этого находим остаточную сумму квадратов:

$$Q_e(\tilde{\beta}_1, \tilde{\beta}_2) = 190 - (-2{,}041\ -0{,}530) \cdot \begin{pmatrix} -113 \\ 208 \end{pmatrix} = 69{,}607.$$

Далее вычисляем оценку дисперсии ошибок наблюдений:

$$s^2 = \frac{Q_e}{n-3} = \frac{69{,}607}{11-3} \approx 8{,}7, \quad s \approx 2{,}95.$$

По таблице П6 находим $t_{0,95}(8) = 1{,}860$. Доверительные интервалы определяем по формуле (40). Для параметра β_1 имеем

$$-2{,}041 \pm 1{,}86 \cdot 2{,}95 \cdot \sqrt{0{,}169} \approx -2{,}041 \pm 0{,}927,$$

следовательно, параметр β_1 накрывается интервалом $(-2{,}968; -1{,}114)$. Для параметра β_2 имеем

$$-0{,}53 \pm 1{,}86 \cdot 2{,}95 \cdot \sqrt{0{,}042} \approx -0{,}53 \pm 1{,}124.$$

Следовательно, параметр β_2 накрывается интервалом $(-1{,}654;\ 0{,}594)$.

б) Для проверки гипотезы $H_0:\ \beta_1 = \beta_2 = 0$ вычислим выборочное значение статистики F. Так как $\tilde{\boldsymbol{\beta}}^{\top} A^{\top}\boldsymbol{Y} = (-2{,}041\ -0{,}53) \begin{pmatrix} -113 \\ 208 \end{pmatrix} =$ $= 120{,}393$, по формуле (39) имеем

$$F_в = \frac{120{,}393/2}{69{,}607/(11-3)} \approx 6{,}918;$$

так как это значение больше $F_{0,90}(2{,}8) = 3{,}113$ (таблица П7), гипотеза H_0 отклоняется.

в) Проверим значимость переменных x_1 и x_2. Доверительный интервал для параметра β_1 не накрывает нуль, следовательно, переменная x_1 значима. Доверительный интервал для параметра β_2 накрывает нуль, следовательно, переменная x_2 может быть исключена из рассмотрения.

г) Коэффициент множественной корреляции вычисляется по формуле (42):

$$R = \sqrt{\frac{120{,}393}{190}} \approx 0{,}796. \quad \triangleright$$

Для приведенных ниже данных (задачи 19.382–19.385) выполнить следующие задания:

а) найти уравнение плоскости регрессии и доверительные интервалы для параметров β_1 и β_2;

б) проверить гипотезу: $H_0 : \beta_1 = \beta_2 = 0$;

в) проверить гипотезы $H_0^{(j)} : \beta_j = 0, \ j = 1, 2$;

г) вычислить коэффициент множественной корреляции. Предполагается, что ошибки наблюдений не коррелированы, имеют равные дисперсии и распределены по нормальному закону. Уровень значимости α задается.

19.382.

x_1	1	−1	2	1	−1	−4	7	0	8	3	6	−2
x_2	−1	1	−2	−6	−8	5	3	−3	0	−10	2	7
y	2	0	1	−4	−8	4	11	−2	9	8	10	5

$$\sum x_{1i} = 20, \quad \sum x_{2i} = -12, \quad \sum y_i = 36, \quad \sum x_{1i}^2 = 186,$$
$$\sum x_{2i}^2 = 302, \quad \sum y_i^2 = 496, \quad \sum x_{1i}x_{2i} = -35, \quad \sum x_{1i}y_i = 215,$$
$$\sum x_{2i}y_i = 118; \quad \alpha = 0{,}10.$$

19.383.

x_1	1	4	0	5	−3	3	−5	−1	2	−2
x_2	4	−6	2	−4	12	−2	14	6	0	8
y	−4	−5	4	−1	4	0	5	1	2	7

$$\sum x_{1i} = 4, \quad \sum x_{2i} = 34, \quad \sum y_i = 13, \quad \sum x_{1i}^2 = 94,$$
$$\sum x_{2i}^2 = 516, \quad \sum y_i^2 = 4, \sum x_{1i}x_{2i} = -174,$$
$$\sum x_{1i}y_i = -77, \quad \sum x_{2i}y_i = 216; \quad \alpha = 0{,}10.$$

19.384.

x_1	31	34	35	41	38	32	29	34
x_2	29,5	14,2	18,0	21,3	47,5	10,0	21,0	36,5
y	22,0	14,0	23,0	43,0	66,0	7,6	12,0	36,0

$$\sum x_{1i} = 274, \quad \sum x_{2i} = 198, \quad \sum y_i = 223{,}6, \quad \sum x_{1i}^2 = 9488,$$

$$\sum x_{2i}^2 = 5979{,}08, \quad \sum y_i^2 = 8911{,}76, \quad \sum x_{1i}x_{2i} = 6875{,}6,$$

$$\sum x_{1i}y_i = 8049{,}2, \quad \sum x_{2i}y_i = 6954{,}7; \quad \alpha = 0{,}05.$$

19.385.

x_1	0	44	4	61	35	64	13	56	18	2
x_2	14	0	29	34	54	16	44	59	49	32
y	0,5	47,2	8	63,8	18,2	47,5	0	60,9	19,2	9

$$\sum x_{1i} = 297, \quad \sum x_{2i} = 331, \quad \sum y_i = 274{,}3, \quad \sum x_{1i}^2 = 14\,627,$$

$$\sum x_{2i}^2 = 14\,207, \quad \sum y_i^2 = 13\,108{,}47, \quad \sum x_{1i}x_{2i} = 9926,$$

$$\sum x_{1i}y_i = 13\,451{,}6, \quad \sum x_{2i}y_i = 8972{,}9; \quad \alpha = 0{,}05.$$

19.386. Записать систему нормальных уравнений для нахождения оценок параметров плоскости регрессии (38) по результатам наблюдений (x_{1i}, x_{2i}, y_i), $i = 1, 2, \ldots, n$.

6. Вычисление и статистический анализ оценок параметров линейной модели при коррелированных и неравноточных наблюдениях. В некоторых случаях ошибки наблюдений ε_i, $i = 1, 2, \ldots, n$, случайной зависимой переменной Y *имеют различные дисперсии* и *коррелированы*. Так случается, когда измерения проводятся при помощи приборов, имеющих разную точность, либо когда ошибки измерения зависят от значений независимой переменной x и т.д. В этих случаях вычисление и статистический анализ МНК-оценок параметров проводится следующим образом.

Предположим, что ошибки наблюдений ε_i, $i = 1, 2, \ldots, n$, имеют нулевые математические ожидания и ковариационную матрицу $\sigma^2 W$, где W — известная симметрическая положительно определенная матрица [6]). Система нормальных уравнений (23) для нахождения оценок параметров модели (21) записывается в виде

$$A^{\top}W^{-1}A\boldsymbol{\beta} = A^{\top}W^{-1}\boldsymbol{Y}.$$

Вектор оценок параметров вычисляется по формуле

$$\tilde{\boldsymbol{\beta}} = (A^{\top}W^{-1}A)^{-1}A^{\top}W^{-1}\boldsymbol{Y}.$$

Остаточная сумма квадратов находится из соотношения

$$Q_e = \boldsymbol{Y}^{\top}W^{-1}\boldsymbol{Y} - \tilde{\boldsymbol{\beta}}^{\top}A^{\top}W^{-1}\boldsymbol{Y}. \tag{43}$$

Ковариационная матрица оценок вычисляется по формуле

$$K = \sigma^2(A^{\top}W^{-1}A)^{-1}. \tag{44}$$

Обычно величина σ^2 неизвестна. Оценка этой величины по результатам наблюдений определяется по формуле

$$\tilde{\sigma}^2 = \frac{Q_e}{n-k}, \tag{45}$$

где k — число оцениваемых параметров.

В случае, когда ошибки наблюдений ε_i *не коррелированы*, матрица W имеет диагональный вид:

$$W = \begin{pmatrix} \frac{1}{w_1} & & & 0 \\ & \frac{1}{w_2} & & \\ & & \ddots & \\ 0 & & & \frac{1}{w_n} \end{pmatrix},$$

где w_i, $i = 1, 2, \ldots, n$ — заданные *«веса» дисперсий ошибок наблюдений*. В этом случае обратная матрица W^{-1} имеет вид

$$W^{-1} = \begin{pmatrix} w_1 & & & 0 \\ & w_2 & & \\ & & \ddots & \\ 0 & & & w_n \end{pmatrix},$$

[6]) Симметрическая матрица положительно определена тогда и только тогда, когда все ее главные миноры положительны (критерий Сильвестра).

и, в частности, для модели $y = \beta_0 + \beta_1 x + \beta_2 x^2$ получим систему нормальных уравнений в виде (см. (7))

$$\begin{aligned}
&\beta_0 \sum w_i + \beta_1 \sum w_i x_i + \beta_2 \sum w_i x_i^2 = \sum w_i y_i, \\
&\beta_0 \sum w_i x_i + \beta_1 \sum w_i x_i^2 + \beta_2 \sum w_i x_i^3 = \sum w_i x_i y_i, \\
&\beta_0 \sum w_i x_i^2 + \beta_1 \sum w_i x_i^3 + \beta_2 \sum w_i x_i^4 = \sum w_i x_i^2 y_i.
\end{aligned} \tag{46}$$

Статистический анализ оценок параметров проводится аналогично тому, как это было сделано в п. 2 настоящего параграфа.

Пример 8. Предположим, что в условиях примера 4 веса дисперсий ошибок наблюдений распределены следующим образом: $w_1 = 2$; $w_2 = 1{,}5$; $w_3 = w_4 = w_5 = 1$; $w_0 = 1{,}5$; $w_7 = 2$, т.е. ошибки наблюдений не коррелированы, но их дисперсии зависят от модуля значения независимой переменной x. При условии, что ошибки наблюдений ε_i имеют нормальное распределение $N\left(0, \dfrac{\sigma}{\sqrt{w_i}}\right)$, $i = 1, 2, \ldots, n$, найти оценки параметров модели $y = \beta_0 + \beta_1 x + \beta_2 x^2$, оценку ковариационной матрицы ошибок наблюдений, а также доверительные интервалы для параметров. Принять $\alpha = 0{,}05$.

◁ Используя данные и результаты решения примера 4, вычислим коэффициенты системы (46):

$$\sum w_i x_i = 0, \quad \sum w_i x_i^2 = 50, \quad \sum w_i x_i^3 = 0, \quad \sum w_i x_i^4 = 374,$$
$$\sum w_i y_i = -8, \quad \sum w_i x_i y_i = 54, \quad \sum w_i x_i^2 y_i = -196, \quad \sum w_i = 10.$$

Система нормальных уравнений (46) имеет вид

$$\begin{aligned}
10\beta_0 \quad + \quad 50\beta_2 &= -8, \\
50\beta_1 \qquad\quad &= 54, \\
50\beta_0 \quad + \quad 374\beta_2 &= -196.
\end{aligned} \tag{47}$$

Решая эту систему, получим

$$\tilde\beta_0 = 5{,}49, \quad \tilde\beta_1 = 1{,}08, \quad \tilde\beta_2 = -1{,}258.$$

Найдем остаточную сумму квадратов. Предварительно вычислим

$$\boldsymbol{Y}^\top W^{-1} \boldsymbol{Y} = \sum w_i y_i^2 = 271.$$

Так как правая часть нормальной системы (47) равна $\begin{pmatrix} -8 \\ 54 \\ -196 \end{pmatrix}$, по формуле (43) находим

$$Q_e = 271 - (5{,}49 \cdot 1{,}08 - 1{,}258) \begin{pmatrix} -8 \\ 54 \\ -196 \end{pmatrix} = 10{,}032.$$

Оценка ковариационной матрицы по формулам (44) и (45) равна

$$K = \frac{10{,}032}{7-3}\begin{pmatrix} 10 & 0 & 50 \\ 0 & 50 & 0 \\ 50 & 0 & 374 \end{pmatrix}^{-1} \approx$$

$$\approx 2{,}508\begin{pmatrix} 0{,}303 & 0 & -0{,}040 \\ 0 & 0{,}02 & 0 \\ -0{,}040 & 0 & 8{,}06\cdot 10^{-3} \end{pmatrix} \approx \begin{pmatrix} 0{,}756 & 0 & -0{,}101 \\ 0 & 0{,}5 & 0 \\ -0{,}101 & 0 & 0{,}02 \end{pmatrix}.$$

Определим доверительные интервалы для параметров модели. Так как $t_{0,975}(4) = 2{,}776$, то по формуле (30) получим границы доверительных интервалов:

$$\text{для } \beta_0:\ 5{,}49 \pm 2{,}776\cdot\sqrt{2{,}508}\cdot\sqrt{0{,}303};$$

$$\text{для } \beta_1:\ 1{,}08 \pm 2{,}776\cdot\sqrt{2{,}508}\cdot\sqrt{0{,}02};$$

$$\text{для } \beta_2:\ -1{,}258 \pm 2{,}776\cdot\sqrt{2{,}508}\cdot\sqrt{8{,}06\cdot 10^{-3}}$$

или

$$\beta_0 \in (3{,}08;\ 7{,}90),\quad \beta_1 \in (0{,}46;\ 1{,}70),\quad \beta_2 \in (-1{,}65;\ -0{,}87).\quad \triangleright$$

19.387. Решить задачу 19.345, если ошибки наблюдений не коррелированы, а веса дисперсий ошибок наблюдений заданы следующим образом:

$$w_1 = w_7 = 2;\quad w_2 = w_6 = 1{,}5;\quad w_3 = w_4 = w_5 = 1.$$

Найти оценку дисперсии результатов наблюдений.

19.388. Масса тела m определяется по результатам n независимых взвешиваний $m_1, m_2, \ldots, m_n$, причем результат i-го взвешивания имеет дисперсию, равную σ^2/w_i, где w_i — заданные значения, $i = 1, 2, \ldots, n$. Найти оценку массы тела по этим результатам. Определить дисперсию полученной оценки.

19.389. Пусть для оценки параметра β модели $y = \beta x$ используется выборка (x_i, y_i), $i = 1, 2, \ldots, n$. Предположим, что дисперсия наблюдаемой случайной величины Y пропорциональна x. Найти оценку параметра β и дисперсию этой оценки.

19.390. Решить задачу 19.389, если дисперсия Y пропорциональна x^2.

19.391. Используя результаты, полученные при решении задачи 19.389 и 19.390, найти оценку параметра β модели $y = \beta x$ по выборке

x_i	1	2	3	4	5	6	7	8	9	10
y_i	0,5	0,9	0,4	2,2	0,6	1,1	0,8	2,4	1,2	4,3

если

а) результаты наблюдений имеют равные дисперсии;

б) дисперсия результатов наблюдений пропорциональна x;

в) дисперсия результатов наблюдений пропорциональна x^2.

Вычислить дисперсии оценки параметра, если коэффициент пропорциональности в б) и в) равен единице.

§ 8. Непараметрические методы математической статистики

1. Основные понятия. Критерий знаков. В практике обработки результатов наблюдений распределение генеральной совокупности часто неизвестно либо (для непрерывных случайных величин) отличается от нормального распределения, так что применение методов из § 3–5 настоящей главы не обосновано и может привести к ошибкам. В этих случаях применяют *методы, не зависящие* (*или свободные*) *от распределения* генеральной совокупности, называемые также *непараметрическими методами.*

Непараметрические методы используют не сами численные значения элементов выборки, а *структурные свойства выборки* (например, отношения порядка между ее элементами). В связи с этим теряется часть информации, содержащаяся в выборке, поэтому, например, мощность непараметрических критериев меньше, чем мощность их аналогов из § 4. Однако непараметрические методы могут применяться при более общих предположениях и более просты с точки зрения выполнения вычислений.

Большая группа непараметрических критериев используется для проверки гипотезы о принадлежности двух выборок $x_1, x_2, \ldots, x_{n_1}$ и $y_1, y_2, \ldots, y_{n_2}$ одной и той же генеральной совокупности, т.е. о том, что функции распределения двух генеральных совокупностей $F_X(x)$ и $F_Y(y)$ равны: $F_X(x) \equiv F_Y(y)\big|_{y=x}$. Такие генеральные совокупности называют *однородными.* Необходимое условие однородности состоит в равенстве характеристик положения и (или) рассеивания у рассматриваемых генеральных совокупностей — таких, как средние, медианы, дисперсии и др. Используемые для этих целей непараметрические критерии в качестве основного предположения используют только непрерывность распределения генеральной совокупности.

Простейший критерий такого рода, *критерий знаков*, применяется для проверки гипотезы H_0 об однородности генеральных совокупностей по попарно связанным выборкам. Такая задача возникает, например, при сравнении двух измерительных приборов. При этом используют n объектов и над каждым из них производят по одному измерению с помощью обоих приборов. Обозначим x_i и y_i, $i = 1, 2, \ldots, n$, результаты измерения i-го объекта, полученные соответственно при помощи первого и второго приборов. Если сравниваемые выборки получены из однородных совокупностей, то значения x_i и y_i взаимосвязаны, и, следовательно, вероятности появления положительных и отрицательных разностей $x_i - y_i$ равны. Вероятности появления нулевых разностей равны нулю в силу предполагаемой непрерывности распределения измеряемого признака. Таким образом, вероятности появления положительных и отрицательных разностей равны 1/2, т.е.

$$\mathrm{P}\left[x_i - y_i > 0\right] = \mathrm{P}\left[x_i - y_i < 0\right] = 1/2, \quad i = 1, 2, \ldots, l,$$

где l — число ненулевых разностей, $l \leqslant n$. Нулевые разности могут появиться из-за случайных погрешностей или ошибок округления, и соответствующие им пары наблюдений исключаются из рассмотрения.

Статистикой критерия знаков является число «+» или «−» в последовательности знаков разностей парных выборок (x_i, y_i), $i = 1, 2, \ldots, l$. В дальнейшем для определенности берется число знаков «+» . При условии, что проверяемая гипотеза H_0 верна, а пары наблюдений (x_i, y_i) и, следовательно, знаки разностей $x_i - y_i$ *независимы*, число знаков «+» имеет *биномиальное распределение* с параметрами $p = 1/2$ и l, т.е. $B\,(l, 1/2)$. Задача сводится к проверке гипотезы $H_0\colon\ p = 1/2$ при одной из альтернативных групп $H_1^{(1)}\colon\ p > 1/2$, $H_1^{(2)}\colon\ p < 1/2$, $H_1^{(3)}\colon p \neq 1/2$.

Пусть r — наблюденное число знаков «+», а α — заданный уровень значимости. Гипотеза H_0 отклоняется, если при $H_1^{(1)}\colon\ p > 1/2$ выполняется неравенство

$$\sum_{i=r}^{l} C_l^i \left(\frac{1}{2}\right)^l \leqslant \alpha, \tag{1}$$

или при $H_1^{(2)}\colon\ p > 1/2$ выполняется неравенство

$$\sum_{i=0}^{r} C_l^i \left(\frac{1}{2}\right)^l \leqslant \alpha, \tag{2}$$

или, наконец, при $H_0^{(3)}\colon\ p \neq 1/2$ выполняется одно из неравенств

$$\begin{gathered}\sum_{i=0}^{l} C_l^i \left(\frac{1}{2}\right)^l \leqslant \frac{\alpha}{2},\\ \sum_{i=1}^{r} C_l^i \left(\frac{1}{2}\right)^l \leqslant \frac{\alpha}{2}.\end{gathered} \tag{3}$$

Если при соответствующих альтернативных гипотезах неравенства (1)–(3) не выполняются, то гипотеза H_0 не противоречит результатам наблюдений и принимается на уровне значимости α. Часто более удобно проводить проверку гипотезы H_0, используя статистику Фишера. Гипотеза H_0 отклоняется, если при $H_1^{(1)}\colon\ p > 1/2$ выполняется неравенство

$$F_в = \frac{r}{l - r + 1} \geqslant F_{1-\alpha}(k_1, k_2), \tag{4}$$

где $k_1 = 2(l - r + 1)$, $k_2 = 2r$, или при $H_1^{(2)}: \ p < 1/2$ выполняется неравенство

$$F_в = \frac{l - r}{r + 1} \geqslant F_{1-\alpha}(k_1, k_2), \tag{5}$$

где $k_1 = 2(r+1)$, $k_2 = 2(l-r)$, или, наконец, при $H_1^{(3)}: \ p \neq 1/2$ должно выполняться одно из неравенств (4) или (5) с заменой α на $\alpha/2$.

Пример 1. Предполагается, что один из двух приборов, определяющих скорость автомобиля, имеет систематическую ошибку. Для проверки этого предположения определили скорость 10 автомобилей, причем скорость каждого фиксировалась одновременно двумя приборами.

В результате получены следующие данные:

v_1, км/ч	70	85	63	54	65	80	75	95	52	55
v_2, км/ч	72	86	62	55	63	80	78	90	53	57

Позволяют ли эти результаты утверждать, что второй прибор действительно дает завышенные значения скорости? Принять $\alpha = 0{,}10$.

◁ В предположении, что скорости движения автомобилей не зависят друг от друга, задачу можно решить, применяя критерий знаков.

Составим последовательность знаков разностей $v_1 - v_2$: $-, -, +, -, +, 0, -, +, -, -$. Число ненулевых разностей $l = 9$, число положительных разностей $r = 3$. Проверим гипотезу о том, что различие в показаниях приборов вызвано случайными ошибками, т.е. гипотезу $H_0: \ p = 1/2$. Альтернативная гипотеза предполагает, что показания второго прибора имеют положительное смещение; в этом случае вероятность появления положительных разностей должна быть меньше $1/2$. Таким образом, альтернативная гипотеза формулируется так: $H_1: \ p < 1/2$. Для проверки гипотезы H_0 используем неравенство (5). Имеем

$$k_1 = 2 \cdot (3+1) = 8, \quad k_2 = 2 \cdot (9-3) = 12, \quad F_в = \frac{9-3}{3+1} = 1{,}5.$$

Так как по таблице П7 $F_{0,90}(8{,}12) = 2{,}24$, гипотеза H_0 не противоречит результатам наблюдений. Следует считать, что различие в показаниях приборов вызвано случайными ошибками. ▷

19.392.** Сформулировать параметрический аналог критерия знаков (для случая нормально распределенной генеральной совокупности).

19.393. Можно ли применить критерий знаков в случае, когда пары (x_i, y_i), $i = 1, 2, \ldots, n$, статистически зависимы? Если нет, то почему?

Для решения задач 19.394–19.399 использовать критерий знаков.

19.394. Ниже приводится время (в секундах) решения контрольных задач одиннадцатью учащимися до и после специальных упражнений по устному счету. Можно ли считать, что эти упражнения улучшили способности учащихся в решении задач? Принять $\alpha = 0{,}10$.

До упражнений	87	61	98	90	93	74	83	72	81	75	83
После упражнений	50	45	79	90	88	65	52	79	84	61	52

19.395. Для 10 человек была предложена специальная диета. После двухнедельного питания по этой диете масса их тела изменилась следующим образом:

Масса до диеты (кг)	68	80	92	81	70	79	78	66	57	76
Масса после диеты (кг)	60	84	87	79	74	71	72	67	57	70

а) Можно ли рекомендовать эту диету для людей, желающих похудеть?

б) Оказывает ли эта диета вообще какое-либо существенное действие на массу тела?

Принять $\alpha = 0{,}10$.

19.396. Сравнивалось действие двух экстрактов вируса табачной мозаики. Для этого каждая из половин листа натиралась соответствующим препаратом. Число пораженных мест приводится ниже:

Экстракт A	20	39	43	13	28	26	17	49	36
Экстракт B	31	22	45	6	21	13	17	46	31

Можно ли считать, что действие этих экстрактов различно? Принять $\alpha = 0{,}01$.

19.397. Изучалось влияние черного и апрельского пара на урожай ржи. Опыт длился шесть лет. Учитывалась масса 1000 зерен в граммах. Результаты опыта следующие:

Год посева	1	2	3	4	5	6
По черному пару	31,1	24,0	24,6	28,6	29,1	30,1
По апрельскому пару	31,6	24,2	24,8	19,1	29,9	31,0

Можно ли считать, что урожай ржи по апрельскому пару значимо выше, чем по черному? Проверить это предположение, если $\alpha = 0{,}05$.

19.398. Проверить предположение о том, что предлагаемый лечебный препарат не меняет состав крови (в частности, числа лейкоцитов), если препарат испытывался на десяти особях, а последующий анализ крови дал следующие результаты:

$$0{,}97;\ 1{,}05;\ 1{,}09;\ 0{,}88;\ 1{,}01;\ 1{,}14;\ 1{,}03;\ 1{,}07;\ 0{,}94;\ 1{,}02$$

(числа выражают отношение числа лейкоцитов в опыте к числу лейкоцитов в норме). Принять $\alpha = 0{,}01$.

19.399. Изменение урожайности при применении одного из видов предпосевной обработки семян характеризуется следующими данными (в центнерах с гектара):

Год	1972	1973	1974	1975	1976	1977	1978	1979	1980
Необработанные семена	20,0	17,9	20,6	22,0	21,4	23,8	21,4	19,8	18,4
Обработанные семена	22,1	18,5	19,4	22,1	21,7	24,9	21,6	20,3	18,3

Можно ли считать, что предпосевная обработка увеличивает урожайность? Принять $\alpha = 0{,}05$.

2. Критерий Вилкоксона, Манна и Уитни. Критерий применяется для сравнения двух независимых выборок объема n_1 и n_2 и проверяет гипотезу H_0, утверждающую, что выборки получены из однородных генеральных совокупностей и, в частности, имеют равные средние и медианы.

Статистика W критерия определяется следующим образом. Расположим $n_1 + n_2$ значений объединенной выборки в порядке возрастания, т.е. в виде вариационного ряда. Каждому элементу ряда поставим в соответствие его номер в ряду — *ранг*. Если несколько элементов ряда совпадают по величине, то каждому из них присваивается ранг, равный среднему арифметическому их номеров. Последний элемент в ранжированной объединенной выборке должен иметь ранг $n_1 + n_2$. Этот факт можно использовать при проверке правильности ранжирования.

Пусть R_1 — сумма рангов первой выборки, R_2 — сумма рангов второй выборки. Вычислим значения w_1 и w_2:

$$w_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1, \tag{6}$$

$$w_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2. \tag{7}$$

Правильность вычислений проверяется по формуле

$$w_1 + w_2 = n_1 n_2. \tag{8}$$

Выборочное значение $w_в$ статистики критерия есть наименьшее из чисел w_1 и w_2. В таблице П10 приводятся вероятности того, что $W < w_в$, при условии, что гипотеза H_0 верна, т.е. значения

$$p = \mathrm{P}\left[W < w_в / H_0\right]$$

для выборок объема n_1 и n_2 $(n_1 \geqslant n_2)$. При односторонней (двусторонней) альтернативной гипотезе H_1 гипотеза H_0 отклоняется, если $p \leqslant \alpha$ $(p \leqslant \alpha/2)$, где α — заданный уровень значимости. В противном случае гипотеза H_0 не противоречит результатам наблюдений.

Если объем каждой из выборок больше 8, то проверку гипотезы H_0 можно проводить, используя статистику

$$Z = \frac{W - \dfrac{1}{2} n_1 n_2}{\sqrt{\dfrac{1}{12} n_2 n_2 (n_1 + n_2 + 1)}}, \tag{9}$$

имеющую (при условии, что верна гипотеза H_0) приблизительно нормальное распределение $N(0, 1)$. В этом случае гипотеза H_0 отклоняется на уровне значимости α, если выборочное значение $z_в$ статистики Z удовлетворяет неравенству

$$z_в < u_\alpha \quad (z_в > u_{1-\alpha})$$

при левосторонней (правосторонней) альтернативной гипотезе H_1 и если

$$|\, z_в \,| > u_{1-\alpha/2}$$

при двусторонней альтернативной гипотезе H_1.

Пример 2. Измерялось напряжение пробоя у диодов, отобранных случайным образом из двух партий. Результаты измерения (в вольтах) следующие:

1-я партия	39	50	61	67	40	40	54	–
2-я партия	60	53	42	41	40	54	63	69

Можно ли считать, что у диодов из второй партии напряжение пробоя выше, чем у диодов из первой партии? Принять $\alpha = 0{,}10$.

◁ Воспользуемся критерием Вилкоксона. Составим вариационный ряд, отмечая принадлежность элемента к первой партии черточкой сверху В результате получим следующую ранжированную последовательность:

Элемент	$\overline{39}$	$\overline{40}$	$\overline{40}$	40	41	42	$\overline{50}$	53	$\overline{54}$	54	60	$\overline{61}$	63	$\overline{67}$	69
Ранг	1	3	3	3	5	6	7	8	9,5	9,5	11	12	13	14	15

Сумма рангов первой выборки $R_1 = 49{,}5$, сумма рангов второй выборки $R_2 = 70{,}5$, $n_1 = 7$, $n_2 = 8$. По формулам (6) и (7) находим

$$w_1 = 7 \cdot 8 + \frac{7 \cdot (7+1)}{2} - 49{,}5 = 34{,}5,$$

$$w_2 = 7 \cdot 8 + \frac{8 \cdot (8+1)}{2} - 70{,}5 = 21{,}5.$$

Используя соотношение (8), проверяем правильность вычислений:

$$34{,}5 + 21{,}5 = 56.$$

Выборочное значение статистики $w_в$ равно меньшему из чисел 34,5 и 21,5, т.е.

$$w_в = 21{,}5.$$

По таблице П10 находим (с интерполяцией)

$$p = \mathrm{P}\,[W < 21{,}5] = 0{,}25.$$

Так как предположение о том, что у диодов второй партии напряжение пробоя выше, соответствует односторонней альтернативной гипотезе H_1, а вероятность $p = 0{,}25$ превышает уровень значимости $\alpha = 0{,}1$, гипотеза H_0 не противоречит результатам измерений. Следовательно, результаты измерений не дают оснований считать, что напряжение пробоя у диодов второй партии выше, чем у диодов первой партии. ▷

П р и м е р 3. В условиях примера 2 получены результаты новой серии измерений напряжения пробоя у диодов (в вольтах):

1-я партия	50	41	48	60	46	60	51	42	62	54	42	46
2-я партия	38	40	47	51	63	50	63	57	59	51	–	–

Имеются ли основания утверждать, что напряжение пробоя у диодов обеих партий различно?

Решить пример, используя:

а) критерий Вилкоксона;

б) проверку гипотезы о равенстве средних (в предположении, что обе выборки получены из нормально распределенных генеральных совокупностей). Принять $\alpha = 0{,}1$.

◁ а) Как и в примере 2, упорядочим результаты измерений и определим ранги каждого результата. Имеем

Элемент	38	40	$\overline{41}$	$\overline{42}$	$\overline{42}$	$\overline{46}$	$\overline{46}$	47	$\overline{48}$	$\overline{50}$	50
Ранг	1	2	3	4,5	4,5	6,5	6,5	8	9	10,5	10,5

Элемент	$\overline{51}$	51	51	$\overline{54}$	57	59	$\overline{60}$	$\overline{60}$	$\overline{62}$	63	63
Ранг	13	13	13	15	16	17	18,5	18,5	20	21,5	21,5

Найдем суммы рангов

$$R_1 = 129{,}5, \qquad R_2 = 123{,}5.$$

Так как $n_1 = 12$, $n_2 = 10$, по формулам (6) и (7) находим

$$w_1 = 12 \cdot 10 + \frac{12 \cdot (12+1)}{2} - 129{,}5 = 68{,}5,$$

$$w_2 = 12 \cdot 10 + \frac{10 \cdot (10+1)}{2} - 123{,}5 = 51{,}5.$$

Выборочное значение $w_в$ статистики критерия таково:

$$w_в = 51{,}5.$$

Так как $n_1 > 8$ и $n_2 > 8$, то для проверки гипотезы H_0 используем статистику Z. Выборочное значение этой статистики определяется по формуле (9):

$$z_в = \frac{51{,}5 - \frac{1}{2}12 \cdot 10}{\sqrt{\frac{1}{12} \cdot 12 \cdot 10\,(12+10+1)}} \approx -0{,}56.$$

Проверяемое предположение соответствует двусторонней альтернативной гипотезе, следовательно, значение $|\,z_a\,|$ сравнивается с квантилью $u_{1-\alpha/2}$, которая определяется по таблице П1:

$$u_{0,95} = 1{,}645.$$

Таким образом, утверждение о том, что напряжение пробоя у диодов обеих партий различно, следует отклонить.

б) Проверим гипотезу о равенстве средних (§ 4 п. 2). По результатам наблюдений вычислим оценки средних и дисперсий для каждой выборки:

$$\overline{x}_1 \approx 50{,}17, \quad s_1^2 \approx 55{,}06, \quad \overline{x}_2 = 51{,}90, \quad s_2^2 \approx 76{,}32.$$

Предварительно следует проверить гипотезу о равенстве дисперсий $H_0 : \sigma_1^2 = \sigma_2^2$ при альтернативной гипотезе $H_1 : \sigma_1^2 \neq \sigma_2^2$. Для этого найдем отношение выборочных дисперсий и сравним его с квантилью $F_{1-\alpha/2}(n_2 - 1, n_1 - 1)$ (таблица П7):

$$\frac{s_2^2}{s_1^2} \approx 1{,}39, \ F_{0,95}(9{,}11) = 2{,}90.$$

Так как $1{,}39 < 2{,}90$, гипотеза о равенстве дисперсий принимается, следовательно, проверку гипотезы о равенстве средних можно проводить на основе статистики Стьюдента (§ 4, таблица 4.2). Выборочное значение этой статистики равно

$$t_в = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} =$$

$$= \frac{50{,}17 - 51{,}90}{\sqrt{\frac{(12-1)\cdot 55{,}06 + (10-1)\cdot 76{,}32}{12+10-2}}\sqrt{\frac{1}{12}+\frac{1}{10}}} \approx -0{,}5.$$

Так как значение квантили $t_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2)$ по таблице П6 равно $t_{0,95} = 1{,}725$ и $|\,t_в\,| < 1{,}725$, гипотеза о равенстве средних принимается. Таким образом, утверждение о том, что напряжение пробоя у диодов обеих партий различно, следует отклонить. Это совпадает с результатом, полученным в а). ▷

19.400. У полевых транзисторов из двух партий, изготовленных с применением различных технологий, измерялось дифференциальное сопротивление канала R_i. Результаты измерений (в микроомах) следующие:

Технология A	0,01	0,02	0,12	0,30	0,29	0,15	0,21	–
Технология B	0,15	0,07	0,25	0,15	0,22	0,18	0,18	0,27

Влияет ли технология изготовления на величину дифференциального сопротивления канала R_i? Принять $\alpha = 0{,}05$.

19.401. В условиях задачи 19.400 у полевых транзисторов измерялась еще одна характеристика: емкость затвор — сток. Увеличилась ли величина емкости затвор — сток у транзисторов, изготовленных по технологии B, если измерения дали следующие результаты (в пикофарадах):

Технология A	2,8	3,0	3,1	3,2	3,3	3,4	3,7	–
Технология B	3,8	3,4	3,6	2,9	2,8	3,0	3,4	3,0

Принять $\alpha = 0{,}05$.

19.402. В биохимическом исследовании, проведенном методом меченых атомов, по результатам изучения 8 препаратов контрольной серии получены следующие показания счетчика импульсов (в импульсах в минуту):

Опыт	340	343	322	349	332	320	313	304
Контроль	318	321	318	301	312	—	—	–

Можно ли считать, что полученные значения опытной и контрольной серий различны? Принять $\alpha = 0{,}10$.

19.403. По выборкам из двух партий микросхем после операции легирования поликремния измерялось удельное сопротивление. Результаты замеров следующие:

1-я партия	52,2	33	76	32,5	49,5	32,5	191,5	112,5
2-я партия	119	17,5	43,5	43,5	90,5	40	50	108

1-я партия	52,9	114,8	33,7	69,1	112,5	48,5	16,5
2-я партия	62,4	16,5	97,5	96	46	–	–

Можно ли утверждать, что обе партии получены из одной генеральной совокупности? Принять $\alpha = 0{,}10$.

19.404. В условиях задачи 19.403 после операции разгонки бора измерена глубина слоя диффузии и получены следующие результаты (мкм):

1-я партия	9,8	9,8	8,6	8,6	9,2	9,2	9,8
2-я партия	8,6	9,2	10,4	9	9,8	9,2	9,6

1-я партия	9	10	9,4	9	11,2	10,8	9,2	9,4
2-я партия	10	9,8	9,0	9,8	8,7	8,6	–	–

Можно ли считать, что глубина слоя диффузии в микросхемах из обеих партий различна? Принять $\alpha = 0{,}10$.

19.405. Длина тела личинок щелкуна, обитающих в посевах озимой ржи и проса (выраженная в мм), варьируется следующим образом:

В посевах ржи	7	10	14	15	12	16	12
В посевах проса	11	12	16	13	18	15	–

На основании этих проб создается впечатление о более крупных размерах личинок щелкунов, обитающих на просе. Проверить это предположение, используя критерий Вилкоксона. Принять $\alpha = 0{,}01$.

19.406. Изучалось влияние кобальта на увеличение массы тела кроликов. Опыт проводился на двух группах животных — опытной и контрольной. Возраст кроликов колебался в пределах от 1,5 до 2 месяцев. Исходная масса тела особей находилась в пределах от 500 до 600 г. Опыт длился 8 недель. Обе группы содержались на одном и том же кормовом рационе, но в отличие от контрольных, опытные кролики ежедневно получали в виде водного раствора по 0,06 г хлористого кобальта на 1 кг массы тела. За время опыта у животных наблюдались следующие прибавки в массе (за 1 неделю):

Контрольные	560	580	600	420	530	490	580	470
Опытные	692	700	621	640	561	680	630	–

Можно ли утверждать, что прибавки хлористого кобальта действительно дают прибавку массы тела? Принять $\alpha = 0{,}10$.

3. Критерий для проверки гипотезы H_0 о равенстве дисперсий двух генеральных совокупностей. Этот критерий может использоваться вместо критерия, основанного на отношении выборочных дисперсий (§ 4, п. 2, таблица 4.1), при условии, что у рассматриваемых генеральных совокупностей равны или близки характеристики положения, т.е. средние или медианы (см. пп. 1, 2 настоящего параграфа). Критерий применяется следующим образом. Объединенная выборка объема $n_1 + n_2$ упорядочивается в порядке возрастания и отмечается принадлежность каждого элемента к той или иной выборке. Ранги присваиваются по следующему правилу: наименьшему значению присваивается ранг 1, два наибольших значения получают ранги 2 и 3, ранги 4 и 5 получают следующие наименьшие значения и т.д. Схема расстановки рангов показана ниже:

$$1,\ 4,\ 5,\ 8,\ 9,\ \ldots,\ 7,\ 6,\ 3,\ 2.$$

Каждому из совпадающих по величине элементов присваивается ранг, равный среднему арифметическому (как в критерии Вилкоксона). При $n_1 > 10$, $n_2 > 10$ статистика Z критерия определяется по формуле

$$Z = \frac{\left| R_2 - \dfrac{n_2(n_1 + n_2 + 1)}{2} \right| - \dfrac{1}{2}}{\sqrt{\dfrac{n_1(n_1 + n_2 + 1)}{12}}}, \tag{10}$$

где R_2 — сумма рангов для выборки меньшего объема n_2 ($n_2 \leqslant n_1$). При условии, что верна гипотеза H_0: дисперсии сравниваемых генеральных совокупностей равны — статистика Z имеет приблизительно нормальное распределение $N(0,\ 1)$. Гипотеза H_0 отклоняется, если выборочное значение $z_в$ статистики Z удовлетворяет неравенству

$$z_в < u_\alpha \quad (z_в > u_{1-\alpha})$$

при левосторонней (правосторонней) альтернативной гипотезе H_1 и если

$$|\, z_в \,| > u_{1-\alpha/2}$$

при двусторонней альтернативной гипотезе $H_1 : \sigma_1^2 \neq \sigma_2^2$. В противном случае гипотеза H_0 не противоречит результатам наблюдений на уровне значимости α.

Если объем одной из выборок меньше или равен десяти, то приведенный выше критерий может использоваться только для приблизительных расчетов.

Пример 4. Проверить гипотезу о равенстве дисперсий по данным примера 3.

◁ При решении примера 3 было установлено, что характеристики положения у рассматриваемых генеральных совокупностей равны, следовательно, критерий для проверки гипотезы H_0 о равенстве дисперсий применим. Воспользуемся упорядоченными результатами измерений из решения примера 3 и расставим ранги по приведенному выше правилу. Имеем

Элемент	38	40	$\overline{41}$	$\overline{42}$	$\overline{42}$	$\overline{46}$	$\overline{46}$	47	$\overline{48}$	$\overline{50}$	50
Ранг	1	4	5	8,5	8,5	12,5	12,5	16	17	20,5	20,5

Элемент	$\overline{51}$	51	51	$\overline{54}$	57	59	$\overline{60}$	$\overline{60}$	$\overline{62}$	63	63
Ранг	19,7	19,7	19,7	15	14	11	8,5	8,5	6	2,5	2,5

Вычислим сумму рангов для 2-й партии ($n_2 = 10$); имеем $R_2 = 110{,}9$. Выборочное значение статистики критерия определяем по формуле (10):

$$z_в = \frac{\left|110{,}9 - \dfrac{10\,(10+12+1)}{2}\right| - \dfrac{1}{2}}{\sqrt{\dfrac{10\,(10+12+1)}{12}}} \approx 0{,}237.$$

Так как при $\alpha = 0{,}10$ имеем $u_{0{,}95} = 1{,}645$, то при двусторонней альтернативной гипотезе $H_1 : \sigma_1^2 \neq \sigma_2^2$ гипотеза H_0 не противоречит результатам измерений. ▷

19.407. Двум группам испытуемых предлагалось провести опознание трех начертаний цифры 5. Результаты эксперимента (время опознания в секундах) следующие:

1-я группа	25	28	27	29	26	24	28	23	30	25	26	25
2-я группа	18	19	31	32	17	15	41	35	38	13	14	–

Можно ли считать, что дисперсии результатов для первой и второй групп различны? Принять $\alpha = 0{,}05$.

19.408. В течение некоторого времени суточная производительность двух автоматов характеризуется следующими данными:

1-й автомат	105	60	83	111	138	71	87	130	93	105	122
2-й автомат	172	45	51	155	117	103	82	93	31	51	—

Проверить гипотезу H_0 о равенстве дисперсий этих показателей. Принять $\alpha = 0{,}10$.

19.409. Для контроля настройки двух станков-автоматов, производящих детали по одному чертежу, определили отклонения от номинальных размеров у нескольких деталей, изготовленных на обоих станках. В результате получили следующие данные (в мкм):

Станок A	44	−14	32	8	−50	20	−35	15	10	−8	−20	5
Станок B	52	−49	61	−35	−48	18	−45	35	23	21	−59	−19

Проверить гипотезу H_0 о равенстве дисперсий по этим данным на уровне значимости $\alpha = 0{,}10$.

19.410. Контролируемый размер нескольких деталей был проверен до и после наладки станка. В результате получены следующие результаты (в мм):

До наладки	36,4	37,5	36,9	37,6	38,1	35,5	37,8	38,3	36,6	38,4	37,5
После наладки	36,8	39,2	37,6	39,9	39,6	34,2	36,5	36,3	39,8	–	–

Можно ли считать, что дисперсии результатов измерений до и после наладки станка различны? Принять $\alpha = 0{,}05$.

4. Критерий серий. Критерий применяется для проверки гипотезы H_0, утверждающей, что элементы выборки получены случайным образом и независимы. Пусть $x_1, x_2, \ldots, x_n$ — выборка результатов наблюдений, а $\tilde{h}_x$ — выборочная медиана, определенная по этим данным. Каждому элементу выборки поставим в соответствие знак «+» либо «−» в зависимости о того, больше или меньше медианы его значение (нулевые значения не учитываются). Тем самым всей выборке поставлен в соответствие набор знаков. Обозначим n_1 число знаков «+», а n_2 — число знаков «−» в полученном наборе знаков. *Серией* в этом наборе называется всякая последовательность, состоящая из одинаковых знаков и ограниченная противоположными знаками, либо находящаяся в начале или конце набора. Например, в наборе

$$+-+++-----++,$$

содержится 5 серий: (+), (−), (+++), (−−−−−), (++), $n_1 = 6$, $n_2 = 6$.

Статистикой критерия серий является число серий N. Критическая область определяется неравенствами $N \leqslant N_1$ и $N \geqslant N_2$. Значения границ критической области N_1 и N_2 для уровня значимости $\alpha = 0,05$ приведены в таблице П11.

Пример 5. Скорости автомобилей в некоторой точке трассы образовали следующий ряд (км/ч):

$$31,\ 39,\ 40,\ 45,\ 27,\ 28,\ 35,\ 55,\ 21,\ 33,\ 42,\ 36.$$

Можно ли считать полученные значения случайными? Принять $\alpha = 0,05$.

$\triangleleft$ Найдем оценку медианы этих данных (§ 1 п. 3). Для этого представим данные в виде вариационного ряда:

$$21,\ 27,\ 28,\ 31,\ 33,\ 35,\ 36,\ 39,\ 40,\ 42,\ 45,\ 55.$$

Оценка медианы $\tilde{h} = \dfrac{35+36}{35,5}$. Исходному ряду наблюдений соответствует следующая последовательность знаков:

$$-+++---+--++,$$

где $n_1 = 6$, $n_2 = 6$, число серий $N = 6$. По таблице П11 при $\alpha = 0,05$ находим $N_1 = 3$, $N_2 = 11$. Таким образом, гипотеза H_0 принимается: полученные значения скорости можно считать случайными. $\triangleright$

При больших объемах выборки, когда либо n_1, либо n_2, либо оба значения больше 20, для проверки гипотезы H_0 можно использовать статистику Z, выборочное значение $z_в$ которой вычисляется по формуле

$$z_в = \frac{\left|N - \dfrac{2n_1n_2}{n_1+n_2} - 1\right| - \dfrac{1}{2}}{\sqrt{\dfrac{2n_1n_2[2n_1n_2 - (n_1+n_2)]}{(n_1+n_2)^2(n_1+n_2-1)}}}. \tag{11}$$

При условии, что верна гипотеза H_0, статистика Z имеет приблизительно нормальное распределение $N(0,\ 1)$. В этом случае критическая область определяется неравенствами

$$z_в \leqslant u_{\alpha/2} \quad \text{или} \quad z_в \geqslant u_{1-\alpha/2}.$$

Замечание. Критерий серий применяется для проверки случайности любой выборки, элементами которой являются два различных символа, например: 1 и 0, A и B, $+$ и $-$. Статистикой критерия является число серий в этой выборке.

Пример 6. Можно ли считать, что последовательность

$$110010001010100111011010010000101000$$

получена из совокупности случайных последовательностей? Принять $\alpha = 0,01$.

◁ В данной последовательности число нулей $n_1 = 21$, а число единиц $n_2 = 15$. Число серий $N = 23$.

Так как $n_1 > 20$, для проверки гипотезы H_0, утверждающей, что данная последовательность получена из совокупности случайных последовательностей, воспользуемся статистикой Z. По формуле (11) выборочное значение $z_в$ этой статистики равно:

$$z_в = \frac{\left|22 - \dfrac{2 \cdot 21 \cdot 15}{21+15}\right| - \dfrac{1}{2}}{\sqrt{\dfrac{2 \cdot 21 \cdot 15 \cdot [2 \cdot 15 \cdot 21 - (21+15)]}{(21+15)^2 \cdot (21+15-1)}}} \approx 1{,}394.$$

Так как $u_{1-\alpha/2} = u_{0,995} = 2{,}576$ (таблица П1), гипотеза H_0 принимается: можно считать, что данная последовательность получена из совокупности случайных последовательностей. ▷

19.411. Глубина слоя диффузии, определенная по выборке из партии микросхем, имеет следующие значения (в мкм):

$$9{,}8;\ 9{,}8;\ 8{,}6;\ 8{,}6;\ 9{,}2;\ 9{,}2;\ 9{,}8;$$

$$9{,}0;\ 10{,}0;\ 9{,}4;\ 9{,}0;\ 11{,}2;\ 10{,}8;\ 9{,}2;\ 9{,}4.$$

Проверить гипотезу H_0 о том, что полученные результаты распределены случайным образом. Принять $\alpha = 0{,}05$.

19.412. При заданном токе 10 мА измерялось прямое падение напряжения на диодах. Получены следующие значения (в вольтах)

$$0{,}917;\ 0{,}918;\ 0{,}921;\ 0{,}909;\ 0{,}919;\ 0{,}917;$$

$$0{,}916;\ 0{,}917;\ 0{,}918;\ 0{,}919;\ 0{,}919;\ 0{,}916.$$

Можно ли считать, что эти значения случайны? Принять $\alpha = 0{,}05$.

19.413. Измерение массы некоторого вещества, полученного в результате химической реакции, дало следующие результаты (в граммах):

$$14,\ 14,\ 14,\ 15,\ 16,\ 17,\ 21,\ 22,\ 17,\ 19,\ 23.$$

Случайны ли эти результаты? Принять $\alpha = 0{,}05$.

19.414. Производительность цеха в течение 12 рабочих дней характеризуется следующими цифрами (в условных единицах):

$$13{,}0;\ 13{,}1;\ 13{,}0;\ 12{,}5;\ 12{,}8;\ 12{,}3;\ 12{,}1;$$

$$12{,}2;\ 12{,}1;\ 12{,}7;\ 12{,}0;\ 12{,}6.$$

Можно ли считать, что изменение производительности вызвано случайными причинами? Принять $\alpha = 0{,}05$.

19.415. Для 13 деталей получены следующие отклонения контрольного размера от номимального значения (в мкм):

$$+8,\ +10,\ +5,\ -5,\ -9,\ +7,\ +6,$$
$$-11,\ -4,\ -4,\ +15,\ +21,\ -3.$$

Можно ли считать, что полученная выборка представляет результаты случайных и независимых наблюдений? Принять $\alpha = 0{,}05$.

19.416. Радист принял следующую последовательность символов:

— — — — · · — · — · · — · — · · — — — — · · — · —

· · — · — — — — · · · — — — — · · — · — ·

Можно ли считать, что эта последовательность есть случайная выборка? Принять $\alpha = 0{,}05$.

19.417. При подбрасывании монеты 45 раз последовательность результатов (*Г* — выпадение герба, *Р* — выпадение решетки) имела следующий вид:

ГГГГГРРРРГГГРРРГРГРРРРРР

ГГГРРРРРГГГРГГГГРРР.

Является ли такая последовательность случайной выборкой? Принять $\alpha = 0{,}05$.

5. Ранговая корреляция. Пусть (x_i, y_i), $i = 1, 2, \ldots, n$ — выборка наблюдений непрерывных случайных величин X и Y. Каждому значению x_i, $i = 1, 2, \ldots, n$ поставим в соответствие ранг x'_i, т.е. номер элемента x_i в вариационном ряду $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$. Аналогичным способом определим ранги y'_i элементов y_i, $i = 1, 2, \ldots, n$. Каждой паре (x_i, y_i) соответствует пара рангов (x'_i, y'_i). Вычислим коэффициент корреляции (§ 1, п. 3, задача 19.62, формула (28)) по выборке рангов (x'_i, y'_i), $i = 1, 2, \ldots, n$. Полученное значение r_s называется *выборочным значением рангового коэффициента корреляции Спирмена* ρ_s. Ранговый коэффициент корреляции, так же как и коэффициент корреляции r, характеризует зависимость между случайными величинами X и Y, но вычисляется значительно проще, а именно, справедлива следующая формула:

$$r_s = 1 - \frac{6\sum (x'_i - y'_i)^2}{n\,(n^2 - 1)}. \tag{12}$$

Поэтому при большом объеме выборки для оценки зависимости между случайными величинами используется ранговая корреляция. Коэффициент r_s является *непараметрической мерой связи* и, следовательно, может использоваться при произвольном непрерывном распределении генеральной совокупности, в то время как использование коэффициента корреляции r предполагает двумерное нормальное распределение генеральной совокупности. Гипотеза H_0 : $\rho_s = 0$ при альтернативной гипотезе H_1 : $\rho_s \neq 0$ и при объеме выборки $n \geqslant 9$ проверяется по значению статистики

$$|r_s| \sqrt{\frac{n-2}{1-r_s^2}} = T\,(n-2). \tag{13}$$

При условии, что верна гипотеза H_0 эта статистика имеет распределение Стьюдента с $n-2$ степенями свободы. Если $t_в > t_{1-\alpha/2}(n-2)$, где α — заданный уровень значимости, то гипотеза H_0 отклоняется, т.е. между X и Y существует ранговая корреляционная зависимость.

Пример 7. Вычислить коэффициент ранговой корреляции для следующей выборки:

x	68,8	63,3	75,5	67,2	71,3	72,8	76,5	63,5	69,9	71,4
y	167,0	113,3	159,9	153,6	150,8	181,2	173,1	115,4	125,6	166,2

Проверить значимость ранговой корреляции при $\alpha = 0{,}10$.

◁ Определим ранги элементов исходной выборки. Предварительно перепишем исходную выборку, упорядочив ее элементы по верхней строке (т.е. по значениям x_i), в результате получим

x	63,3	63,5	67,2	68,8	69,9	71,3	71,4	72,8	75,7	76,5
y	113,3	115,4	153,6	167,0	125,6	150,8	166,2	181,2	159,9	173,1

Определим ранги для значения y_i. Вариационный ряд для y_i имеет вид

i	1	2	3	4	5	6	7	8	9	10
$y^{(i)}$	113,3	115,4	125,6	150,8	153,6	159,9	166,2	167,0	173,1	181,2

Таким образом, упорядоченной по элементам x_i выборке соответствует следующая последовательность пар рангов и их разностей:

x_i'	1	2	3	4	5	6	7	8	9	10
y_i'	1	2	5	8	3	4	7	10	6	9
$x_i' - y_i'$	0	0	-2	-4	2	2	0	-2	3	1

Так как $\sum\limits_{i=1}^{10}(x_i' - y_i')^2 = 42$, по формуле (12) находим

$$r_s = 1 - \frac{6 \cdot 42}{10 \cdot (10^2 - 1)} \approx 0{,}745.$$

Проверим значимость полученного результата. Найдем выборочное значение статистики (13):

$$t_в = 0{,}745\sqrt{\frac{10-2}{1-(0{,}745)^2}} \approx 3{,}159.$$

Так как $t_{0,95}(8) = 1{,}860$ (таблица П6), ранговая корреляция значима. ▷

19.418. Бегуны, ранги которых при построении по росту были 1, 2, ..., 10, заняли на состязаниях следующие места:

$$6,\ 5,\ 1,\ 4,\ 2,\ 7,\ 8,\ 10,\ 3,\ 9.$$

Как велика ранговая корреляция между ростом и быстротой бега?

19.419. Цветные диски, имеющие порядок оттенков 1, 2, ..., 15, были расположены испытуемым в следующем порядке:

$$7,\ 4,\ 2,\ 3,\ 1,\ 10,\ 6,\ 8,\ 9,\ 5,\ 11,\ 15,\ 14,\ 12,\ 13.$$

Охарактеризовать способность испытуемого различать оттенки цветов с помощью коэффициентов ранговой корреляции между действительными и наблюдаемыми рангами.

19.420. Найти коэффициент ранговой корреляции между урожайностью пшеницы и картофеля на соседних полях по следующим данным:

Годы	1926	1927	1928	1929	1930	1931	1932	1933	1934	1935	1936	1937
Пшеница (ц)	20,1	23,6	26,3	19,9	16,7	23,2	31,4	33,5	28,2	35,3	29,3	30,5
Картофель (т)	7,2	7,1	7,4	6,1	6,0	7,3	9,4	9,2	8,8	10,4	8,0	9,7

19.421. Для контрольной партии интегральных схем по нескольким параметрам определено значение критерия годности K. Найти коэффициент ранговой корреляции между значениями K и удельного сопротивления p-кармана R_p, а также между значениями R_p и напряжением отсечки V_0 по следующим данным:

K	0,226	0,187	0,678	0,141	0,197	0,339	0,421	0,141	0,127	0,819
R_p, $\frac{Ом \cdot мм^2}{м}$	905	1004	1119	1200	1340	1261	1140	1190	1060	1130
V_0, В	1,2	1,9	1,7	1,5	4,5	2,2	2,3	2,4	1,8	1,4

Проверить значимость полученных коэффициентов при $\alpha = 0{,}10$.

19.422. Измерения длины головы (x) и длины грудного плавника (y) у 16 окуней дали результаты (в мм):

x	66	61	67	73	51	59	48	47	58	44	41	54	52	47	51	45
y	38	31	36	43	29	33	28	25	36	26	21	30	20	27	28	26

а) Найти коэффициент ранговой корреляции. Проверить значимость полученного результата при $\alpha = 0{,}05$.

б) Найти коэффициент корреляции и проверить его значимость при $\alpha = 0{,}05$ в предположении, что выборка наблюдений получена из нормально распределенной двумерной совокупности.

19.423. Связь между массой тела (x) и количеством гемоглобина в крови (y) у павианов-гамадрилов характеризуется следующими данными:

Масса тела, кг	18,3	17,7	19	18	19	22	21	21	20	30
Гемоглобин (по Сали)	70	74	72	80	77	80	80	89	76	86

а) Найти коэффициент ранговой корреляции.

б) Найти коэффициент корреляции.

19.424. Предположим, что между переменными x и y существует линейная зависимость: $y = ax + b$, $a > 0$. Показать, что коэффициент ранговой корреляции $r_s = 1$.

19.425. Пусть ранги двух характеристик соответствуют обратному порядку. Показать, что ранги x_i' и y_i' связаны следующим соотношением:

$$x_i' + y_i' = n + 1, \qquad i = 1, 2, \ldots, n,$$

и, следовательно, $r_s = -1$.

19.426. Изменяется ли коэффициент ранговой корреляции, если вместо исходных значений выборки (x_i, y_i), $i = 1, 2, \ldots, n$, рассмотреть значения $(\varphi(x_i), \psi(y_i))$, $i = 1, 2, \ldots, n$, где φ и ψ — монотонно возрастающие функции?

19.427*. Коэффициент корреляции по выборке (x_i, y_i), $i = 1, 2, \ldots, n$, вычисляется следующим образом:

$$r = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}.$$

Используя этот результат, вывести формулу (12) для определения коэффициента ранговой корреляции.

ОТВЕТЫ И УКАЗАНИЯ

ГЛАВА 18

18.1. $\Omega = \{(k, m) \,|\, 1 \leqslant k, m \leqslant 6\}$, $A = \{(3, 3), (6, 3), (3, 6), (6, 6)\}$, $B = \{(k, m) \,|\, 1 \leqslant k, m \leqslant 5\}$, $C = \{(k, m) \,|\, 4 \leqslant k, m \leqslant 6\}$, $D = \{(k, k) \,|\, 1 \leqslant k \leqslant 6\}$. **18.2.** $\Omega =$ (ггг), (цгг), (гцг), (ггц), (гцц), (цгц), (ццг), (ццц)}. $A = \{$(гцц), (цгц), (ццг)$\}$, $B = \{$(ггг)$\}$, $C = \{$(ггг), (цгг), (гцг), (ггц)$\}$, $D = \{$(ггг), (цгг), (ггц)$\}$. **18.3.** $\Omega = \{n \,|\, n \in N\}$, $A = \{3\}$,

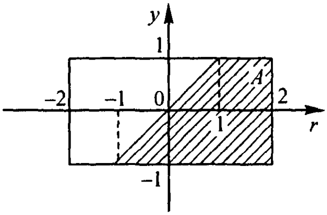


Рис. 37

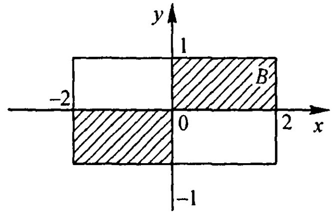


Рис. 38

$B = \{n \,|\, n = 3, 4, \ldots\} = \Omega \backslash \{1, 2\}$. **18.4.** $\Omega = \{(i, j, k) \,|\, 1 \leqslant i \leqslant 3, 1 \leqslant j \leqslant 3, 1 \leqslant k \leqslant 3\}$, $A = \{(i, j, k) \,|\, 2 \leqslant i \leqslant 3, 2 \leqslant j \leqslant 3, 2 \leqslant k \leqslant 3\}$, $B = \{(i, j, k) \,|\, 1 \leqslant i \leqslant 3, 1 \leqslant j \leqslant 3, 1 \leqslant k \leqslant 3, i \neq j \neq k\}$, $C = \{(i, i, i) \,|\, 1 \leqslant i \leqslant 3\}$.

18.5. $\Omega = \{(x, y) \,|\, \leqslant -2 \leqslant x \leqslant 2, -1 \leqslant y \leqslant 1\}$. Множества, соответствующие указанным событиям, см. на рис. 37–39. Пары совместных событий: A и B, B и C, A и C. **18.6.** Множество эле-

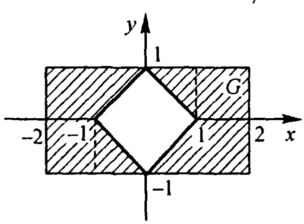


Рис. 39

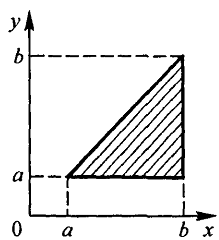


Рис. 40

ментарных исходов Ω и множества, соответствующие указанным событиям, изображены на рис. 40–43. Пара несовместных событий: A и C. **18.7.** $\Omega = \{(x, y) \,|\, 0 \leqslant x, y \leqslant 60\}$. Множество, соответствующее событию A, изображено на рис. 44. **18.8.** Множества, соответствующие событиям B и C, изображены на рис. 45 и 46. **18.9.** Множества, соответствующие событиям D, E и F, изображены на рис. 47–49. **18.10.** $\Omega = \{(x, y) \,|\, x \in \mathbb{Z}_0, y \in \mathbb{Z}_0\}$. $A = \{(x, y) \,|\, x \in \mathbb{Z}_0, y \in \mathbb{Z}_0, x > y\}$,

$B = \{(x, y) \mid x \in \mathbb{Z}_0,\ y \in \mathbb{Z}_0,\ x \neq y\}$, $C = \{(x, y) \mid x = 1,\ y = 3\} = \{(1, 3)\}$, $D = \{(x, y) \mid x \in \mathbb{Z}_0,\ y \in \mathbb{Z}_0,\ x + y \geqslant 3\}$. Указание. Пусть x — количество голов, забитых командой «Динамо», y — количество голов, забитых командой «Спартак», $\mathbb{Z}_0$ — множество неотрицательных целых чисел. При записи множества Ω и его подмножеств-событий предполагается некоторая идеализация реальных условий, позволяющая формализовать модель футбольного матча как вероятностного эксперимента.

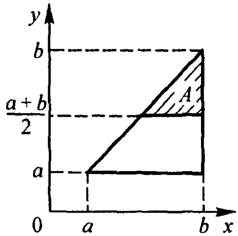


Рис. 41

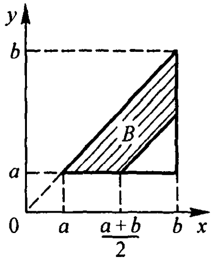


Рис. 42

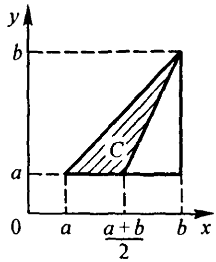


Рис. 43

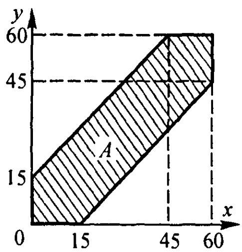


Рис. 44

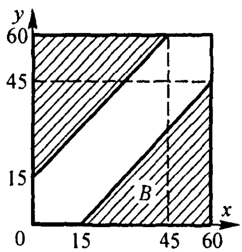


Рис. 45

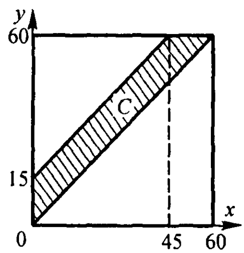


Рис. 46

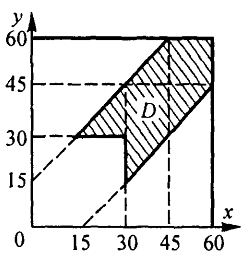


Рис. 47

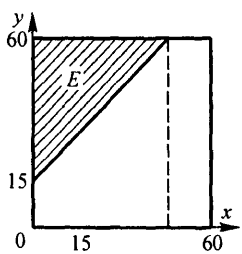


Рис. 48

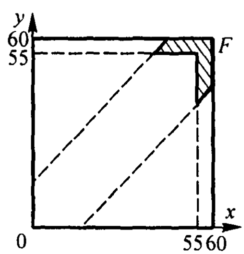


Рис. 49

В частности, множества возможных значений x и y следует считать неограниченными, поскольку до игры нет никаких оснований установить верхнюю границу счета. Практически, конечно, сколь угодно большой счет ни в какой игре не осуществим. **18.11.** Событие B наблюдаемо, A и C не наблюдаемы. Указание. Множество элементарных исходов данного эксперимента можно записать в виде $\Omega = \{(v, \varphi) \mid v \geqslant 0,$

$(2n)^\circ \leqslant \varphi < (2n+2)^\circ,\ n = 0, 1, \ldots, 179\}$. **18.14.** б) $\overline{A_1 + A_2 + \cdots + A_n} = \overline{A}_1\overline{A}_2 \ldots \overline{A}_n$, $\overline{A_1 A_2 \ldots A_n} = \overline{A}_1 + \overline{A}_2 + \cdots + \overline{A}_n$. **18.15.** $\triangleleft$ Преобразуем правую часть. Используя доказанное в примере 4 свойство дистрибутивности умножения относительно сложения, получим $(A + C)(B + C) = A(B + C) + C(B + C) = AB + AC + CB + C = AB + (A + B)C + C\Omega = AB + (A + B + \Omega)C = AB + C$. $\triangleright$ **18.17.** Указание. Воспользоваться тождеством из задачи 18.16. **18.21.** Учесть, что $B = A + B\overline{A}$. Далее перейти к противоположному событию $\overline{B}$ и воспользоваться правилом де Моргана. **18.26.** Указание. Преобразовать правую часть, используя тождество из задачи 18.16 и правила де Моргана. **18.27.** Указание. Воспользоваться тождеством из задачи 18.16 и правилами де Моргана. **18.32.** $X = \overline{B}$. **18.33.** $\triangleleft$ Необходимость. Пусть $A - B = \varnothing$, тогда $A\overline{B} = \varnothing$ и $\overline{A\overline{B}} = \overline{\varnothing}$. Следовательно, $\overline{A} + B = \Omega$ и $\overline{A}A + BA = A\Omega = A$. Последнее соотношение доказывает, что $AB = A$. Отсюда в силу доказанного в задаче 18.19 утверждения следует, что $A \subset B$. Достаточность. Пусть $A \subset B$. Тогда $AB = A$ (см. задачу 18.19), т.е. $AB + \overline{A} = A + \overline{A} = \Omega$. Следовательно, $\overline{\overline{A} + AB} = \varnothing$ или $A \cdot \overline{AB} = \varnothing$, откуда $A(\overline{A} + \overline{B}) = A\overline{B} = A - B = \varnothing$. $\triangleright$ **18.34.** Указание. Обозначим $\omega_k = \{$среди $2n$ собравшихся ровно k человек знакомы вошедшему$\}$. $\Omega = \{\omega_0, \omega_1, \ldots, \omega_{2n}\}$. Показать, что $A + B = \Omega$. Для доказательства второго утверждения учесть, что $A - B = \overline{B}$. **18.35.** $E_1 = A\overline{B}\,\overline{C} + \overline{A}B\overline{C} + \overline{A}\,\overline{B}C$, $F_1 = AB\overline{C} + A\overline{B}C + \overline{A}BC$. **18.36.** $E_2 = A + B + C$, $F_2 = F_1 + ABC$. **18.37.** $E_3 = \overline{A}\,\overline{B}\,\overline{C}$, $F_3 = F_1 + ABC = F_2$, $G = \overline{A} + \overline{B} + \overline{C} = \overline{ABC}$. **18.38.** а) $\Omega = \{\omega_k,\ k = 1, 2, \ldots, 8\}$, где $\omega_1 = \overline{D}_1\overline{D}_2\overline{K}$, $\omega_2 = D_1\overline{D}_2\overline{K}$, $\omega_3 = \overline{D}_1 D_2\overline{K}$, $\omega_4 = \overline{D}_1\overline{D}_2 K$, $\omega_5 = \overline{D}_1 D_2 K$, $\omega_6 = D_1\overline{D}_2 K$, $\omega_7 = D_1 D_2\overline{K}$, $\omega_8 = D_1 D_2 K$. б) $A = D_1 D_2 + K = \omega_4 + \omega_5 + \cdots + \omega_8$. в) $\triangleleft$ $\omega_4 + \omega_5 + \cdots + \omega_8 = D_1 D_2(K + \overline{K}) + K(D_1\overline{D}_2 + \overline{D}_1\overline{D}_2 + \overline{D}_1\overline{D}_2 + \overline{D}_1 D_2) = D_1 D_2 + K(\overline{D}_1 + \overline{D}_2 + D_1 D_2) = D_1 D_2 + K(\overline{D_1 D_2} + D_1 D_2) = D_1 D_2 + K$. $\triangleright$ **18.39.** $B = A_1 + A_2 A_3 + A_4$. **18.40.** $B = (A_1 + A_2 + A_3(A_4 + A_5)$. **18.41.** Множество Ω и указанные события изо-

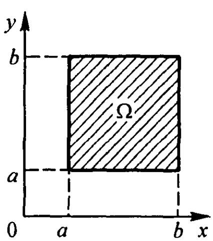


Рис. 50

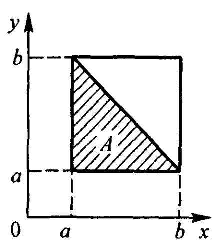


Рис. 51

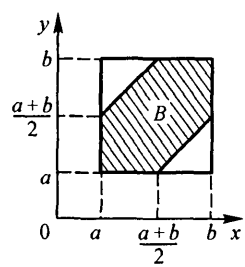


Рис. 52

бражены на рис. 50–55. **18.42.** а) $\Omega = \{\omega_1, \omega_2, \ldots, \omega_8\} = \{\overline{A}_1\overline{A}_2\overline{A}_3,\ A_1\overline{A}_2\overline{A}_3,\ \overline{A}_1 A_2\overline{A}_3,\ \overline{A}_1\overline{A}_2 A_3,\ A_1 A_2\overline{A}_3,\ A_1\overline{A}_2 A_3,\ \overline{A}_1 A_2 A_3,\ A_1 A_2 A_3\}$; б) $A = \omega_2 + \omega_3 + \omega_4 = A_1\overline{A}_2\overline{A}_3 + \overline{A}_1 A_2\overline{A}_3 + \overline{A}_1\overline{A}_2 A_3$, $B = \omega_2 + \cdots + \omega_8 =$

$= \overline{\omega}_1 = \overline{\overline{A}_1\overline{A}_2\overline{A}_3} = A_1 + A_2 + A_3$, $C = \omega_1 + \dots + \omega_7 = \overline{\omega}_8 = \overline{A_1A_2A_3} = \overline{A}_1 + \overline{A}_2 + \overline{A}_3$, $D = \omega_5 + \omega_6 + \omega_7 + \omega_8 = A_1A_2\overline{A}_3 + A_1\overline{A}_2A_3 + \overline{A}_1A_2A_3 + A_1A_2A_3$, $E = \omega_1 + \omega_4 = \overline{A}_1\overline{A}_2\overline{A}_3 + \overline{A}_1\overline{A}_2A_3 = \overline{A}_1\overline{A}_2$. **18.43.** а) $\Omega = \{\omega_1, \overline{\omega}_1\omega_2, \overline{\omega}_1\overline{\omega}_2\omega_3, \dots, \overline{\omega}_1 \dots \overline{\omega}_7\omega_8\}$. б) $A = \overline{\omega}_1\overline{\omega}_2\omega_3 + \overline{\omega}_1\overline{\omega}_2\overline{\omega}_3\omega_4 + \dots + \overline{\omega}_1 \dots \overline{\omega}_7\omega_8 = \overline{\omega}_1\overline{\omega}_2$. **18.44.** $\Omega = \{A_1, \overline{A}_1B_1, \overline{A}_1\overline{B}_1A_2, \dots\}$,

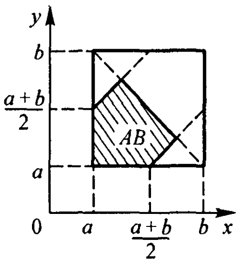


Рис. 53

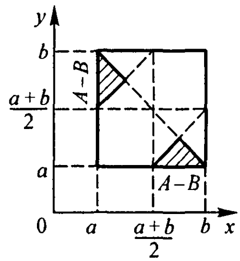


Рис. 54

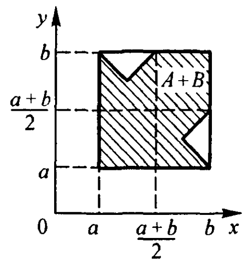


Рис. 55

$A = A_1 + \overline{A}_1\overline{B}_1A_2 + \dots$, $B = \overline{A}_1B_1 + \overline{A}_1\overline{B}_1\overline{A}_2B_2 + \dots$ **18.46.** Нет. **18.48.** 16. **18.52.** Использовать тождество $\varnothing + \Omega = \Omega$. **18.54.** Указание. Воспользоваться результатом предыдущей задачи. **18.55.** $\lhd$ Используя тождество $A + B = A + \overline{A}B$ (см. задачу 18.20в), находим $\mathrm{P}(A + B) = \mathrm{P}(A) + \mathrm{P}(\overline{A}B)$, так как события A и $\overline{A}B$ несовместны. Из тождества $B = AB + \overline{A}B$ следует $\mathrm{P}(B) = \mathrm{P}(AB) + \mathrm{P}(\overline{A}B)$ в силу несовместности событий AB и $\overline{A}B$. Отсюда $\mathrm{P}(\overline{A}B) = \mathrm{P}(B) - \mathrm{P}(AB)$ и $\mathrm{P}(A + B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(AB)$. $\rhd$ **18.59.** Указание. Использовать результаты задач 18.20б и 18.55. **18.60.** Указание. Использовать результаты задач 18.20 и 18.55. **18.64.** Указание. Положить $B + C = D$ и применить дважды формулу сложения для двух событий (см. задачу 18.55). **18.65.** $\lhd$ Из условия задачи следует, что $AB \subset C$, откуда $\mathrm{P}(C) \geqslant \mathrm{P}(AB)$ (см. пример 6). Отсюда получаем $\mathrm{P}(C) \geqslant \mathrm{P}(AB) = \mathrm{P}(\overline{\overline{A} + \overline{B}}) = 1 - \mathrm{P}(\overline{A}) - \mathrm{P}(\overline{B}) + \mathrm{P}(\overline{A}\,\overline{B}) = \mathrm{P}(A) + \mathrm{P}(B) - 1 + \mathrm{P}(\overline{A}\,\overline{B}) \geqslant \mathrm{P}(A) + \mathrm{P}(B) - 1$. $\rhd$ **18.66.** 5/6. **18.67.** 0,05. **18.68.** $\mathrm{P}(A) = 1/6$, $\mathrm{P}(B) = 1/3$, $\mathrm{P}(C) = 1/2$, $\mathrm{P}(D) = 2/3$, $\mathrm{P}(E) = 2/3$. **18.69.** $\mathrm{P}(A) = 1/6$, $\mathrm{P}(B) = 5/12$. **18.70.** $\mathrm{P}(C) = 1/2$, $\mathrm{P}(D) = 35/36$. **18.71.** $\mathrm{P}(E) = 5/6$, $\mathrm{P}(F) = 11/36$, $\mathrm{P}(G) = 1/9$. **18.72.** $\mathrm{P}(A) = 0{,}01$, $\mathrm{P}(B) = 0{,}2$, $\mathrm{P}(C) = 5/144$. **18.73.** 1/720. **18.74.** $1/m^{n-1}$. **18.75.** $\mathrm{P}(A) = 1 - 1/6^4$, $\mathrm{P}(B) = 5!/6^4$. **18.76.** 7/9. **18.77.** 1/28. **18.78.** $1 - C_{21}^7/C_{28}^7 \approx 0{,}932$. **18.79.** $\mathrm{P}(A) = 1/2$, $\mathrm{P}(B) = 1/42$. **18.80.** $\mathrm{P}(A) = 1/143$, $\mathrm{P}(B) = 2/91$, $\mathrm{P}(C) = 12/143$. **18.81.** $\mathrm{P}(A) = \dfrac{C_{m_1}^m}{C_{m_1+m_2}^m}$, $\mathrm{P}(B) = \dfrac{C_{m_1}^k C_{m_2}^{m-k}}{C_{m_1+m_2}^m}$.

18.82. $\mathrm{P}(C) = 1 - \dfrac{C_{m_2}^m}{C_{m_1+m_2}^m}$, $\mathrm{P}(D) = \dfrac{1}{C_{m_1+m_2}^m} \displaystyle\sum_{n=k}^{m} C_{m_1}^n C_{m_2}^{m-n}$. **18.83.** $q_n = \dfrac{n}{2n-1}$. **18.84.** $\mathrm{P}(A) \approx 0{,}264 \cdot 10^{-2}$, $\mathrm{P}(B) \approx 0{,}2813$. **18.85.** $\mathrm{P}(A) =$

$= \dfrac{C_{m_1}^{n_1} C_{m_2}^{n_2} \ldots C_{m_s}^{n_s}}{C_{m_1+m_2+\dots+m_s}^{n}}$, $\mathrm{P}(B) = \dfrac{C_{m_3+m_4+\dots+m_s}^{n}}{C_{m_1+m_2+\dots+m_s}^{n}}$. **18.86.** $C_n^m/2^n$. **18.87.** $1/9!$. **18.88.** $\mathrm{P}(B) = 1/9$, $\mathrm{P}(C) = 1/12$. **18.89.** $\mathrm{P}(D) = 1/126$, $\mathrm{P}(E) = 1/945$. **18.90.** $2/7$. **18.91.** а) $1/4$; б) $1/6$. **18.92.** $\mathrm{P}(A) = 1/60$, $\mathrm{P}(B) = 2/5$. **18.93.** $\mathrm{P}(C) = 1/20$, $\mathrm{P}(D) = 2/5$, $\mathrm{P}(E) = 9/10$. **18.94** б) $\dfrac{(n-m)(n-m-1)}{n(n-1)}$; в) C_m^k/C_n^k.

18.95. $\mathrm{P}(A) = \dfrac{2(n!)^2}{(2n)!}$, $\mathrm{P}(B) = \dfrac{(n+1)(n!)^2}{(2n)!}$ **18.96.** $\mathrm{P}(A) = 1/45$, $\mathrm{P}(B) = 7/45$, $\mathrm{P}(C) = 1/15$. **18.97.** $1/11$. **18.98.** а) $1/30$; б) $1/6$; в) $1/10$.

18.99. $p_n = \dfrac{4}{n+3}$; $p_6 = 4/9$, $p_9 = 1/3$. **18.100.** $\mathrm{P}(A) = \left(\dfrac{5}{6}\right)^{10} \approx 0{,}1615$, $\mathrm{P}(B) \approx 0{,}8385$, $\mathrm{P}(C) \approx 0{,}155$. **18.101.** $0{,}0016$. **18.102.** $1/4$. **18.103.** $\mathrm{P}(A) = 0{,}001$, $\mathrm{P}(B) \approx 0{,}0605$. **18.104.** $\mathrm{P}(C) = 0{,}1$, $\mathrm{P}(D) \approx 2{,}1 \cdot 10^{-5}$.

18.105. $\mathrm{P}(A) = 1/216$, $\mathrm{P}(B) = \dfrac{5}{48} \approx 0{,}16$, $\mathrm{P}(C) = 5/324$, $\mathrm{P}(D) = 1/6^{-5}$.

18.106. $\mathrm{P}(A) = 1/64$, $\mathrm{P}(B) = 1/64$, $\mathrm{P}(C) = 29/32$. **18.107.** Указание. Представить данный эксперимент как последовательность испытаний, каждое из которых состоит в случайном заполнении очередного ящика, и воспользоваться формулой для числа элементов прямого произведения множеств. **18.108.** $N(\Omega) = 4200$, $\mathrm{P}(A) = 2/15$. **18.109.** $30/143$. **18.110.** $\mathrm{P}(A) = 14/323$, $\mathrm{P}(B) = 125/969$. **18.111.** а) $\triangleleft$ n_1 элементов e_1 можно разместить на n мест $C_n^{n_1}$ способами, при этом n_2 элементов e_2 можно разместить на $n - n_1$ оставшихся мест $C_{n-n_1}^{n_2}$ способами и т.д. Совмещая все эти возможности, получим $N(\Omega) = C_n^{n_1} C_{n-n_1}^{n_2} \ldots C_{n-n_1-\dots-n_{s-1}}^{n_s} = \dfrac{n!}{n_1! n_2! \ldots n_s!}$. $\triangleright$

б) $p = \dfrac{n_2! n_3! \ldots n_s!}{n!} \times \dfrac{(n-n_1)!}{n_2! n_3! \ldots n_s!} = \dfrac{1}{C_n^{n_1}}$. **18.112.** $\mathrm{P}(A) \approx 0{,}0013$, $\mathrm{P}(B) \approx 0{,}0154$, $\mathrm{P}(C) \approx 0{,}3215$, **18.113.** $2!3!2!/10! \approx 6{,}6 \cdot 10^{-6}$.

18.114. $\mathrm{P}(A) = \dfrac{4!48!(13!)^4}{(12!)^4 52!} \approx 0{,}105$, $\mathrm{P}(B) = \dfrac{16 \cdot 39!(13!)^4}{(13!)^3 52!} \approx 8{,}4 \cdot 10^{-12}$.

18.115. $\mathrm{P}(C) \approx 0{,}01056$, $\mathrm{P}(D) \approx 0{,}0552$. **18.116.** Сочетания: $\Omega = \{ab, ac, bc\}$, $N(\Omega) = 3 = C_3^2$ (выбор без возвращения и без упорядочивания); размещения: $\Omega = \{ab, ba, ac, ca, bc, cb\}$, $N(\Omega) = 6 = A_3^2$ (выбор без возвращения и с упорядочиванием); сочетания с повторениями; $\Omega = \{aa, bb, cc, ab, ac, bc\}$, $N(\Omega) = 6 = C_4^2$ (выбор с возвращением и без упорядочивания); размещения с повторениями; $\Omega = \{aa, bb, cc, ab, ba, ac, ca, bc, cb\}$, $N(\Omega) = 9 = 3^2$ (выбор с возвращением и с упорядочиванием). **18.117.** $E = \{aa, ab, ac, ba, bb, bc\}$, $p = 1/3$. **18.118.** C_n^k/n^k. **18.119.** $\dfrac{(m-1)\ldots(m-n+1)}{m^{n-1}}$. **18.120.** а) $1/n$; б) $1/n(1-1/n)^{k-1}$. **18.121.** а) $3/55$; б) $12/55$. **18.122.** а) $\mathrm{P}(A) =$

$= 1 - \dfrac{C_{n-m}^k}{C_n^k}$, если $k \leqslant n-m$; 1, если $k > n-m$; $\mathrm{P}(B) = \dfrac{C_m^1 C_{n-m}^{k-1}}{C_n^k}$, если $k \leqslant n-m+1$; 0, если $k > n-m+1$; $\mathrm{P}(C) = \dfrac{C_m^{k_1} C_{n-m}^{k-k_1}}{C_n^k}$, если $k \leqslant n-m+k_1$; 0, если $k > n-m+k_1$. **18.123.** $p_n = 1 - \dfrac{A_{365}^n}{365^n}$, $p_{24} \approx 0{,}538$, $p_{50} \approx 0{,}97$. **18.125.** $\mathrm{P}(A) = 10^{-8}$, $\mathrm{P}(B) = 10^{-7}$. **18.126.** $\mathrm{P}(C) = 28A_9^6 \cdot 10^{-7} \approx 0{,}17$, $\mathrm{P}(D) = 378A_8^4 \cdot 10^{-6} \approx 0{,}64$. **18.127.** $\mathrm{P}(E) = 56A_9^5 \cdot 10^{-7} \approx 0{,}08$, $\mathrm{P}(F) = C_{10}^3 \cdot 3^8 \cdot 10^{-8} \approx 0{,}008$. **18.128.** $\mathrm{P}(A) = 25/546$, $\mathrm{P}(B) = 24/91$. **18.129.** а) $1/2$; б) $\dfrac{n-1}{2n}$. **18.130.** а) $1/3$; б) $\dfrac{(n-1)(n-2)}{3n^2}$. **18.131.** $\mathrm{P}(A) = \dfrac{2^n}{(2n)!}$, $\mathrm{P}(B) = \dfrac{n!}{1 \cdot 3 \cdot \ldots \cdot (2n-1)}$. **18.132.** $N(\Omega) = 10^6$, $p = 0{,}72 10^{-3}$. **18.133.** $N(\Omega) = A_{10}^6$, $p = 1/210$. **18.134.** $\mathrm{P}(A) = \dfrac{1}{m^{n-1}}$, $\mathrm{P}(B) = \left(1 - \dfrac{1}{m}\right)^n$. **18.135.** $\mathrm{P}(C) = \dfrac{A_n^{m_1}(m-m_1)^{n-m_1}}{m^n}$, $\mathrm{P}(D) = \dfrac{(m-1)^{n-n_1}}{m^{n-1}}$. **18.136.** $\mathrm{P}(E) = 9/64$, $\mathrm{P}(F) = 27/128$. **18.137.** $\mathrm{P}(E) = 4/27$, $\mathrm{P}(F) = 2/9$. **18.138.** $(3/4)^n$. Указание. Учесть, что в состав всех подмножеств конечного множества E входит и пустое множество, поэтому общее число всех возможных пар равно $N(\Omega) = 4^n$. Пусть S_k — число всех пар подмножеств, в которых первое подмножество имеет ровно k элементов. Число таких пар $S_k = C_n^k \cdot 2^{n-k}$. Кроме того, необходимо учесть, что пересекаются только те множества, которые имеют общие элементы (следовательно, пустые множества не пересекаются), поэтому если событие $A = \{$два произвольно взятых подмножества множества E пересекаются$\}$, то $N(A) = \sum\limits_{k=0}^{n} S_k$.

$$\textbf{18.139.}\ \mathrm{P}(B) = \begin{cases} \dfrac{\pi a^2}{4}, & 0 < a \leqslant 1, \\ \sqrt{a^2-1} + a^2\left(\dfrac{\pi}{4} - \arccos\dfrac{1}{a}\right), & 1 < a \leqslant \sqrt{2}, \\ 1, & \sqrt{2} < a. \end{cases}$$

$$\textbf{18.140.}\ \mathrm{P}(B) = \begin{cases} a(1 - \ln a), & 0 < a \leqslant 1, \\ 1, & 1 < a. \end{cases}$$

$$\textbf{18.141.}\ \mathrm{P}(C) = \begin{cases} a^2, & 0 < a \leqslant 1, \\ 1, & 1 < a; \end{cases} \qquad \mathrm{P}(D) = a(2-a).$$

18.142. $2/3$. **18.143.** $5/6$. **18.144.** $\alpha/2\pi$. **18.145.** $\dfrac{a}{l}\left(2 - \dfrac{a}{l}\right)$. **18.146.** $\sin^2\dfrac{\alpha}{2}$.

18.147. а) $\approx 0{,}08$; б) $\dfrac{\sqrt{3}}{4} \approx 0{,}433$, в) $\dfrac{\pi}{12}$. **18.148.** $\mathrm{P}(A) = 7/16, \mathrm{P}(C) = 7/32$. **18.149.** $\mathrm{P}(D) = 1/4$, $\mathrm{P}(E) = 9/32$, $\mathrm{P}(F) = 1/24$. **18.150.** 139/1152. **18.151.** а) $1 - \dfrac{1}{2}\left(1 - \dfrac{t_1}{T}\right)^2 - \dfrac{1}{2}\left(1 - \dfrac{t_2}{T}\right)^2$; б) $1 - \dfrac{1}{2}\left(1 - \dfrac{t_1 - t_3}{T}\right)^2 - \dfrac{1}{2}\left(1 - \dfrac{t_2}{T}\right)^2$. **18.152.** $\mathrm{P}(A) = 2/3$, $\mathrm{P}(B) = 1/12$. Указание. Показать, что $A = \{(a, b) \mid |a| \leqslant 1, b \leqslant 1, a^2 \geqslant b\}$ и $B = \{(a, b) \mid -1 \leqslant a \leqslant 0, 0 \leqslant b \leqslant 1, a^2 \geqslant b\}$. **18.153.** $\dfrac{h}{\sqrt{h^2 + 4r^2}}$. Указание. Положение цилиндра полностью характеризуется положением оси цилиндра при различных поворотах вокруг неподвижного центра масс. Достаточно проследить за значениями какой-либо иной, помимо центра масс, точки оси цилиндра, например центра круга, лежащего в основании цилиндра. При этом множество всех возможных значений данной точки представляет собой полусферу радиуса $h/2$, а множество тех значений, при которых цилиндр упадет на боковую поверхность, — сферический пояс. **18.154.** 1/4. Указание. Пусть длины отрезков x, y, $l - x - y$. Показать, что область, соответствующая искомому событию, выделяется на плоскости условиями $x < l/2$, $y < l/2$, $x + y < l/2$. **18.155.** $\sqrt{5}\, d/(\sqrt{5} - \sqrt{2})$. Указание. Воспользоваться результатом примера 11. **18.156.** 1/4. Указание. Обозначим событие $D = \{\Delta ABC$ остроугольный$\}$. Пусть точки A, B и C фиксированы. Условимся измерять длины дуг между точками в таком направлении, чтобы при движении по окружности за точкой A следовала точка B, а за точкой B — точка C. Показать, что если при таком способе измерения x — длина дуги AB, y — длина дуги BC, то $\Omega = \{(x, y) \mid 0 \leqslant x < 2\pi, 0 \leqslant y < 2\pi, x + y < 2\pi\}$, $D = \{(x, y) \mid 0 < x < \pi, 0 < y < \pi, 2\pi - (x + y) < \pi\}$. **18.157.** $\dfrac{2}{\pi} \arccos \dfrac{r_1}{r_2}$. **18.158.** $\dfrac{2}{9} \ln 2 + \dfrac{1}{12} \approx 0{,}237$. **18.159.** $\dfrac{2l}{(\pi/a)}$. Указание. Положение иглы характеризуется двумя координатами (r, φ), где r — расстояние от центра иглы до ближайшей прямой линии, φ — угол между направлением иглы и прямой линией. **18.159(1).** 1) 1/4. 2) 1/3. 3) 1/2. **18.163.** 1/3. **18.164.** 1/4. **18.165.** 3/4. **18.166.** $\mathrm{P}(A_1/C) = 11/32$, $\mathrm{P}(A_2/C) = 9/16$. **18.167.** 1/3. Указание. Учесть, что множество элементарных исходов Ω можно представить в виде $\Omega = \Omega_1 \times \Omega_2$, где $\Omega_i = \{M, Д\}$, $i = 1, 2$. **18.168.** 1/3. **18.169.** $\mathrm{P}(B/A) = 0{,}5$, $\mathrm{P}(A/B) = 60/91$. **18.173.** Не следует. **18.174.** Указание. Показать сначала, что независимы события A и $\overline{B}$. **18.176.** A и B, F и B независимы, A и F зависимы. **18.177.** Зависимы. $\mathrm{P}(D/C) = 9/16$. **18.179.** а) E и F независимы; б) не являются. **18.181.** 2/9. **18.182.** 48/95. **18.183.** 8/9. **18.184.** 20/21. **18.185.** 1/4. **18.187.** а) 0,216; б) 1/6. **18.188.** $\mathrm{P}(BF) = 2/9$, $\mathrm{P}(AF) = 1/9$, $\mathrm{P}(ABF) = 1/18$. **18.189.** $1/n$. **18.190.** 5/8. **18.191.** $\mathrm{P}(A) = (1 - p_1)(1 - p_2)$, $\mathrm{P}(B) = p_1 + p_2 - 2p_1p_2$. **18.192.** 0,664.

18.193. $\mathrm{P}(A) = 0{,}24$, $\mathrm{P}(\overline{A}\,\overline{B}) = 0{,}48$, $\mathrm{P}(\overline{A} + \overline{B}) = 0{,}88$, $\mathrm{P}(A\Delta B) = 0{,}4$. **18.194.** $\mathrm{P}(B + F) = 13/18$, $\mathrm{P}(F - A) = 1/3$, $\mathrm{P}(F - AB) = 7/18$. **18.195.** $1-(1-p_1^2)(1-p_2)$. **18.196.** $\mathrm{P}(A) = 0{,}8$, $\mathrm{P}(B) = 0{,}4$, $\mathrm{P}(C) = 0{,}6$. **18.197.** $228/253 \approx 0{,}901$. **18.198.** $0{,}104$. **18.199.** $\mathrm{P}(A) = 1/2$, $\mathrm{P}(B) = 3/4$, $\mathrm{P}(C) = 1/2$, $\mathrm{P}(D) = 25/36$. **18.200.** $0{,}22$. **18.201.** $\mathrm{P}(A) = 0{,}729$, $\mathrm{P}(B) = 0{,}972$, $\mathrm{P}(C) = 0{,}891$. **18.202.** $(p_1p_2+(1-p_1)(1-q_2))(q_1q_2+(1-q_1)(1-p_2))$. **18.203.** $n \geqslant 25$. **18.204.** $n = \left[\dfrac{\ln(1-p_2)}{\ln(1-p_1)} + 1\right]$, где $[x]$ — целая часть числа x; $n(0{,}3;\ 0{,}9) = 7$, $n(0{,}3;\ 0{,}95) = 9$. **18.205.** $1-3p_2p_3^2$. **18.206.** С мелкой мишени. **18.206(1).** В отношении $7:1$. **18.207.** $\mathrm{P}(A) = 1-(1-pp_1)^{n-1}$, $\mathrm{P}(B) = 1-(1-pp_1)^{n-k}$. **18.208.** $1-q_1q_2q_3$. **18.209.** $1 - (1-p_1p_2p_3)(1-p_4p_5p_6)$. **18.210.** $p_1p_4(1-q_2q_3)$. **18.211.** $(1-q_1q_2)(1-q_3q_4)$. **18.212.** $p_5(1-q_1q_2)(1-q_3q_4) + q_5(p_1p_3+p_2p_4-p_1p_2p_3p_4)$. **18.213.** $\dfrac{p(m+n)-p(m)}{1-p(m)}$. **18.214.** $p_1 = 2/3$, $p_2 = 1/3$. **18.215.** $p_1 = \dfrac{2}{3}\left(1-\left(\dfrac{1}{2}\right)^{20}\right) \approx \dfrac{2}{3}$, $p_2 = \dfrac{1}{3} + \dfrac{2}{3}\left(\dfrac{1}{2}\right)^{20} \approx \dfrac{1}{3}$. **18.216.** $p_1 = 4/7$, $p_2 = 3/7$. **18.217.** $\mathrm{P}(A) = 7/9$, $\mathrm{P}(B) = 2/9$. **18.218.** $\mathrm{P}(C) = 5/9$, $\mathrm{P}(D) = 4/9$. **18.219.** Могут существовать 0, 1, 2, 3, 4 амебы соответственно с вероятностями $11/32$, $4/32$, $9/32$, $4/32$, $4/32$. **18.220.** Указание. Использовать метод математической индукции. **18.221.** $p_n = \sum\limits_{k=1}^{n}\dfrac{(-1)^{k-1}}{k!}$, $p_5 = \dfrac{19}{30}$, $p_{10} = 1-e^{-1}+R_{11} \approx 0{,}649$, где R_n — остаточный член разложения в ряд Тейлора функции e^x в точке $x=-1$. Указание. Пусть событие $A_k = \{k$-й адресат получил свое письмо$\}$. Тогда $\mathrm{P}(A_k) = \dfrac{(n-1)!}{n!}$, $\mathrm{P}(A_kA_m) = \dfrac{(n-2)!}{n!}$ $(k \neq m), \ldots, \mathrm{P}(A_1 \ldots A_n) = \dfrac{1}{n!}$. Для вычисления искомой вероятности $p = \mathrm{P}(A_1+A_2+\cdots+A_n)$ воспользоваться формулой сложения для n слагаемых. **18.222.** p. **18.223.** $p > 1/2$. **18.224.** p. **18.225.** $0{,}84$. **18.226.** $0{,}895$. **18.227.** $\dfrac{\alpha n_1+\beta n_2}{100(n_1+n_2)}$. **18.228.** $1-npq_1q^{n-1}-q^n$, где $q_1 = 1-p_1$, $q = 1-p$. **18.229.** $\approx 0{,}574$. **18.230.** $\dfrac{1}{2}(q_1(1-q_2)+(1-q_1)p_2+(1-p_1)(1-q_2)+p_1p_2)$. **18.231.** $67/120$. **18.232.** $0{,}452$. **18.233.** $0{,}445$. **18.234.** $0{,}763$. **18.235.** В одну урну поместить один белый шар, в другую — два белых и три черных. $P_{\max}(a) = 0{,}7$. **18.236.** В первый район следует послать m_0 геологоразведочных партий, где m_0 — ближайшее целое к числу $\dfrac{n}{2}+\dfrac{1}{2}\ln\dfrac{1-p_1}{p_1}$. Указание. Пусть событие $A = \{$хотя бы одна из n посланных геолого разведочных партий обнаружила нефть на заданной территории$\}$.

Показать, что $\mathrm{P}(A) = 1 - p_1(1-p)^m - (1-p_1)(1-p)^{n-m}$, где m — число геологоразведочных партий, посланных в первый район. Далее рассмотреть функцию $f(x) = 1 - p_1(1-p)^x - (1-p_1)(1-p)^{n-x}$ и найти ее максимум при $x \in [0, n]$. **18.237.** 0,5. **18.238.** $\dfrac{(n-m)(n-m+1)}{2n(n-1)}$.

18.239. 5/36. **18.239(1).** $\dfrac{4(n+1)}{(n+2)(n+3)}$. **18.240.** $p_1 \approx 0{,}457$, $p_2 \approx 0{,}32$, $p_3 \approx 0{,}223$. $\lhd$ Пусть событие $A_i = \{i\text{-й игрок выиграл}\}$, $i = 1, 2, 3$. Обозначим $\mathrm{P}(A_i) = p_i$. Очевидно, что

$$p_1 + p_2 + p_3 = 1. \tag{$*$}$$

Рассмотрим две гипотезы: $H_1 = \{$при первом подходе первый игрок не достал черный шар$\}$ и $H_2 = \overline{H}_1$. Тогда $\mathrm{P}(H_1) = 0{,}7$, $\mathrm{P}(H_2) = 0{,}3$, $\mathrm{P}(A_2/H_2) = 0$ и $\mathrm{P}(A_2/H_1) = \mathrm{P}(A_1) = p_1$, так как если первый игрок не достал черный шар, то второй оказывается в положении начинающего игру. Следовательно, $\mathrm{P}(A_2) = p_2 = \mathrm{P}(A_2/H_1)\,\mathrm{P}(H_1) + \mathrm{P}(A_2/H_2)\,\mathrm{P}(H_2) = 0{,}7p_1$. Вводя теперь гипотезы $H_1' = \{$при первом подходе как первый, так и второй игроки не достали черный шар$\}$ и $H_2' = \overline{H}_1'$, аналогично получаем $\mathrm{P}(A_3) = p_3 = 0{,}49p_1$. Подставляя полученные выражения в уравнение ($*$), находим $p_1 \approx 0{,}457$, $p_2 \approx 0{,}32$, $p_3 \approx 0{,}223$. $\rhd$ **18.241.** Шансы одинаковы. **18.242.** 2/3. **18.243.** $\approx 0{,}903$. **18.244.** $\mathrm{P}(A_1) \approx 0{,}285$, $\mathrm{P}(A_2) \approx 0{,}0714$. **18.245.** $\mathrm{P}(A/B) = 2/9$. **18.246.** 1/3.

18.247. $\dfrac{p_1p_3(1-p_2)}{(1-p_1)\,p_2p_3 + (1-p_2)\,p_1p_3 + (1-p_3)\,p_1p_2}$.

18.248. $\dfrac{n_1p_1}{n_1p_1 + n_2p_2 + n_3p_3}$. **18.249.** $\approx 0{,}29$. **18.250.** $\mathrm{P}(H_1/A) \approx 0{,}8677$, $\mathrm{P}(H_2/A) \approx 0{,}1052$, $\mathrm{P}(H_3/A) \approx 0{,}0271$. **18.251.** $135/139 \approx 0{,}971$. **18.252.** H_1. **18.253.** $p(1-p)^2/4$. **18.254.** 1/3. **18.255.** $\approx 0{,}0826$. **18.256.** $\approx 0{,}467$. **18.258.** $m_X = 45/16$, $\mathrm{P}\{X > 2\} = 11/16$. **18.259.** $D_X = 167/256$, $d_X = 3$. **18.260.** График функции распределения изображен на рис. 56.

18.261.

x_i	0	1
$\mathrm{P}\{X = x_i\}$	$1-p$	p

,

$$F_X(x) = \begin{cases} 0, & x \leqslant 0; \\ q, & 0 < x \leqslant 1, \\ 1, & x > 1; \end{cases} \qquad m_X = p.$$

18.262. $D_X = pq$, $\mu_3 = pq(q-p)$, $p = 1/2$. **18.263.** $c = 3/4$; график функции распределения изображен на рис. 56, $m_X = -1/3$, $h_X = 1-\sqrt{2}$.

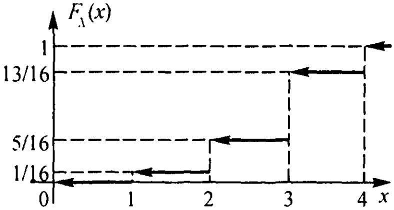


Рис. 56

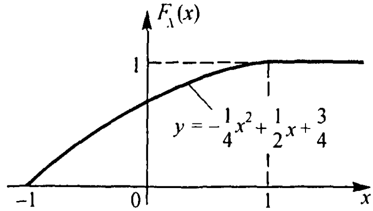


Рис. 57

18.264. $P\{X > 0\} = 1/4$, $P\{-1/2 < X \leqslant 1/2\} = 1/2$. **18.265.** $P\{X \geqslant 3{,}5\} = 1/2$, $P\{|X| < 2{,}5\} = 0{,}3$.

18.266.

x_i	2	3	4
$P\{X = x_i\}$	0,3	0,2	0,5

$$m_X = 3{,}3, \quad D_X = 0{,}76.$$

18.267.

x_i	0	1	2
$P\{X = x_i\}$	q_1q_2	$p_1q_2 + p_2q_1$	p_1p_2

$$m_X = p_1 + p_2, \quad D_X = p_1q_1 + p_2q_2.$$

18.268.

x_i	-1	0	1
$P\{X = x_i\}$	p_1	p_2	p_3

$$p_1 = \left(\frac{2}{3}\right)^3 = \frac{8}{27}, \quad p_2 = \frac{61}{216}, \quad p_3 = \frac{91}{216}.$$

$$F_X(x) = \begin{cases} 0, & x \leqslant -1, \\ \dfrac{8}{27}, & -1 < x \leqslant 0, \\ \dfrac{125}{216}, & 0 < x \leqslant 1, \\ 1, & x > 1, \end{cases} \qquad m_X = \frac{1}{8}, \quad d_X = 1.$$

18.269. $c = 1/2$, $P\{|X| < \pi/4\} = \sqrt{2}/2$, $m_X = 0$, $D_X = \pi^2/4 - 2$. **18.270.** $(2+\sqrt{2})/4$. **18.271.** $P\{X \geqslant 1\} = 3/4$, $m_X = 4/3$, $h_X = \sqrt{2}$, $D_X = 2/9$.

18.272.

x_i	1	2	3	4	5
p_i	q	pq	p^2q	p^3q	p^4

$$d_X = \begin{cases} 1, & p < p_0, \\ 5, & p \geqslant p_0, \end{cases} \qquad p_0 \approx 0{,}725, \quad m_X = \frac{1-p^5}{1-p}.$$

18.273. $\mu_n = \sum\limits_{i=0}^{n}(-1)^i C_n^i \alpha_1^i \alpha_{n-i}$. **18.274.** $\alpha_n = \sum\limits_{i=0}^{n} C_n^i m_x^i \mu_{n-i}$.

18.275. Указание. Использовать соотношение, справедливое для С. В. Д. Т. $\mathrm{P}\{x = x_k\} = F_x(x_{k+1}) - F_x(x_k)$, и преобразовать сумму, выражающую математическое ожидание. **18.276.** Указание. Показать, что математическое ожидание может быть записано в виде $m_x = -\lim\limits_{a\to+\infty}\int\limits_0^a x\,\dfrac{d}{dx}[1 - F_x(x)]\,dx$. При вычислении интеграла воспользоваться неравенством $a\int\limits_a^{+\infty} f(x)\,dx \leqslant \int\limits_a^{+\infty} x f_x(x)\,dx$, вытекающим из конечности математического ожидания.

18.278.

x_i	0	1	2	3
$\mathrm{P}\{X = x_i\}$	1/6	1/2	3/10	1/30

18.279. $m_x = 6/5$, $D_x = 14/25$. **18.280.** 11/42. **18.281.** $p_6 \approx 1{,}22\cdot 10^{-7}$, $p_5 \approx 2{,}85\cdot 10^{-5}$. **18.282.** $m_x = 5/2$, $D_x = 25/12$.

18.283. $F_x(x) = \begin{cases} 0, & x \leqslant 0, \\ x/5, & 0 < x \leqslant 5, \\ 1, & 5 < x, \end{cases}$ $\mathrm{P}\{X > 3\} = 2/5$.

18.284. $m_x = \dfrac{a+b}{2}$, $D_x = \dfrac{(b-a)^2}{12}$. **18.285** 1/3. **18.286** а) $\dfrac{1}{\sqrt{3}} \approx 0{,}577$; б) $1 - \dfrac{1}{\sqrt{3}} \approx 0{,}423$. **18.287.**

$$f_x(x) = \begin{cases} \dfrac{1}{a}\left(1 - \dfrac{|x|}{a}\right), & |x| < a; \\ 0, & |x| \geqslant a; \end{cases} \quad F_x(x) = \begin{cases} 0, & x \leqslant -a, \\ \dfrac{(a+x)^2}{2a^2}, & -a < x \leqslant 0, \\ 1 - \dfrac{(a-x)^2}{2a^2}, & 0 < x \leqslant a, \\ 1, & x > a. \end{cases}$$

18.288. $m_x = d_x = h_x = 0$, $D_x = a^2/6$, $e_x = -3/5$. **18.289.** $m_x = 1/\lambda$, $D_x = 1/\lambda^2$. **18.290.** $\mathrm{P}\{X \geqslant m_x\} = e^{-1} \approx 0{,}368$, $m_x = t_{0{,}632}$. **18.291.** $\mathrm{P}(A) = \exp(-1/2) - \exp(-3/2) \approx 0{,}3834$, $\mathrm{P}(B) = \exp(-2) \approx 0{,}135$. **18.292.** $F_x(x) = \begin{cases} 0, & x \leqslant 0, \\ 1 - e^{-\lambda x}, & x > 0. \end{cases}$ Указание. Записать в алгебре событий дифференциал $dF_x(x) = \mathrm{P}\{x \leqslant X < x + dx\}$ при $x > 0$ и решить полученное дифференциальное уравнение для искомой функции распределения с начальным условием $F_x(0) = 0$. **18.293.** $\mu_{k+1} = (-1)^{k+1} m_x^{k+1} + \dfrac{k+1}{\lambda}\mu_k$ $(k = 0, 1, \ldots)$, $a_x = 2$, $e_x = 6$. **18.294.** $a > 0$,

$b = 1/2$, $c = 1/\pi$. **18.295.** $f_X(x) = \dfrac{a}{\pi\,(x^2 + a^2)}$, m_X и моменты более высокого порядка не существуют. **18.296.** $d_X = h_X = 0$, $t_{0,75} = a$.

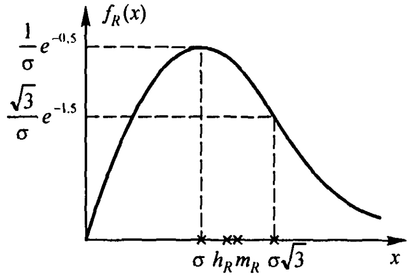


Рис. 58

18.297. График функции $f_R(x)$ изображен на рис. 58; $m_R = \sigma\sqrt{\pi/2} \approx 1{,}253\sigma$, $D_R = \sigma^2(2-\pi/2) \approx 0{,}429\sigma^2$. **18.298.** $d_R = \sigma$, $h_R = \sigma\sqrt{2\ln 2} \approx 1{,}177\sigma$, $a_R = 2\sqrt{\dfrac{\pi}{4-\pi}} \cdot \dfrac{\pi - 3}{4-\pi} \approx 0{,}631$, $d_R < h_R < m_R$. **18.299.** $m_V = 2\sqrt{2/(\beta\pi)}$, $d_V = \sqrt{2/\beta}$, $D_V = \dfrac{1}{\beta}\left(3 - \dfrac{8}{\pi}\right)$.

18.300.
$$F_X(x) = \begin{cases} 0, & x \leqslant -a, \\ \dfrac{1}{2} + \dfrac{1}{\pi}\arcsin\dfrac{x}{a}, & |x| < a, \\ 1, & x > a; \end{cases} \qquad m_X = 0,\ D_X = \frac{a^2}{2}.$$

18.301. d_X не существует, $\hat{h}_X = 0$, $\varkappa_{0,75} = a\sin\dfrac{\pi}{8} \approx 0{,}3827a$. **18.302.** $m_X = m$, $\sigma_X = \sigma$. **18.303.**

$$F_X(x) = \begin{cases} \dfrac{1}{2}e^{x\sqrt{2}/\sigma}, & x \leqslant 0, \\ 1 - \dfrac{1}{2}e^{-x\sqrt{2}/\sigma}, & x > 0; \end{cases} \qquad p_k = 1 - e^{-k\sqrt{2}} \approx \begin{cases} 0{,}7569, & k = 1, \\ 0{,}94, & k = 2, \\ 0{,}9856, & k = 3. \end{cases}$$

18.304. $a_X = 0$, $e_X = 3$. Указание. Использовать интеграл $\displaystyle\int_0^{+\infty} x^n e^{-ax}\,dx = \frac{n!}{a^{n+1}}$. **18.305.** $m_X = \dfrac{a}{a-1}x_0$, если $a > 1$, $D_X = \dfrac{a}{(a-1)^2(a-2)}x_0^2$, если $a > 2$. **18.306.** $d_X = x_0$, $h_X = x_0\sqrt[a]{2}$, $t_{0,75} \approx 156{,}2$. **18.307.** $\mathrm{P}(A) \approx 0{,}1836$, $\mathrm{P}(B) \approx 0{,}9057$, $m_X = \varkappa_{0,3164}$. **18.308.** $m_X = a/b$, $D_X = a/b^2$, $a_X = 2/\sqrt{a}$, $e_X = 6/a$. Указание. При вычислении центральных моментов вывести предварительно рекуррентную формулу для начальных моментов гамма-распределения: $\alpha_{k+1} =$

$= \dfrac{a+k}{b}\alpha_k$ $(k = 0, 1, \ldots)$ и использовать результат задачи 18.273. **18.309.** $m_X = 4$, $D_X = 8$, $h_X \approx 3{,}3567$, $m_X = \varkappa_{0{,}406}$. Указание. Уравнение для отыскания медианы привести к виду $z = \ln 2 + \ln(1+z)$ $(h_X = 2z)$, который допускает итеративное решение, начиная, например, со значения $z_0 = 1$. **18.310.** $m_X = a + b\Gamma\left(1+\dfrac{1}{n}\right)$, $d_X = a + b\sqrt[n]{\dfrac{n-1}{n}}$. Указание. Использовать интеграл $\displaystyle\int_0^{+\infty} x^n e^{-(\alpha x)^m}\,dx = \dfrac{\Gamma\left(\dfrac{n+1}{m}\right)}{m\alpha^{n+1}}$. **18.311.** $m_X = \dfrac{a}{a+b}$, $D_X = \dfrac{ab}{(a+b)^2(a+b+1)}$. Указание. Использовать интеграл $\displaystyle\int_0^1 x^{a-1}(1-x)^{b-1}dx = \dfrac{\Gamma(a)\,\Gamma(b)}{\Gamma(a+b)}$, выражающий условие нормировки для любых допустимых значений параметров $a > 0$, $b > 0$, и свойство гамма-функции: $\Gamma(a+1) = a\Gamma(a)$. **18.312.** $P(A) \approx 0{,}3955$, $P(B) \approx 0{,}2637$. **18.313.** $P(C) \approx 0{,}7627$, $P(D) \approx 0{,}1035$. **18.314.** $\approx 0{,}6513$. **18.315.** $P(A) \approx 0{,}234$, $P(B) \approx 0{,}721$. **18.316.** $P(C) = (15/36)^7 \approx 0{,}00218$, $P(D) \approx 0{,}821$.

18.317.

x_i	0	1
p_i	q	p

$m_X = p$, $D_X = pq$. **18.318.** $P\{X \geqslant 2\} \approx 0{,}363$.

18.319. $$F_X(x) = \begin{cases} 0, & x \leqslant 0,\\ 27/125, & 0 < x \leqslant 1,\\ 81/125, & 1 < x \leqslant 2,\\ 117/125, & 2 < x \leqslant 3,\\ 1, & 3 < x.\end{cases}$$

18.320. $P(A) = 35/127$, $P(B) = 46/127$. **18.321.** $M(X) = 50$, $M(Y) = 75$. **18.322.** Более вероятно выиграть матч из 12 партий. **18.323.** $\approx 0{,}203$. **18.324.** $1 - (pt+q)^n$. **18.325.** $n \geqslant 59$. **18.326.** Более вероятны три дождливых дня. **18.328.** Один или два прибора. **18.329.** $\approx 0{,}5638$.

18.330.

x_i	1	2	3	...	n	...
p_i	p	qp	q^2p	...	$q^{n-1}p$	...

$$P\{X \leqslant 3\} = p(1+q+q^2).$$

18.331. $m_X = 1/p$, $D_X = q/p^2$. **18.332.** 1000 изделий, $P\{X > 3m_X\} \approx 0{,}0498$. **18.333.** $m_X = 1/q$, $d_X = 1$, $D_X = p/q^2$. Указание. Воспользоваться результатом задач 18.330 и 18.331. **18.334.** $P(A) = (n -$

$-k+1)p^kq^{n-k}$. **18.335.** $\mathrm{P}(B)=C_{m-1}^{k-1}p^kq^{m-k}$, $\mathrm{P}(C)=p^k\sum\limits_{i=0}^{m-1}C_{i+k-1}^{i}q^i$.

18.335(1). $\mathrm{P}(A)=\sum\limits_{k=0}^{m}p_kC_n^kp^kq^{n-k}$. **18.336.** $\mathrm{P}\{X=m\}=C_{m-1}^{k-1}p^kq^{m-k}$,

$m\geqslant k$, $m_X=k/p$. **18.337.** 23/648. **18.338.**

x_i	0	1	2
p_i	0,12	0,56	0,32

18.339. $m_X=1{,}2$, $D_X=0{,}4$. **18.340.** $m_X=1{,}5$, $D_X=0{,}65$. **18.341.** $\mathrm{P}(A)\approx\approx 0{,}4344$, $\mathrm{P}(B)\approx 0{,}9136$. **18.342.** а) $m_X=2$, $D_X=1{,}1$; б) 0,9328. **18.343.** 0,297. **18.344.** Указание. При выводе формул для $\mathrm{P}_{m,m}$ и $\mathrm{P}_{m,m-1}$ использовать формулу (8). При доказательстве формулы для $\mathrm{P}_{m,m-k}$ использовать формулу полной вероятности. **18.345.** 0,2144. Указание. Воспользоваться рекуррентной формулой из предыдущей задачи. **18.346.** $\mathrm{P}(A)=0{,}216$, $\mathrm{P}(B)=0{,}189$. **18.347.** $\mathrm{P}(C)=0{,}135$. **18.348.** $\mathrm{P}(A)\approx 0{,}012$, $\mathrm{P}(B)\approx 1{,}144\cdot 10^{-4}$. **18.349.** $\mathrm{P}(C)\approx 0{,}246$, $\mathrm{P}(D)\approx 2^{-10}\approx 10^{-3}$. **18.350.** $\mathrm{P}(A)\approx 0{,}055$, $\mathrm{P}(B)\approx 0{,}055$. **18.351.** $\mathrm{P}(C)\approx 0{,}0638$. **18.352.** $\mathrm{P}(A)\approx 0{,}394$, $\mathrm{P}(B)\approx 0{,}013$, $\mathrm{P}(C)\approx\approx 0{,}998$. **18.353.** $\mathrm{P}(A)\approx 0{,}018$, $\mathrm{P}(B)\approx 0{,}092$. **18.354.** $\mathrm{P}(C)\approx 0{,}18$, $\mathrm{P}(D)\approx 0{,}938$. **18.355.** $\mathrm{P}(A)=F_X(m)=e^{-\lambda t_1}\sum\limits_{k=0}^{m-1}\frac{(\lambda t_1)^k}{k!}$, $\mathrm{P}(B)=$ $=\frac{1}{2}(1+e^{-2\lambda t_2})$. **18.356.** Две опечатки с вероятностью $\approx 0{,}251$. **18.357.** 0,0047. **18.358.** Закон Пуассона с параметром $\lambda=np$, $m_X==D_X=np$. **18.359.** $n\geqslant 300$. **18.360.** $n\geqslant 475$. Указание. Получить уравнение для λ и привести его к виду, допускающему метод итераций. **18.361.** $F_X(x)=\Phi\left(\frac{x-1}{2}\right)$. **18.362.** $p_1\approx 0{,}683$, $p_2\approx 0{,}954$, $p_3\approx 0{,}997$.

18.363. а) $(-5;\ 25)$, б) $(0;\ 20)$, в) $(6{,}65;\ 13{,}35)$. **18.364.** 0,898. **18.365.** $m==15{,}39$, $\sigma=3{,}26$. **18.366.** $\approx 95\,\%$. **18.367.** $\approx 4{,}299$. **18.368.** $\approx 99{,}95\,\%$. **18.369.** $\approx 0{,}0196$. Указание. Относительной точностью изделия называется величина $\frac{|x-m_X|}{m_X}$. **18.370.** $2(1-\Phi(1/(2\alpha)))$. **18.371.** $\alpha\leqslant\leqslant 0{,}2146$. **18.372.** $\mathrm{M}[X^2]\approx 1{,}1842$, $\mathrm{P}\{X^2>2\}\approx 0{,}8328$. **18.373.** $p_1=$

$=0{,}1587$, $p_2=0{,}68$. **18.374.** $\sqrt{\dfrac{(b-a)\left(\dfrac{a+b}{2}-m\right)}{\ln\dfrac{b-m}{a-m}}}$. $\triangleleft$ $Q(\sigma)=$

$=\mathrm{P}\{a\leqslant X<b\}=\Phi\left(\frac{b-m}{\sigma}\right)-\Phi\left(\frac{a-m}{\sigma}\right)$. Необходимое условие

экстремума: $Q'(\sigma)=0$. Так как

$$Q'(\sigma) = -\frac{1}{\sigma^2}\left[(b-m)\,\Phi'(x)]\right|_{x=(b-m)/\sigma} - (a-m)\,\Phi'(x)]\Big|_{x=(a-m)/\sigma},$$

то искомое значение σ является корнем уравнения

$$(b-m)\exp\left\{-\frac{(b-m)^2}{2\sigma^2}\right\} - (a-m)\exp\left\{-\frac{(a-m)^2}{2\sigma^2}\right\} = 0,$$

решая которое, получим σ. ▷ **18.375.** $\sigma\sqrt{2/\pi}$. **18.376.** $a_X = e_X = 0$. **18.377.** $\mathrm{M}[X^4] = 10$, $\mathrm{M}[X^6] = 76$. Указание. Воспользоваться результатом задачи 18.274 и рекуррентным соотношением (12).

18.378. а)

x_i	1	2
$P\{X=x_i\}$	0,8	0,2

y_i	−1	0	1
$P\{Y=y_i\}$	0,2	0,35	0,45

б) зависимы; в) $P\{X=2,\ Y\geqslant 0\} = 0{,}15$, $P\{X>Y\} = 0{,}65$.

18.379.

x	y			
	$y \leqslant -1$	$-1 < y \leqslant 0$	$0 < y \leqslant 1$	$1 < y$
$x \leqslant 1$	0	0	0	0
$1 < x \leqslant 2$	0	0,15	0,45	0,8
$2 < x$	0	0,2	0,55	1

$$m_X = 1{,}2, \quad m_Y = 0{,}25.$$

18.380. $K = \begin{pmatrix} 0{,}16 & 0 \\ 0 & 0{,}1875 \end{pmatrix}$.

18.381.

y_j	−1	0	1
$P\{Y=y_i/X=1\}$	0,12	0,24	0,64

,

$$\mathrm{M}\{Y/X=1]=0{,}52\}.$$

18.382. $P\{X+Y>1=5/12\}$.

18.383.

x_i	y: 0	1	$\mathrm{P}\{X = x_i\}$
0	1/3	1/6	1/2
1	1/3	1/6	1/2
$\mathrm{P}\{Y = y_j\}$	2/3	1/3	

. **18.384.** $\mu_{2,2} = 1/18$.

18.385.

x_i	y: 0	1	$\mathrm{P}\{X = x_i\}$
0	2/15	4/15	2/5
1	4/15	1/3	3/5
$\mathrm{P}\{Y = y_j\}$	2/5	3/5	

.

18.386.

x_i	y: 0	1	$\mathrm{P}\{X = x_i\}$
0	0,16	0,24	0,4
1	0,24	0,36	0,6
$\mathrm{P}\{Y = y_j\}$	0,4	0,6	

Вероятность события $\{X > Y\}$ больше в опыте из задачи 18.385.

18.387.

x_i	y: −1	0	1
0	0,06	0,1	0,04
1	0,21	0,35	0,14
2	0,03	0,05	0,02

,

$F_{X,Y}(1{,}5;\ -0{,}5) = 0{,}27$
$F_{X,Y}(0{,}5;\ 4) = 0{,}2$.

18.388. Функция распределения может быть представлена следующей таблицей:

x	y		
	$y \leqslant 0$	$0 < y \leqslant 1$	$1 < y$
$x \leqslant 0$	0	0	0
$0 < x \leqslant 1$	0	$q_1 q_2$	q_1
$1 < x$	0	q_2	1

18.389.

x_i	y_j		
	0	1	2
0	0	0	2/15
1	0	8/15	0
2	1/3	0	0

, $\rho_{X,Y} = -1$.

18.390.

x_i	y_j		
	0	1	2
1	p^2	pq	0
2	0	pq	q^2

, $(m_X, m_Y) = (1 + q,\ 2q)$. $\sigma_X^2 = pq$, $\sigma_Y^2 = 2pq$.

18.391. $\mathrm{P}\{X \geqslant Y + 1\} = p - p_{11} - p_{12} = 1 - pq$.

18.392. $\rho_{XY} = 1/\sqrt{2} \approx 0{,}707$.

18.393.

x_i	y_j		
	1	3	5
2	0,25	0,05	0,1
4	0,15	0,3	0,15

, $(m_X, m_Y) = (3{,}2;\ 1{,}25)$.

18.394.

x_i	y_j		
	0	1	$P\{X = x_i\}$
0	0	0,5	0,5
1	0,25	0,25	0,5
$P\{Y = y_j\}$	0,25	0,75	1

18.395.

x_i	y_j		
	$y \leqslant 0$	$0 < y \leqslant 1$	$1 < y$
$x \leqslant 0$	0	0	0
$0 < x \leqslant 1$	0	0	0,5
$1 < x$	0	0,25	1

18.396. $K = \begin{pmatrix} -0{,}25 & -0{,}125 \\ -0{,}125 & -0{,}1875 \end{pmatrix}$.

18.397.

x_i	y_j		
	1	2	3
1	1/9	1/9	1/9
2	0	1/6	1/6
3	0	0	1/3

зависимы.

18.398. $\rho_{XY} = \sqrt{6/17} \approx 0{,}594$.

18.399.

x_i	1	2	3
$P\{X = x_i/Y = 3\}$	2/11	3/11	6/11

.

18.400. 2,5 %. У к а з а н и е. Ввести индикаторные случайные величины: X — индикатор брака вследствие дефекта α при испытании одной детали, Y — индикатор брака вследствие дефекта β при испытании одной детали — и описать закон распределения случайного вектора (X, Y).
18.401. 0,669. **18.402.** 0,0118. У к а з а н и е. См. указание к задаче 18.400.
18.403. 0,066. **18.404.** $c = 1$, $P\{X + Y < 1\} = 1/3$.

18.405. $f_X(x) = \begin{cases} 0, & \text{если } x \leqslant 0 \text{ или } x > 1, \\ x + \dfrac{1}{2}, & \text{если } 0 < x \leqslant 1; \end{cases}$ компоненты X и Y зависимы.

18.406. $F_X(x) = \begin{cases} 0, & \text{если } x \leqslant 0, \\ \dfrac{1}{2}x(x+1), & \text{если } 0 < x \leqslant 1, \\ 1, & \text{если } x > 1; \end{cases}$ $(m_X, m_Y) = (7/12;\ 7/12)$.

18.407. $c = 3/28$, $\mathrm{P}\{X+Y < 2\} = 3/14$. **18.408.** $(8/7;\ 10/7)$. **18.409.** $p_1 = \sqrt{3}/4$, $p_2 = (3\sqrt{3} - 4)/2$.

18.410. $f_{XY}(x, y) = \begin{cases} \dfrac{1}{2}\sin(x+y), & 0 \leqslant x \leqslant \pi/2,\ 0 \leqslant y \leqslant \pi/2, \\ 0 & \text{в остальных случаях;} \end{cases}$

$(m_X, m_Y) = (\pi/4,\ \pi/4)$.

18.411. $f_{XY}(x, y) = \begin{cases} \dfrac{1}{3}, & x \in (-1, 2),\ y \in (1, 2), \\ 0 & \text{в остальных случаях;} \end{cases}$

$$F_{XY}(x, y) = \begin{cases} 0, & x \leqslant -1 \text{ или } \quad y \leqslant 1, \\ \dfrac{1}{3}(x+1)(y-1), & -1 \leqslant x \leqslant 2, \quad 1 \leqslant y \leqslant 2, \\ \dfrac{1}{3}(x+1), & -1 \leqslant x \leqslant 2, \quad 2 < y, \\ y - 1, & 2 < x, \quad 1 \leqslant y \leqslant 2, \\ 1, & 2 < x, \quad 2 < y; \end{cases}$$

компоненты независимы. **18.412.** $(m_X, m_Y) = (0{,}5;\ 1{,}5)$, $D_X = 3/4$, $D_Y = 1/12$. **18.413.** $p_1 = 1/6$, $p_2 = \pi/12$.

18.414. $f_{XY}(x, y) = \begin{cases} 0, & (x, y) \notin \Pi, \\ 1/4, & (x, y) \in \Pi; \end{cases}$

$$\Pi = \{(x, y) \,\big|\, -1 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 2\},\ \mathrm{P}(A) = 1/4.$$

18.415. $(1/3; 2/3)$. **18.416.** Зависимы. **18.417.** $\begin{pmatrix} a^2/12 & 0 \\ 0 & a^2/12 \end{pmatrix}$; компоненты некоррелированы. **18.418.** $p_1 = 1/2$, $p_2 = \pi/4$. **18.419.** Не коррелированы, но зависимы. **18.420.** Каждая компонента распределена равномерно на отрезке $[0, 1]$, компоненты зависимы. **18.421.** $\mathrm{P}(A) = \dfrac{1}{4}(1 - e^{-4}) \approx 0{,}432$, $\mathrm{P}(B) = \dfrac{1}{8}(7 + e^{-8}) \approx 0{,}875$. **18.422.** $e^{-2} \approx 0{,}135$. **18.423.** $\approx 0{,}279$. **18.424.** $\mathrm{P}(A) = 0{,}5$, $\mathrm{P}(B) = 0{,}75$, $\mathrm{P}(C) = 0{,}5$. **18.425.** $\mathrm{P}(D) \approx 0{,}6826$, $\mathrm{P}(E) \approx 0{,}8221$. **18.426.** $\approx 3{,}4\sigma$. **18.427.** $\mathrm{P}(G_1) \approx 0{,}0328$, $\mathrm{P}(G_2) \approx 0{,}2818$. **18.428.** $\approx 0{,}5052$. **18.429.** $\mathrm{P}(G_1) \approx 0{,}0582$,

$\mathrm{P}(G_2) \approx 0{,}0887$. Указание. Воспользоваться симметрией кругового рассеивания. Например, $\mathrm{P}(G_1) = \mathrm{P}(G^{(1)}) = \dfrac{1}{2}\mathrm{P}(G)$, где $G^{(1)}$ — треугольник с вершинами $(0,0)$, $(1,1)$, $(0,1)$, а G — квадрат с вершинами $(0,0)$, $(1,0)$, $(1,1)$, $(0,1)$. Аналогично, $\mathrm{P}(G_2) = \mathrm{P}(G^{(2)})$, где $G^{(2)}$ — треугольник с вершинами $(0,0)$, $(\sqrt{2}, 0)$, $(\sqrt{2}, \sqrt{2})$, и далее как в предыдущей задаче. **18.430.** $\mathrm{P}(G_3) \approx 0{,}05566$, $\mathrm{P}(G_4) \approx 0{,}0252$.

$$\textbf{18.431}.\ f_{XY} = \frac{1}{32\pi}\exp\left\{-\frac{25}{32}\left(\frac{(x+2)^2}{16} - \frac{3(x+2)(y-3)}{50} + \frac{(y-3)^2}{25}\right)\right\},$$

$p \approx 0{,}865$. **18.432.** $y = 3 + \dfrac{3}{4}(x+2)$. **18.433.** $(x+1)^2 + \dfrac{(y-1)^2}{4} = 4{,}6052$. **18.434.** $f_X(x) = \sqrt{\dfrac{2}{5\pi}}\,e^{-2x^2/5}$, $m_X = m_Y = 0$, $D_X = 5/4$, $D_Y = 1/4$, $\rho_{XY} = -1/\sqrt{5}$. **18.435.** $n \geqslant 24$. **18.436.** $m_S = 7$, $D_S = 35/6$. **18.437.** $m_U = 18$, $m_V = 17$. **18.438.** $D_Z = 108$. **18.439.** $\mathrm{M}[X+Y] = 1{,}5$, $\mathrm{M}[X-Y] = 0{,}5$, $\mathrm{M}[X^2+Y^2] = 5/3$, $\mathrm{M}[XY] = 0{,}5$. **18.440.** $\mathrm{D}[X+Y] = \mathrm{D}[X-Y] = 5/12$, $\mathrm{D}[XY] = 7/36$. **18.441.** $\rho_{UV} = \dfrac{\alpha^2 - \beta^2}{\alpha^2 + \beta^2}$. **18.443.** $1/\sqrt{10} \approx 0{,}3162$. **18.444.** $\mathrm{D}[X-Y] = 13/3$, $\mathrm{M}[XY^2 + X^2Y] = 31/3$. **18.445.** $m_Z = 1$, $D_Z = 32/3$.

18.446. $\begin{pmatrix} 1 & 2 & 6 \\ 2 & 6 & 18 \\ 6 & 18 & 60 \end{pmatrix}$. Указание. При вычислении моментов использовать формулу (12) § 2 и результат задачи 18.274. **18.447.** $\mathrm{M}[Y] = 0$, $\mathrm{D}[Y] = \dfrac{n}{2(n+1)}$. Указание. Воспользоваться формулами

$$\sum_{k=0}^{n}\sin\left(\frac{2\pi k}{n}\right) = \mathrm{Im}\left(\sum_{k=0}^{n}\exp\left\{i\frac{2\pi k}{n}\right\}\right),$$
$$\sum_{k=0}^{n}\sin^2\left(\frac{2\pi k}{n}\right) = \frac{1}{2}\left(n + 1 - \mathrm{Re}\left(\sum_{k=0}^{n}\exp\left\{i\frac{4\pi k}{n}\right\}\right)\right).$$

18.448. $m_Y = \dfrac{pq}{1+q^2}$, $D_Y = \dfrac{pq}{1-q^2} - \dfrac{p^2q^2}{(1+q^2)^2}$. **18.449.** $m_Z = \dfrac{p}{1+q^2}$, $D_Z = \dfrac{p}{1-q^2} - \dfrac{p^2}{(1+q^2)^2}$. **18.450.** $\mathrm{M}[L] = 4r/\pi$, $\mathrm{D}[L] = 2r^2(1 - 8/\pi^2)$. **18.451.** $\mathrm{M}[L] = \pi l$, $\mathrm{D}[L] = \pi^2 l^2/3$. **18.452.** $\mathrm{M}[S] = \pi l^2/3$, $\mathrm{D}[S] = 4\pi^2 l^4/45$. **18.453.** $l^2/6$. **18.454.** $\mathrm{M}[R] = l/3$, $\mathrm{D}[R] = l^2/18$. **18.455.** $\mathrm{M}[R] = 2a/3$, $\mathrm{D}[R] = a^2/18$. **18.456.** $\mathrm{M}[R^2] = \dfrac{1}{3}(a^2 + b^2)$,

$D[R^2] = \dfrac{4}{45}(a^4 + b^4)$. **18.457.** 4,084. **18.458.** 4,084. **18.459.** $M[Z] = p$, $D[Z] = pq/n$. **18.460.** $m_z = k/p$, $D_z = kq/p^2$. Указание. Для вычисления m_z и D_z положить $Z = \sum\limits_{m=1}^{k} X_m$, где X_m — число деталей, сошедших с автоматической линии от момента получения $(m-1)$-й по счету нестандартной детали до момента получения m-й нестандартной (последняя включительно). Далее воспользоваться тем, что X_m подчиняется геометрическому распределению с параметром p, и применить свойства математического ожидания и дисперсии. **18.461.** $M[Z_n] = npq + p^2$. Указание. Ввести индикаторы: I_k — индикатор успеха в k-м испытании, J_k — индикатор начала серии для k-го испытания, т.е. $J_k = 1$, если k-е испытание является началом очередной серии успехов, и $J_k = 0$, если не является. Показать, что $Z_n = I_1 + \sum\limits_{k=2}^{n} J_k$. **18.462.** Указание. Вначале вычислить условные математические ожидания $M[Z/Y = n]$ и $M[Z^2/Y = n]$, а затем воспользоваться формулой полного математического ожидания (8) § 3. **18.463.** $M[Z] = 49/4$, $D[Z] = 735/16$. Указание. Представить Z в виде $Z = \sum\limits_{n=1}^{Y} X_n$, где Y — число очков при одном подбрасывании игральной кости, X_n — число очков при n-м подбрасывании, и воспользоваться результатом задачи 18.462. **18.464.** $m_z = m/p$, $D_z = \dfrac{\sigma^2}{p} + \dfrac{m^2 q}{p^2}$. Указание. Использовать результат задачи 18.462. **18.467.** Воспользоваться свойством дисперсии суммы случайных величин. **18.468.** Указание. Пусть $Z = \max(X^2, Y^2)$. Убедиться, что Z может быть представлена в виде $Z = \dfrac{X^2 + Y^2}{2} + \left|\dfrac{Y^2 - X^2}{2}\right|$. Левая часть доказываемого неравенства вытекает из свойств 2) и 4) математического ожидания. Для доказательства правой части неравенства применить свойства 4) и 6) математического ожидания. **18.469.** $1 - \dfrac{1}{2^n}$. Указание. Найти сначала условные математические ожидания $M[X_k/X_{k-1}]$, $k = 2, 3, \ldots, n$, а затем воспользоваться формулой полного математического ожидания (6) § 3. **18.471.** а) $M[X_n] = n\sum\limits_{k=1}^{n}\dfrac{1}{k}$; б) $M[X_2] = 3$, $M[X_5] = 11{,}4$, $M[X_{10}] = 27{,}86$, $M[X_{100}] \approx 518{,}2$. Указание. Доказать формулу $X_n = 1 + \sum\limits_{k=1}^{n-1} Y_k$, где Y_k — число писем, опущенных вплоть до момента, пока одно из писем попадет в один из k ящиков, до тех пор оставшихся пустыми. Далее воспользоваться свойством математического ожидания и учесть, что случайные величины Y_k соответствуют опытам до первого успеха. При больших значениях n использовать формулу $\sum\limits_{k=1}^{n} \frac{1}{k} = \ln n + c$, где $c \approx 0{,}577$ — постоянная Эйлера. **18.472.** $E_X(t) = \cos t$. **18.473.** $E_X(t) = \cos^2 t$, $\sigma_X^2 = 2$. **18.474.** $E_{I_k}(t) = pe^{it} + q$, $E_X(t) = (pe^{it} + q)^n$. **18.475.** $E_{I_k}(t) = p_k e^{it} + q_k$, $E_X(t) = \prod\limits_{k=1}^{n} (p_k e^{it} + q_k)$. **18.476.** $E_X(t) = (pe^{it}+q)^n$. **18.477.** $m_X = np$, $D_X = npq$. **18.478.** $E_X(t) = pe^{it}/(1-qe^{it})$, $m_X = p$. **18.479.** $E_Z(t) = [E(t)]^2$. **18.483.** Указание. Использовать неравенство $\cos x \geqslant 1 - x^2/2$. **18.484.** $E_X(t) = \dfrac{\lambda}{\lambda - it}$. **18.485.** $m_X = 1/\lambda$, $D_X = 1/\lambda^2$, $a_X = 2$. **18.486.** $E_X(t) = \dfrac{(e^{itb} - e^{ita})}{it(b-a)}$. **18.487.** $E_X(t) = e^{itc - a|t|}$. Указание. Получить для характеристической функции выражение $E_X(t) = \dfrac{a}{\pi} e^{itc} \displaystyle\int\limits_{-\infty}^{+\infty} \frac{e^{itx}}{x^2 + a^2}\, dx$. Интеграл, стоящий в правой части, может быть вычислен с помощью теории вычетов, если контур интегрирования замкнуть полуокружностью радиуса $r \to \infty$ с центром в начале координат, лежащей в верхней (при $t > 0$) или нижней (при $t < 0$) полуплоскости. **18.488.** $E_Y(t) = (e^{it} - 1)/it$, что соответствует равномерному распределению $R(0,\,1)$. **18.489.** $E_Y(t) = 1/(1-it)$. Характеристическая функция соответствует показательному распределению с параметром $\lambda = 1$. **18.490.** $E_Y(t) = e^{itm}\, e^{-\sigma^2 t^2/2}$, что соответствует нормальному закону распределения $N(m,\,\sigma)$. **18.491.** $E_Z(t) = e^{-|t|}$, Z распределена по закону Коши с параметрами $c = 0$ и $a = 1$. **18.492.** $f_X(x) = \dfrac{1}{2} e^{-|x|}$, что соответствует закону распределения Лапласа с параметрами $m = 0$ и $\sigma = \sqrt{2}$. **18.493.** $E_Y(t) = (1 - 2it)^{-1/2}$. **18.494.** Указание. Используя результат 18.493, показать, что $E_{Z_n}(t) = (1 - 2it)^{-n/2}$. С другой стороны, установить, что

$$E'_{Z_n}(t) = \frac{\left(\frac{1}{2}\right)^n}{\Gamma\left(\frac{n}{2}\right)} \int\limits_0^{+\infty} x^{n/2-1} e^{-x/2+itx}\, dx,$$

где Z'_n подчиняется закону распределения $\chi^2(n)$ и, вычислив этот интеграл, получить тот же результат. **18.495.** $\mathrm{M}[Z_n] = n$, $\mathrm{D}[Z_n] = 2n$.

18.496. $E_Y(t) = e^{itm}\Big/\left(1+\dfrac{\sigma^2t^2}{2}\right)$, $m_Y = m$, $\sigma_Y^2 = \sigma^2$. **18.497.** $f_X(x) = \dfrac{1-\cos x}{\pi x^2}$.

18.498.

y_k	-1	$-\sqrt{2}/2$	0	$\sqrt{2}/2$	1
$\mathrm{P}\{Y=y_k\}$	$1/9$	$2/9$	$2/9$	$2/9$	$2/9$

18.499.

x_k	1	2	4	6
p_k	5/16	15/32	3/16	1/32

18.500.

x_k	2	3	4	5	6	7	8	9	10	11	12
p_k	1/36	1/18	1/12	1/9	5/36	1/6	5/36	1/9	1/12	1/18	1/36

18.501. Случайная величина Y подчиняется закону $N(0,\,1)$. Указание. Воспользоваться монотонностью линейной функции.

18.502. $$F_Y(y) = \begin{cases} 0, & -y \leqslant -4, \\ F_X\left(\dfrac{\sqrt{4+y}}{3}\right) - F_X\left(-\dfrac{\sqrt{4+y}}{3}\right), & -4 < y; \end{cases}$$

$$F_Z(z) = \begin{cases} 0, & z \leqslant 0, \\ F_X(1+z) - F_X(1-z), & 0 < z; \end{cases}$$

$$F_V(v) = \begin{cases} 0, & v \leqslant 0, \\ 1 - F_X\left(\dfrac{1}{2}\ln\dfrac{1}{v}\right), & 0 < v \leqslant 1. \\ 1, & 1 < v \end{cases}$$

18.503. $$f_Z(z) = \begin{cases} 0, & z \leqslant 0, \\ \dfrac{2}{\sigma\sqrt{2\pi}}\exp\{-z^2/2\sigma^2\}, & z > 0. \end{cases}$$

18.504. $f_W(w) = \begin{cases} 0, & w \notin [0,\,1], \\ 1, & w \in [0,\,1], \end{cases}$ что соответствует закону $R(0,\,1)$. Указание. Вычислить сначала функцию распределения $F_W(w)$. **18.505.** $f_Y(y) = \dfrac{a}{\pi\,(y^2+a^2)}$ (закон Коши с параметрами $c = 0$ и a).

18.506. $$f_X(x) = \begin{cases} 0, & x \leqslant 2, \\ \dfrac{4}{\pi x\sqrt{x^2-4}}, & x > 2. \end{cases}$$

18.507. $f_{Y_1}(y)=\begin{cases}\dfrac{1}{3l\sqrt[3]{3y^2}}, & y\in\left(-\dfrac{3l^3}{8},\dfrac{3l^3}{8}\right),\\ 0, & y\notin\left(-\dfrac{3l^3}{8},\dfrac{3l^3}{8}\right);\end{cases}$

$$f_{Y_2}(y)=\begin{cases}\dfrac{1}{\pi}\cdot\dfrac{1}{\sqrt{a^2-y^2}}, & |y|<a,\\ 0, & |y|\geqslant a.\end{cases}$$

18.508. $f_Y(y)=\begin{cases}0, & y\leqslant 0,\\ 2\lambda ye^{-\lambda y^2}, & y>0;\end{cases}$

$$f_Z(z)=\begin{cases}0, & z\leqslant 0,\\ \dfrac{\lambda}{2\sqrt{z}}e^{-\lambda\sqrt{z}} & z>0;\end{cases}$$

$$f_U(u)=\begin{cases}0, & u\leqslant 0 \quad \text{или} \quad u>1,\\ 1, & 0<u\leqslant 1.\end{cases}$$

18.509. $f_{Y_1}(y)=\dfrac{3}{\pi}\cdot\dfrac{1}{y^2+4y+13};$

$$f_{Y_2}(y)=\frac{3}{\pi}\cdot\frac{1}{13y^2+4y+1}.$$

18.510. $f_{Y_3}(y)=\begin{cases}1, & y\in\left(-\dfrac{1}{2},\dfrac{1}{2}\right),\\ 0, & y\notin\left(-\dfrac{1}{2},\dfrac{1}{2}\right).\end{cases}$

$$f_{Y_4}(y)=\begin{cases}0, & y\leqslant 0,\\ \dfrac{2}{\pi}\dfrac{1}{\sqrt{y}\,(4+y)}, & y>0.\end{cases}$$

18.511. Показательный с параметром a. **18.512.** Указание. Записать выражение для $F_T(t)$ в виде двойного интеграла по соответствующей области. **18.513.** $m_{\max}=n-1$.

18.514.

u_k	0	1	2	3
p_k	0,3	0,4	0,25	0,05

v_k	−1	0	3
p_k	0,35	0,55	0,1

.

18.515.

z_k	0	1	2
p_k	q^2	$2pq$	p^2

v_k	0	1
p_k	q^2+2pq	p^2

.

18.516. $F_Z(z) = \begin{cases} 0, & -z \leqslant -1, \\ 0{,}2, & -1 < z \leqslant 0, \\ 0{,}9, & 0 < z \leqslant 1, \\ 1, & -1 < z. \end{cases}$

18.517. $f_Z(z) = \begin{cases} 0, & z \leqslant 0, \\ 1/3, & 0 < z \leqslant 1, \\ \dfrac{1}{3z^{3/2}}, & 1 < z. \end{cases}$

18.518. $f_Z(z) = \begin{cases} -0, & z \leqslant 0 \text{ или } z > 2, \\ -z^2, & 0 < z \leqslant 1; \\ -z^2+2z, & 1 < z \leqslant 2; \end{cases}$

$$f_U(u) = \begin{cases} 0, & u \leqslant 0 \text{ или } u \geqslant 1, \\ 2(1-u), & 0 < u < 1. \end{cases}$$

18.519. $g_{U,V}(u, v) = \begin{cases} \dfrac{1}{4}\left(\dfrac{1}{\sqrt{v}} + \dfrac{1}{\sqrt{u}}\right), & 0 < u < v < 1, \\ 0 & \text{в остальных случаях.} \end{cases}$

18.520. $f_Z(z) = \begin{cases} 0, & z \leqslant 0 \text{ или } z \geqslant r, \\ \dfrac{2z}{r^2}, & 0 < z < r. \end{cases}$

18.521. $f_Z(r) = \begin{cases} 0, & r \leqslant 0, \\ \dfrac{r}{\sigma^2}\exp\{-r^2/2\sigma^2\}, & 0 < r, \end{cases}$ что соответствует закону Рэлея. **18.522.** По закону Коши с параметрами $c = 0$, $a = 1$.

18.523. $f_{R,\Phi}(r, \varphi) =$

$$= \begin{cases} 0, & r \leqslant 0, \\ \dfrac{1}{2\pi\sigma_X\sigma_Y} r \exp\left(-\dfrac{r^2}{2}\left(\dfrac{\cos^2\varphi}{\sigma_X^2} + \dfrac{\sin^2\varphi}{\sigma_Y^2}\right)\right), & r > 0. \end{cases}$$

18.524. Указание. Рассмотреть функцию распределения $F_{X^nY^n}(u, v) = \mathrm{P}\{X^n < u, Y^n < v\}$ отдельно для нечетных и четных n и показать, что как в том, так и в другом случае справедливо равенство

$F_{X^n, Y^n}(u, v) = F_{X^n}(u)\, F_{Y^n}(v)$. **18.525.** $F_U(u) = 1 - (1 - F(u))^2$, $F_V(v) = F^2(v)$. Указание. Использовать эквивалентность событий $\{U \geqslant u\}$ и $\{X \geqslant u\}\{Y \geqslant u\}$, $\{V < v\}$ и $\{X < v\}\{Y < v\}$.

18.526. $F_{U,V}(u, v) = \begin{cases} 0, & v \leqslant u, \\ F^2(v) - (F(v) - F(u))^2, & u < v. \end{cases}$

18.527. $f_{U,V}(u, v) = \begin{cases} \dfrac{2}{(b-a)^2}, & a < u < v \leqslant b, \\ 0 & \text{в остальных случаях.} \end{cases}$

Указание. Использовать результат задачи 18.526.

18.528.

z_k	0	1	2	3	4
p_k	1/4	3/8	17/64	3/32	1/64

18.529. $\mathrm{P}\{Z = k\} = (k-1)p^2q^{k-2}$, $k = 2, 3, \ldots$ **18.530.** $m_z = D_z = \lambda_1 + \lambda_2$.

18.533. $f_Z(z) = \begin{cases} 0, & z \leqslant -2 \\ \dfrac{z}{4} + \dfrac{1}{2}, & -2 < z \leqslant 0, \\ -\dfrac{z}{4} + \dfrac{1}{2}, & 0 < z \leqslant 2, \\ 0, & z > 2 \end{cases}$ (закон Сипсона с параметром $a = 2$).

Указание. Сначала найти функцию распределения суммы.

18.534. $f_Z(z) = \dfrac{1}{b-a}(F_X(z-a) - F_X(z-b))$.

18.535. $f_Z(z) = \dfrac{1}{6}\left(\Phi\left(\dfrac{z+3}{2}\right) - \Phi\left(\dfrac{z-3}{2}\right)\right)$.

Указание. Воспользоваться результатом задачи 18.534.

18.536. $f_Z(z) = \begin{cases} 0, & z \leqslant 0 \\ \dfrac{\lambda_1\lambda_2}{\lambda_2 - \lambda_1}\left(e^{-\lambda_1 z} - e^{-\lambda_2 z}\right), & z > 0. \end{cases}$

18.538. $f_Z(z) = \begin{cases} 0, & z \leqslant 0 \\ \lambda^2 z e^{-\lambda z}, & z > 0 \end{cases}$ (закон Эрланга 1-го порядка).

Указание. Найти плотность с помощью обратного преобразования Фурье от характеристической функции (свойство 6). **18.539.** Указание. Использовать результат задачи 18.538 и метод математической индукции. **18.540.** Указание. Использовать свойства 4 и 6 характеристической функции и результат задачи 18.494. **18.541.** Показательное распределение с параметром $p\lambda$. Указание. Найти условную характеристическую функцию $\mathrm{M}\left[e^{itZ}/Y = n\right]$ и использовать формулу полного математического ожидания. **18.542.** $\mathrm{P}(A) \geqslant 0{,}84$, $\mathrm{P}(B) \geqslant 0{,}36$, $\mathrm{P}(C) \geqslant 0{,}96$. **18.543.** $\mathrm{P}(A) \leqslant 0{,}2$. **18.544.** $\mathrm{P}(A) \leqslant 0{,}004$. **18.545.** $\mathrm{P}(A) \leqslant 0{,}002$. **18.546.** $\mathrm{P}(A) \leqslant 0{,}16$, $\mathrm{P}(B) \leqslant 0{,}04$. **18.547.** $p_1 \geqslant 0$, $p_2 \geqslant 0{,}75$, $p_3 \geqslant 8/9$. **18.549.** Указание. Воспользоваться результатом задачи 18.548. **18.550.** Применим, так как $\mathrm{D}[X_n] \leqslant 2$, $n \in \mathbb{N}$. **18.551.** Применим. **18.552.** Применим. **18.553.** Применим, так как выполняется условие (1). **18.553(1).** Применим. **18.554.** 0,666. **18.555.** $\mathrm{P}\{X < 15\} \geqslant 0{,}96$. **18.556.** $\mathrm{P}(B) \approx 0{,}9066$, $\mathrm{P}(C) \approx 0{,}8858$. **18.557.** $\mathrm{P}(D) \approx 0{,}1051$. **18.558.** $\mathrm{P}(A) \approx 0{,}9964$, $\mathrm{P}(B) \approx 0{,}0176$. **18.559.** $\mathrm{P}(A) \approx 0{,}0797$, $\mathrm{P}(B) \approx 0{,}5160$. **18.560.** $\mathrm{P}(C) \approx 0{,}6689$. **18.561.** $(-0{,}0283;\ 0{,}0283)$. **18.562.** а) $\approx 0{,}8944$; б) 0,4. Указание. Относительной точностью измерения называется величина $\left|\dfrac{X - m_X}{m_X}\right|$. **18.563.** $\approx 0{,}0392$. **18.564.** $(792, 828)$. **18.565.** $\mathrm{P}(A) \approx 0$ с точностью до пяти знаков. **18.566.** $\mathrm{P}(B_{40000}) \approx 0{,}995$, $\mathrm{P}(B_{60000}) = 0{,}5$, $\mathrm{P}(B_{80000}) \approx 0{,}005$. **18.567.** $\approx 0{,}9974$. **18.568.** $\approx 0{,}3888$. **18.569.** Не менее 1000 раз. **18.570.** Не менее 127 раз. **18.570(1).** 29. **18.571.** 753. Указание. Применить интегральную теорему Муавра–Лапласа. **18.573.** 0,9998. **18.574.** $\mathrm{P}(A) = 0{,}5$, $\mathrm{P}(B) = 0{,}9742$. **18.575.** 0,0008. **18.576.** 0,84. **18.578.** $\approx 0{,}0062$. **18.579.** $\lhd$ Если X распределена по закону $\mathrm{St}(n)$, то

$$\ln f_X(x) = \ln\Gamma\left(\frac{n+1}{2}\right) - \ln\Gamma\left(\frac{n}{2}\right) - \frac{1}{2}\ln(n\pi) - \frac{n+1}{2}\ln\left(1 + \frac{x^2}{n}\right).$$

Воспользуемся формулой Стирлинга для гамма-функции [1])

$$\ln\Gamma(u) = \left(u - \frac{1}{2}\right)\ln u - u + \frac{1}{2}\ln(2\pi) + O\left(\frac{1}{u}\right).$$

Отсюда следует, что

$$\ln f_X(x) = -\frac{1}{2}\ln(\pi n) - \frac{1}{2} + \left(\frac{n-1}{2}\right)\ln\frac{n-1}{n+2} + \\ + \frac{1}{2}\ln\left(\frac{n-1}{2}\right) - \frac{n+1}{2}\ln\left(1 + \frac{x^2}{n}\right) \underset{n\to+\infty}{\longrightarrow} -\frac{1}{2}\ln(2\pi) - \frac{x^2}{2},$$

и, следовательно, $\lim\limits_{n\to+\infty} f_X(x) = \dfrac{1}{\sqrt{2\pi}}\, e^{-x^2/2}$ для всех $x \in \mathbb{R}$. $\rhd$

[1]) Из этой формулы, в частности, следует известная формула Стирлинга для факториалов $n! = \Gamma(n+1) \approx (n/e)^n\sqrt{2\pi n}$ при больших n.

18.580. $\approx$ 1,28. Точное табличное значение 1,289. **18.581.** 5,3%. **18.582.** Не менее 33 испытаний. Указание. Представить $\Phi(1)$ в виде $\Phi(1) = 0{,}5 + I$, где

$$I = \mathrm{M}[Y] = \frac{1}{\sqrt{2\pi}} \int\limits_0^1 e^{-x^2/2}\, dx, \quad Y = \varphi(X) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

X подчиняется закону распределения $R(0,\,1)$. Кроме того, учесть, что $\Phi(1) = 0{,}8413$, $I = 0{,}3413$. **18.583.** 0,7372. **18.584.** $n_{min} = 27427$. Указание. Вероятность пересечения иглой любой из параллельных прямых равна $p = 1/\pi$. Так как в данном эксперименте измеряется число пересечений (т.е. случайная величина X_n), то $\dfrac{1}{n} X_n = \dfrac{1}{\pi^*}$, где π^* — приближенное значение числа π в методе статистических испытаний. По центральной предельной теореме случайная величина

$$T = \frac{\dfrac{1}{n} X_n - \mathrm{M}\left[\dfrac{1}{n} X_n\right]}{\sqrt{\mathrm{D}\left[\dfrac{1}{n} X_n\right]}} = \frac{\sqrt{n}\left(\dfrac{\pi}{\pi^*} - 1\right)}{\sqrt{\pi - 1}}$$

асимптотически (при больших n) распределена по закону $N(0,\,1)$. Так как события

$$\left\{\frac{|\pi - \pi^*|}{\pi} < \delta_{\min} = 0{,}03\right\} \text{ и } \left\{|T| < \frac{\delta_{\min}}{1 - \delta_{\min}} \cdot \frac{\sqrt{n}}{\sqrt{\pi - 1}}\right\}$$

равносильны, то

$$\mathrm{P}\left\{\frac{|\pi - \pi^*|}{\pi} < 0{,}03\right\} = 2\Phi\left(\frac{0{,}03}{0{,}97} \cdot \frac{\sqrt{n}}{\sqrt{\pi - 1}}\right) - 1 = 0{,}9996.$$

Решая полученное уравнение, находим наименьшее значение n.

18.585. $m_X(t) = vt$, $D_X(t) = 1 + t^2$,

$$K_X(t_1,\, t_2) = \begin{cases} 0, & t_1 \neq t_2, \\ 1 + t^2, & t_1 = t_2 = t. \end{cases}$$

18.586. $m_X(t) = 1/2$, $\quad D_X(t) = 1/4$,

$$F_2(x,\, y/t_1,\, t_2) = \begin{cases} 0, & x \leqslant 0 \quad \text{или} \quad y \leqslant 0, \\ F_X(x) = 1 - e^{-2x}, & 0 < x < y, \\ F_Y(y) = 1 - e^{-2y}, & 0 < y < x. \end{cases}$$

18.587. $F_1(x/t) = \begin{cases} 0, & x \leqslant 0, \\ \dfrac{1}{3}\dfrac{x}{t^2}, & 0 < x \leqslant 3t^2, \\ 1, & x > 3t^2; \end{cases}$

$$f_1(x/t) = \begin{cases} \dfrac{1}{3t^2}, & x \in (0,\, 3t^2), \\ 0, & x \notin (0,\, 3t^2). \end{cases}$$

18.588. $f_1(x/t) = \dfrac{1}{\sigma\,|t|\,\sqrt{2\pi}} \exp\left\{-\dfrac{(x - tm - b)^2}{2t\sigma^2}\right\}$, $m_X(t) = tm + b$, $\sigma_X(t) = \sigma\,|t|$, $K_X(t_1,\, t_2) = t_1 t_2 \sigma^2$. **18.589.** $m_X(t) = m\,(1+t)$, $D_X(t) = \sigma^2(1+t^2)$, $K_X(t_1,\, t_2) = \sigma^2(1 + t_1 t_2)$.

18.590. $f_1(x/t) = \dfrac{1}{\sigma\sqrt{2\pi(1+t^2)}} \exp\{-(x - m\,(1+t))^2/2\sigma^2(1+t^2)\}$.

18.591. $f_1(x/t) = \displaystyle\int_{-\infty}^{+\infty} f_V(v)\, f_U(x - vt)\, dv$, $t > 0$. **18.591(1).** $m_X(t) = 0$, $D_X(t) = \sqrt{2}\, e^{4t^2}$, $K_X(t_1,\, t_2) = \sqrt{2}\, e^{(t_1+t_2)^2}$. **18.592.** $K_Z(t_1,\, t_2) = K_X(t_1,t_2)K_Y(t_1,t_2) + K_X(t_1,t_2)m_Y(t_1)m_Y(t_2) + K_Y(t_1,t_2)m_X(t_1)m_X(t_2)$.

18.595. $K_Y(t_1,\, t_2) = \dfrac{e^{-(t_1^2+t_2^2)}}{1+(t_1-t_2)^2}$, $D_Y(t) = e^{2t^2}$. **18.596.** $m_Z(t) = 4 + t + t^2$, $D_Z(t) = 9 + 4t^2$. **18.597.** $\rho_Z(t_1,\, t_2) = \exp\{-2\,(t_2 - t_1)\}\dfrac{9 + 4t_1t_2 + 6(t_1+t_2)}{|3+2t_1|\cdot|3+2t_2|}$.

18.598. $\rho_X(t_1,\, t_2) = \rho_Y(t_1,\, t_2) = \cos(t_2 - t_1)$, $\rho_{XY}(t_1,\, t_2) = \sin(t_2 - t_1)$.

18.599. $m_X(t) = \displaystyle\int_{-\infty}^{+\infty} \varphi(t, y)\, f_Y(y)\, dy$,

$$K_X(t_1,\, t_2) = \int_{-\infty}^{+\infty} \varphi(t_1,\, y)\, \varphi(t_2,\, y)\, f_Y(y)\, dy - m_X(t_1)\, m_X(t_2),$$

$$D_X(t) = \int_{-\infty}^{+\infty} \varphi^2(t,\, y)\, f_Y(y)\, dy - m_X^2(t).$$

18.600. $m_X(t) = 0$; $D_X(t) = 1$;

$$K_X(\tau) = \begin{cases} 1, & \text{если} \quad t \in (t_k,\, t_{k+1}), \quad t+\tau \in (t_k,\, t_{k+1}), \quad k = 1, 2, \ldots, \\ 0, & \text{если} \quad t \in (t_k,\, t_{k+1}), \quad t+\tau \notin (t_k,\, t_{k+1}), \quad k = 1, 2, \ldots \end{cases}$$

18.601. $m_X(t) = m_A F_T(t)$, $K_X(t,\, t+\tau) = m_A^2 F_T(t)(1-F_T(t+\tau)) + \sigma_Z^2 F_T(t)$, где $F_T(t) = \int\limits_{-\infty}^{t} f_T(s)\,ds$. **18.602.** $f_1(x/t) = \frac{1}{\sqrt{2\pi}}\sigma \exp\left(-\frac{x^2}{2\sigma^2}\right)$, где $x(t) = a\cos\omega t + b\sin\omega t$ — реализация случайного процесса;

$$f_2(x,\, y/t,\, t+\tau) = \frac{1}{2\pi\sigma^2\,|\sin\omega t|}\exp\left(-\frac{x^2 - 2xy\cos\omega\tau + y^2}{2\sigma^2\sin^2\omega\tau}\right),$$

где $x = x(t) = a_1\cos\omega t + b_1\sin\omega t$, $y = x(t+\tau) = a_1\cos\omega(t+\tau) + b_1\sin\omega(t+\tau)$ — реализация процесса в сдвинутые на τ моменты времени.

18.603. $f_1(z/t) =$

$$= \frac{1}{\sqrt{2\pi\,(4a^2 + 9b^2 + 8ab)}}\exp\left\{-\frac{(z - at - b\,(1+t^2))^2}{2\,(4a^2 + 9b^2 + 8ab)}\right\}.$$

18.604.
$$\rho_Z(t_1,\, t_2) = \frac{1}{4a^2 + 9b^2 + 8ab}\left(\frac{4a^2}{1 + 2\,(t_1 - t_2)^2} + {}\right.$$
$$\left. + 9\,b^2\exp\left\{-2\,(t_2 - t_1)^2\right\} + 8ab\cos\omega\,(t_2 - t_1)\right).$$

18.605. $\mathrm{P}\left\{X(t_2) > \beta\right\} = 1 - \Phi\left(\frac{\beta - \alpha\rho(\tau)}{\sigma\sqrt{1-\rho^2(\tau)}}\right)$. Указание. Использовать выражение для условной плотности двумерного нормального распределения, полученное в примере 8 §3 п. 2. **18.607.** $1{,}15\cdot 10^{-3}$. **18.608.** $\approx 0{,}725$. **18.609.** Не менее 128. **18.610.** $m_X(t) = D_X(t) = \lambda t$. **18.611.** $\mathrm{P}\left\{X(t_2) =\right.$

$$\left. = m/X(t_1) = n\right\} = \begin{cases} 0, & m < n, \\ \dfrac{(\lambda\,(t_2 - t_1))^{m-n}}{(m-n)!}\exp\left\{-\lambda\,(t_2 - t_1)\right\}, & n \leqslant m. \end{cases}$$

18.612. $K_X(t_1,\, t_2) = \lambda\min(t_1,\, t_2)$. Указание. Учесть, что

$$\mathrm{M}\left[X(t_1)\,X(t_2)\right] = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\mathrm{P}\left\{X(t_1) = n\right\}\mathrm{P}\left\{X(t_2) = m\,|\,X(t_1) = n\right\},$$

и использовать результат предыдущей задачи.

18.613. $f_1(x/t) = \begin{cases} 0, & x \leqslant 0, \\ \lambda e^{-x\lambda}, & x > 0. \end{cases}$ **18.614.** $\mathrm{P}\left\{X(t) \geqslant 200\right\} \approx 0{,}9$, $m_X = 2\cdot 10^3$, $D_X = 4\cdot 10^6$. Указание. Использовать результат предыдущей задачи.

18.615. $f_1^{(k)}(x/t) = \begin{cases} 0, & x \leqslant 0, \\ \dfrac{(\lambda x)^{k-1}\lambda e^{-\lambda x}}{(k-1)!}, & x > 0 \end{cases}$ (закон Эрланга k-го порядка).

18.616. $X(t) = -\psi_1(t) + \psi_2(t) + V_1\varphi_1(t) + V_2\varphi_2(t)$, $K_X(t_1, t_2) = 3\varphi_1(t_1)\,\varphi_1(t_2) + \varphi_2(t_1)\,\varphi_2(t_2)$, где $\varphi(t) = A\psi(t)$, $V = A\mathring{U}$, $A = \dfrac{\sqrt{2}}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. Указание. Воспользоваться методом ортогонального преобразования, рассмотренным в примере 3. **18.617.** $X(t) = t + \cos t + 2\sin t + V_1\varphi_1(t) + V_2\varphi_2(t)$, $K_X(t_1, t_2) = \varphi_1(t_1)\,\varphi_1(t_2) + \dfrac{13}{8}\varphi_2(t_1)\,\varphi_2(t_2)$, где $\varphi(t) = A\psi(t)$, $V = (A^{\mathsf{T}})^{-1}\mathring{U}$, $A = \dfrac{1}{5}\begin{pmatrix} 25 & 18 \\ 0 & 24 \end{pmatrix}$, $(A^{\mathsf{T}})^{-1} = \dfrac{1}{120}\begin{pmatrix} 24 & 0 \\ -18 & 25 \end{pmatrix}$. Указание. Применить метод Лагранжа, рассмотренный в примере 4. **18.618.** $X(t) = m_X(t) + V_1\varphi_1(t) + V_2\varphi_2(t) + V_3\varphi_3(t)$; $K_X(t_1, t_2) = 8\varphi_2(t_1)\varphi_2(t_2) + 4\varphi_3(t_1)\varphi_3(t_2)$, где $\varphi(t) = A\psi(t)$, $V = A\mathring{U}$,

$$A = \begin{pmatrix} \dfrac{1}{\sqrt{3}} & -\dfrac{1}{\sqrt{3}} & -\dfrac{1}{\sqrt{3}} \\ \dfrac{1}{\sqrt{2}} & -0 & -\dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{6}} & -\dfrac{2}{\sqrt{6}} & -\dfrac{1}{\sqrt{6}} \end{pmatrix}.$$

Указание. Использовать метод ортогонального преобразования, рассмотренный в примере 3. **18.619.** $m_Z(t) = 1$, $D_Z(t) = 1 + \cos^2 t$, $K_Z(t_1, t_2) = \sin t_1 \sin t_2 + 2\cos t_1 \cos t_2$. **18.620.** $m_Y(t) = t^2/2$, $D_Y(t) = (1 - \cos t)^2 + 2(1 - \cos t)$, $K_Y(t_1, t_2) = \sin t_1 \sin t_2 + 2\cos t_1 \cos t_2 - 2(\cos t_1 + \cos t_2) + 2$. **18.621.** $m_Z(t) = 0$, $K_Z(t_1, t_2) = 3 + 4t_1t_2$, где $Z(t) = \dfrac{dX(t)}{dt}$. **18.622.** $m_Y(t) = t$, $D_Y(t) = \dfrac{3}{4}t^4 + \dfrac{1}{9}t^6$. **18.623.** Непрерывен, но не дифференцируем. Указание. Показать, что необходимое условие дифференцируемости не выполняется. **18.624.** Дифференцируем. $\mathrm{D}\left[\dfrac{dX(t)}{dt}\right] = \sigma^2\alpha^2$. **18.625.** $m_Y(T) = 2\cos\omega t - 6\omega\sin\omega t$, $K_Y(t_1, t_2) = 9(\cos\omega t_1 - 3\omega\sin\omega t_1)(\cos\omega t_2 - 3\omega\sin\omega t_2)$, $D_Y(t) = 9(\cos\omega t - 2\omega\sin\omega t)^2$. Указание. Использовать общие формулы преобразования (13). **18.626.** $\mathrm{D}\left[\dfrac{dX(t)}{dt}\right] = (2\alpha + \beta^2)\,D_X$. **18.627.** $m_Y(t) = 7t^2 + 2t$, $D_Y(t) = 4t^2$. **18.628.** $m_Y(t) = \dfrac{3}{2}t - \dfrac{1}{4\omega}\sin 2\omega t$, $D_Y(t) = \dfrac{4}{\omega^2}\sin^2\omega t$. **18.629.** $D_Y(t) = \sigma^2 t\left(t + \dfrac{1}{\lambda}e^{-\lambda t}\right)$. **18.631.** $K_{XY}(t_1, t_2) =$

$= t_1^2 t_2 + \dfrac{1}{2} t_1 t_2^2 + \dfrac{2}{3} t_2^3$. **18.632.** $K_{XZ}(t_1, t_2) = t_1 + 4t_2$. **18.633.** $K_{XY}(t_1, t_2) = \varphi(t_2) K_X(t_1, t_2) + \psi(t_2) \dfrac{\partial^2}{\partial t_2^2} K_X(t_1, t_2)$. **18.634.** $f_{XX'}(x,y) = \dfrac{1}{2\pi\sqrt{2}} \exp\left\{-\dfrac{1}{2}(x-1)^2 + \dfrac{1}{4}y^2\right\}$. **18.635.** Не является.

18.636. Указание. Так как $m_X(t) = 0$ и $K_X(t, t+\tau) = \cos\omega\tau$, то процесс $X(t)$ стационарен в широком смысле. Для доказательства второго утверждения показать, что существует сдвиг τ такой, что, например, $\mathrm{P}\{X(0) = 1\} \neq \mathrm{P}\{X(0+\tau) = 1\}$, т.е. уже одномерный закон распределения процесса $X(t)$ не является инвариантным относительно произвольного сдвига по времени.

18.637. $m_Y(t) = \lambda$, $K_Y(t, t+\tau) = \begin{cases} \lambda(1-\tau), & 0 < \tau \leqslant 1, \\ 0, & 1 < \tau. \end{cases}$

Указание. При вычислении $K_Y(t, t+\tau) = \mathrm{M}[Y(t)Y(t+\tau)] - m_Y(t)m_Y(t+\tau)$ рассмотреть два случая: 1) $\tau > 1$ и 2) $0 < \tau \leqslant 1$ — и воспользоваться тем, что пуассоновский процесс — процесс с независимыми приращениями. **18.638.** $2(K_X(0) - K_X(\tau))$.

18.639. $K_X(t, t+\tau) = \sum\limits_{k=0}^{n} D_k \cos\omega_k\tau$; $D_X = \sum\limits_{k=0}^{n} D_k$. **18.640.** $\mathrm{D}[X(t)] = \alpha_2 \sum\limits_{k=0}^{n} D_k$, $\mathrm{D}\left[\dfrac{dX(t)}{dt}\right] = \alpha_2 \sum\limits_{k=0}^{n} \omega_k D_k$, где $\alpha_2 = \int\limits_{-\infty}^{\infty} \cos^2\varphi f_\Phi(\varphi)\, d\varphi$.

18.641. $\mathrm{D}\left[\int\limits_0^t X(t)\, dt\right] = \alpha_2 \sum\limits_{k=0}^{m} \dfrac{D_k}{\omega_k}$, где α определено в предыдущей задаче. **18.643.** Не является. **18.644.** Является. **18.646.** $K_X(\tau) = e^{-2\lambda|\tau|}$, $\tau = t_2 - t_1$. Указание. Для вычисления ковариационной функции описать закон распределения двумерного случайного вектора $(X(t), X(t+\tau))$. **18.647.** $K_X(\tau) = p(1-p)e^{-\lambda|\tau|}$. Указание. При вычислении $K_X(t, t+\tau)$ воспользоваться формулой полного математического ожидания на разбиении $\{A, \overline{A}\}$, где событие $A = \{$за время τ процесс не совершит ни одного скачка$\}$:

$$\mathrm{M}[X(t) \times X(t+\tau)] = \mathrm{M}[X(t)\, X(t+\tau)/A]\, \mathrm{P}(A) + \mathrm{M}[X(t)\, X(t+\tau)/\overline{A}]\, \mathrm{P}(\overline{A}).$$

18.648. $m_X(t) = 0$, $D_X(t) = \sigma^2$, $K_X(\tau) = \sigma^2 e^{-\lambda|\tau|}$. Указание. См. указание к предыдущей задаче. **18.649.** Непрерывен, но не дифференцируем. Указание. См. указание к задаче 18.623. **18.650.** Дифференцируем бесконечное число раз. Если $Y(t) = \dfrac{d^2X(t)}{dt^2}$, то $D_Y(t) = \left.\dfrac{d^4 K_X(\tau)}{d\tau^4}\right|_{\tau=0} = 8\alpha^2\sigma_X^2$. **18.651.** Дифференцируем дважды. $D_Y = 2\alpha^2\sigma_X^2$.

18.652. Не дифференцируем. **18.653.** $K_Y(t,\,t+\tau) = K_X(\tau) + \dfrac{d^2}{d\tau^2}K_X(\tau) + \dfrac{d^4}{d\tau^4}K_X(\tau)$. **18.655.** $K_{XY}(\tau) = \dfrac{\partial}{\partial\tau}K_X(\tau)$. **18.656.** $K_{XY}(t,\,t+\tau) = \sigma^2\dfrac{\alpha^2+\beta^2}{\beta}e^{-\alpha|t+t_0|}\sin\beta|\tau+t_0|$. Указание. Использовать результат предыдущей задачи. **18.657.** $K_{XY}(t,\,t+\tau) = -\dfrac{1}{3}\sigma^2\alpha^2e^{-\alpha|\tau|}(1+\alpha|\tau|-\alpha^2\tau^2)$. **18.659.** Указание. В двойном интеграле, определяющем дисперсию $\sigma_Y^2(t)$, сделать замену переменных и изменить порядок интегрирования. **18.660.** $D_Y(t) = 4t - 2(1-e^{-2t})$. Указание. Использовать результат предыдущей задачи. **18.661.** Процессы не являются стационарно связанными. Указание. При вычислении ковариационной функции использовать результат задачи 18.658. **18.662.** $m_Y = 0$, $K_Y(\tau) = \sigma_X^2\dfrac{\alpha^2+\beta^2}{\beta}\left(\beta\cos\beta\tau - \dfrac{\alpha}{\beta}\sin\beta|\tau|\right)$, $D_Y = \sigma_X^2(\alpha^2+\beta^2)$.
Указание. При доказательстве нормальности воспользоваться определением производной случайного процесса и свойствами линейной функции нормальной случайной величины.

18.663. $\mathrm{P}\left\{\left|\dfrac{dX(t)}{dt}\right| < \sqrt{13}\right\} = 2\Phi\left(\dfrac{1}{2}\right) - 1 \approx 0{,}383$.

18.664. $f_2(x,\,y/t,\,t) = \dfrac{1}{2\pi\alpha\sigma_X^2\sqrt{2}}\exp\left\{-\dfrac{1}{2\sigma_X^2}\left(x^2+\dfrac{y^2}{\alpha^2}\right)\right\}$.

Указание. Использовать результат задачи 18.662.

18.669. $K_X(\tau) = \sigma_X^2\cos\left(\dfrac{\omega_1+\omega_2}{2}\tau\right)\dfrac{\sin\left(\dfrac{\omega_2-\omega_1}{2}\tau\right)}{\dfrac{\omega_2-\omega_1}{2}\tau}$.

18.670. $K_X(\tau) \to \sigma_X^2\cos\omega_1\tau$ при $\omega_2\to\omega_1$, что соответствует гармоническому колебанию на частоте ω_1. **18.671.** $K_X(\tau) = \dfrac{b}{2}e^{-\alpha|\tau|}$, $D_X = \dfrac{b}{2}$, $\Delta\tau = 2/\alpha$. **18.672.** $K_X(\tau) = \dfrac{b}{4}e^{-\alpha|\tau|}(1+\alpha|\tau|)$, $D_X = \dfrac{b}{4}$, $\Delta\tau = 2/\alpha$.

18.673. $K_X(\tau) = \dfrac{a\omega_0}{1+\omega_0^2\tau^2}$, $D_X = a\omega_0$, $\Delta\tau = \dfrac{\pi}{\omega_0}$.

18.674. $K_X(\tau) = \dfrac{4a}{\omega_0\tau^2}\left(\dfrac{\sin\omega_0\tau}{\omega_0\tau} - \cos\omega_0\tau\right)$, $D_X = \dfrac{4}{3}a\omega_0$, $\Delta\tau \geqslant \dfrac{3\pi}{4\omega_0}$.

18.675. $S_X(\omega) = \dfrac{D_X}{\alpha\sqrt{\pi}}\exp\left\{-\dfrac{\omega^2}{4\alpha^2}\right\}$, $\Delta\omega = 2\alpha\sqrt{\pi}$, $\Delta\tau = \dfrac{\sqrt{\pi}}{\alpha}$.

18.676. $S_X(\omega) = \dfrac{\alpha D_X}{\pi}\left(\dfrac{1}{\alpha^2+(\beta-\omega)^2} + \dfrac{1}{\alpha^2+(\beta+\omega)^2}\right)$, $\Delta\omega = 2\pi$.

18.678. Указание. Использовать результат предыдущей задачи и правила дифференцирования несобственных интегралов по параметру.

18.679. $b\alpha^2/4$. **18.680.** $S_Y(\omega) = (a^2 + b^2\omega^2)\,S_X(\omega)$. **18.681.** $S_X(\omega) = \dfrac{2\sigma_X^2}{\pi\omega^2 t_0}(1 - \cos\omega t_0)$, $\Delta\omega = \dfrac{2\pi}{t_0}$, процесс не дифференцируем. **18.682.** $S_X(\omega) = \dfrac{2\lambda p\,(1-p)}{\pi\,(\lambda^2 + \omega^2)}$, $\Delta\omega = \pi\lambda$, процесс не дифференцируем. **18.683.** $S_X(\omega) = \dfrac{4\alpha^3 D_X}{\pi\,(\alpha^2 + \omega^2)^2}$, $\Delta\omega = \dfrac{1}{2}\pi\alpha$, процесс дифференцируем. **18.684.** $S_X(\omega) = \dfrac{\sigma_X^2}{\pi}\,\dfrac{4\alpha\,(\alpha^2+\beta^2)}{\pi\,(\alpha^2+\beta^2+\omega^2)^2 - 4\beta^2\omega^2}$, $\Delta\omega = \dfrac{\pi\,(\alpha^2+\beta^2)}{2\alpha}$, процесс дифференцируем. **18.685.** $m_Y = 0$, $K_Y(t,\,t{+}\tau) = 2K_X(\tau) - K_X(\tau + t_1) - K_X(\tau - t_1)$. Стационарность сохраняется. **18.686.** $S_Y(\omega) = 4\sin^2\dfrac{\omega t_1}{2}\,S_X(\omega)$. **18.687.** $6K_X(0) + 2K_X(2t_1) + 8K_X(t_1)$. **18.688.** $D_Y = D_X\left(6 + 2e^{-4\alpha^2 t_1^2} - 8e^{-\alpha^2 t_1^2}\right)$. **18.689.** $S_Z(\omega) = \dfrac{2\,(\lambda_1+\lambda_2)}{\pi\,(\omega^2 + 4\,(\lambda_1+\lambda_2)^2)}$. Указание. Использовать результаты задач 18.692, 18.646 и примера 7. **18.690.** $m_Y = 3$, $D_Y = 9/5$. **18.691.** $m_Y = 0$, $D_Y = \dfrac{2\beta}{\omega_0}\operatorname{arctg}\dfrac{\omega_0}{\beta}$. **18.692.** $D_Y = \pi\beta c^2/2$. Указание. В выражении для D_Y, полученном в предыдущей задаче, положить $\sigma_X^2 = c^2\omega_0$ и перейти к пределу при $\omega_0 \to \infty$. **18.693.** $\dfrac{\sigma_Y^2}{\sigma_X^2} = \dfrac{\pi}{\beta}\,e^{\alpha^2\beta^2}(1 - \Phi(\sqrt{2}\,\alpha\beta))$; в указанном частном случае это отношение равно $\approx 0{,}671$.

Указание. Использовать определенный интеграл

$$\int\limits_0^{+\infty}\frac{e^{-\mu^2x^2}\,dx}{x^2+\beta^2} = \frac{\pi}{\beta}e^{\beta^2\mu^2}(1-\Phi(\beta\mu\sqrt{2})).$$

18.694. $\sigma_Y^2 = \dfrac{a^2\pi}{4\alpha}$, $\dfrac{\sigma_Y^2}{\sigma_X^2} = 2\alpha^2$.

18.695. $m_Y = 2$, $S_Y(\omega) = \dfrac{2}{\pi}\sigma_X^2\,\dfrac{\omega^2+4}{(\omega^2+1)(\omega^2+25)}$, $D_Y = \dfrac{3}{10}\,\dfrac{\sigma_X^2}{\pi}$. **18.696.** $D_Y = \dfrac{1}{\pi}\left(\operatorname{arctg}\dfrac{\omega_2}{\lambda} - \operatorname{arctg}\dfrac{\omega_1}{\lambda}\right)$. **18.697.** $S_Y(\omega) = \dfrac{2\sigma_X^2}{\pi}\alpha^3\,\dfrac{\omega^2+\beta^2}{(\omega^2+\alpha^2)^2}$, $D_Y = \sigma_X^2(\alpha^2+\beta^2)$, $\Delta\omega = \dfrac{\pi\alpha}{4}(\alpha^2+\beta^2)$. Указание. Использовать результат задачи 18.672. При вычислении D_Y применить теорию вычетов. **18.698.** $D_Y = D_X\left(1 - \dfrac{R}{L\omega_0}\operatorname{arctg}\dfrac{\omega_0 L}{R}\right)$. **18.699.** $D_Y = \dfrac{\sigma_X^2}{\pi t_0}(1 - e^{-\beta t_0})$. **18.700.** $D_Y = \dfrac{\lambda^2}{(\lambda+\beta)^2}$. **18.701.** $D_Y = \dfrac{c^2\pi}{4\beta\omega_0^2}$.

ГЛАВА 19

19.1. 0, 3, 5, 6, 8, 9, 11, 11, 11, 12, 12, 13, 14, 15, 16, 16, 19 — вариационный ряд,

x_i	0	3	5	6	8	9	11	12	13	14	15	16	19
n_i	1	1	1	1	1	1	3	2	1	1	1	2	1

— статистический ряд, $w = 19$.

19.2. 15, 16, 16, 16, 16, 17, 17, 17, 17, 18, 18, 18, 18, 18, 19, 19 — вариационный ряд,

x_i	15	16	17	18	19
n_i	1	4	4	5	2

— статистический ряд, $w = 4$.

19.5. $w = 63$, $b = 9$, $k = 7$, $n = 50$, результаты группировки сведены в таблицу:

Номер интервала	Границы интервала	x_i^*	n_i^*	$\sum\limits_{j=1}^{i} n_j^*$	$\frac{n_i^*}{n}$	$\frac{1}{n}\sum\limits_{j=1}^{i} n_j^*$
1	14–23	18,5	2	2	0,04	0,04
2	23–32	27,5	3	5	0,06	0,1
3	32–41	36,5	6	11	0,12	0,22
4	41–50	45,5	17	28	0,34	0,56
5	50–59	54,5	10	38	0,2	0,76
6	59–68	63,5	9	47	0,18	0,94
7	68–77	72,5	3	50	0,06	1

19.6. $w = 13{,}5$, $b = 2$, $k = 7$, $n = 65$, результаты группировки сведены в таблицу:

Номер интервала	Границы интервала	x_i^*	n_i^*	$\sum\limits_{j=1}^{i} n_j^*$	$\frac{n_i^*}{n}$	$\frac{1}{n}\sum\limits_{j=1}^{i} n_j^*$
1	8,4–10,4	9,4	3	3	0,0462	0,0462
2	10,4–12,4	11,4	7	10	0,1077	0,1539
3	12,4–14,4	13,4	13	23	0,2	0,3539
4	14,4–16,4	15,4	21	44	0,3231	0,6770
5	16,4–18,4	17,4	17	61	0,2615	0,9385
6	18,4–20,4	19,4	2	63	0,0308	0,9693
7	20,4–22,4	21,4	2	65	0,0308	1,0001

19.7. График эмпирической функции распределения изображен на рис. 59, гистограмма частот — на рис. 60, полигон частот — на рис. 61. **19.8.** График эмпирической функции распределения изображен на рис. 62, гистограмма частот — на рис. 63, полигон частот — на рис. 64. **19.9.** График эмпирической функции распределения изображен на рис. 65, гистограмма частот — на рис. 66, полигон частот — на рис. 67. **19.10.** График эмпирической функции распределения изображен на рис. 68, гистограмма частот — на рис. 69, полигон частот — на рис. 70. **19.11.** а) Гистограмма изображена на рис. 71, полигон

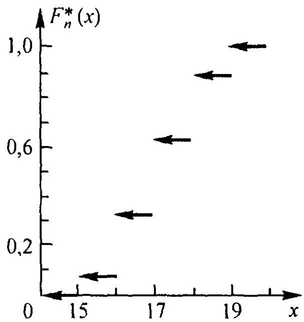


Рис. 59

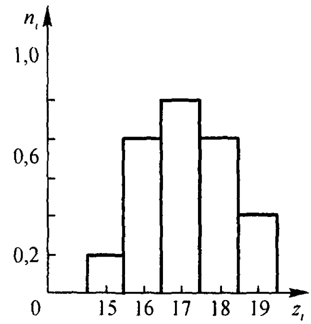


Рис. 60

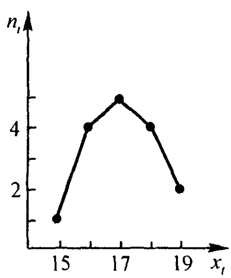


Рис. 61

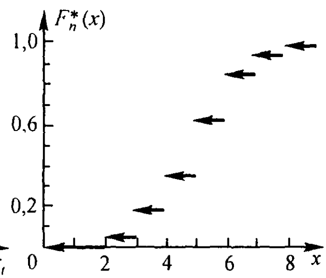


Рис. 62

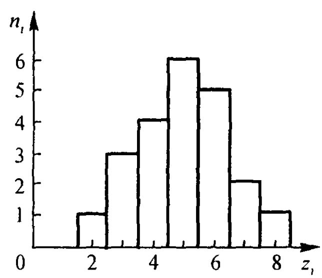


Рис. 63

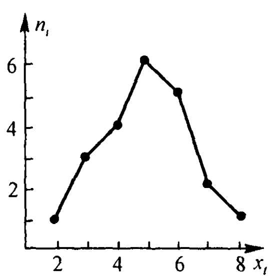


Рис. 64

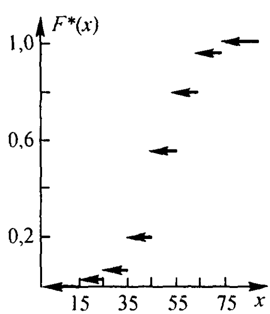


Рис. 65

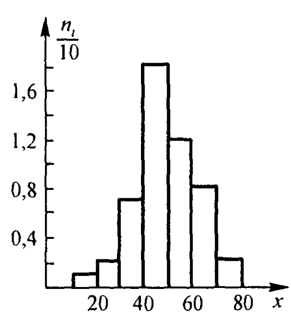


Рис. 66

накопленных частот — на рис. 72; б) гистограмма изображена на рис. 73, полигон накопленных частот — на рис. 74; в) гистограмма изображена на рис. 75, полигон накопленных частот — на рис. 76. **19.23.** $\triangleleft$ Из определения выборки и формулы (1) следует, что $F_n^*(x)$ есть относительная частота события $A = \{z_i < x\}$ в n независимых испытаниях, причем в каждом испытании $\mathrm{P}(A) = F_X(x)$. Утверждение теоремы следует теперь из теоремы Бернулли (гл. 18, § 5). ▷ **19.24.** $d_X^* = h_X^* = 3$, $\overline{x} = 3{,}5$, $D_X^* = 3{,}65$. **19.25.** $d_X^* = 3{,}1$, $h_X^* = 2{,}5$, $\overline{x} \approx 2{,}39$, $D_X^* \approx 0{,}43$.

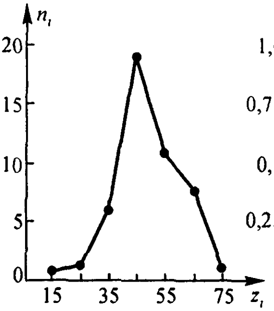


Рис. 67

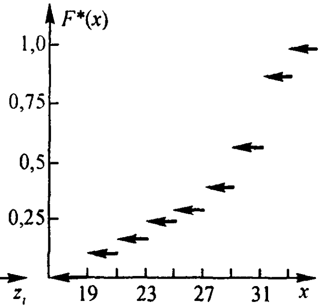


Рис. 68

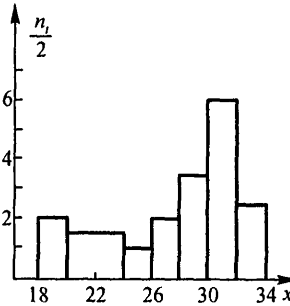


Рис. 69

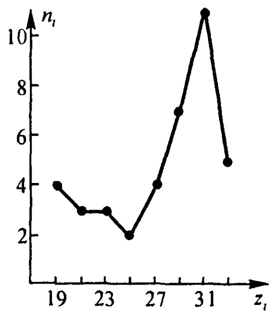


Рис. 70

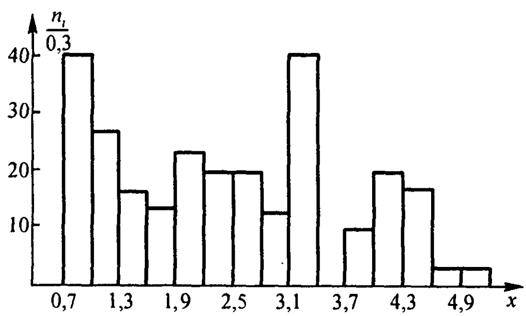


Рис. 71

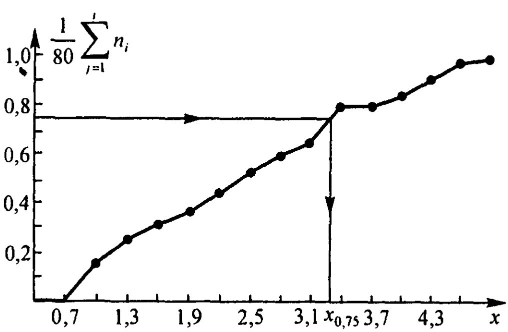


Рис. 72

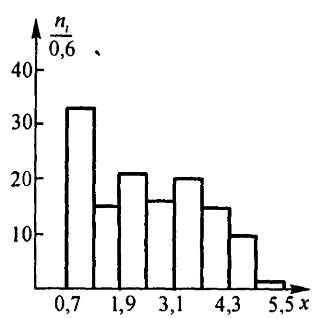


Рис. 73

19.26. а) $d_X^* = 5$, $h_X^* = 4$, $\overline{x} \approx 4{,}14$, $D_X^* \approx 5{,}84$. б) $d_X^* = 5$, $h_X^* = 4$, $\overline{x} \approx 4{,}57$, $D_X^* \approx 11{,}10$. Для выборки б) среднее и дисперсия увеличились, а мода и медиана не изменились. **19.31.** $\overline{x} \approx 6{,}54$, $d_X^* = 6$, $h_X^* = 6$, $h_X^* = 6$, $D_X^* \approx 7{,}34$. **19.32.** $\overline{x} \approx 11{,}78$, $d_X^* = 10$, $h_X^* = 10$, $D_X^* \approx 32{,}39$.

19.33. $\overline{x} \approx 10{,}49$, $d_x^* = 10$, $h_x^* = 10$, $D_x^* \approx 6{,}29$. **19.34.** б) Указание. Найти значение параметра a, при котором $\sum\limits_{i=1}^{n}(x_i - a)^2$ принимает

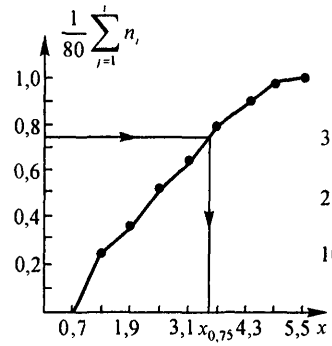


Рис. 73

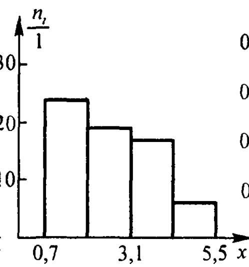


Рис. 74

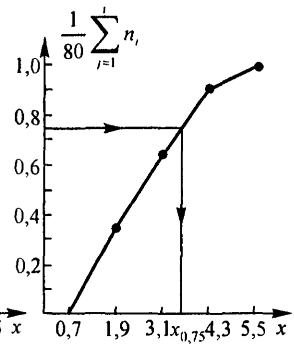


Рис. 75

наименьшее значение. **19.35.** $\triangleleft$ Пусть $x^{(1)} \leqslant x^{(2)} \leqslant \ldots \leqslant x^{(n)}$ — вариационный ряд выборки. Рассмотрим сумму $\sum\limits_{i=1}^{n}\left|x^{(i)} - a\right|$, $a \in \mathbb{R}$, и покажем, что она достигает минимума при $a = h_x^*$. Положим $a = h_x^* + \alpha$ и рассмотрим случай $\alpha > 0$. Тогда для некоторого $l < n$ выполняется соотношение $x^{(l)} \leqslant a \leqslant x^{(l+1)}$, причем $2l > n$. Получаем

$$\sum_{i=1}^{n}\left|x^{(i)} - \alpha\right| = \sum_{i=1}^{l}(\tilde{h}_x + \alpha - x^{(i)}) + \sum_{i=l+1}^{n}(x^{(i)} - \tilde{h}_x - \alpha) =$$

$$= \sum_{i=1}^{n}\left|x^{(i)} - \tilde{h}_x\right| + \alpha\,(2l - n) > \sum_{i=1}^{n}\left|x^{(i)} - \tilde{h}_x\right|.$$

Случай $\alpha < 0$ рассматривается аналогично. $\triangleright$ **19.36.** $\overline{x}$, d_x^*, h_x^* изменятся на величину a, D_x^* не изменится. **19.37.** $\overline{x}$, d_x^*, h_x^* увеличатся (уменьшатся) в k раз, D_x^* — в k^2 раз. **19.38.** 1) $\overline{x} = 17{,}7$, $D_x^* \approx 20{,}08$; 2) $\overline{x} \approx 17{,}20$, $D_x^* \approx 19{,}17$. **19.39.** 1) $\overline{x} \approx 70{,}4$, $D_x^* \approx 9{,}6$; 2) $\overline{x} \approx 69{,}8$, $D_x^* \approx 9{,}2$. **19.40.** 1) $\overline{x} \approx 5{,}96$, $D_x^* \approx 1{,}35$; 2) $\overline{x} \approx 5{,}94$, $D_x^* \approx 1{,}31$. **19.49.** $\overline{x} \approx 17{,}92$, $D_x^* \approx 52{,}15$, $a_x^* \approx 0{,}89$, $e_x^* \approx 7{,}55$. **19.50.** $\overline{x} = 61{,}6$, $D_x^* \approx 19{,}53$, $a_x^* \approx -0{,}47$, $e_x^* \approx -0{,}25$. **19.51.** $\overline{x} = 35{,}72$, $D_x^* \approx 19{,}96$, $a_x^* \approx -0{,}03$, $e_x^* \approx -1{,}09$. **19.52.** $\overline{x} \approx 0{,}05$, $D_x^* \approx 0{,}02$, $a_x^* \approx -0{,}60$, $e_x^* \approx 0{,}37$. **19.53.** $\overline{x} = 9{,}245$, $D_x^* \approx 0{,}02$, $a_x^* \approx 0{,}44$, $e_x^* \approx -0{,}61$. **19.54.** $\overline{x} = 8{,}6$, $D_x^* \approx 377{,}04$, $a_x^* \approx -0{,}12$, $e_x^* \approx -0{,}18$. **19.55.** $x_{0,1}^* = 4{,}4$, $x_{0,5}^* = 6$, $x_{0,9}^* = 7{,}4$. **19.63.** 0,866. **19.64.** 0,806. **19.65.** 0,997. **19.66.** $k_{xy}^* \leqslant 12$. Указание. Показать, что справедливо неравенство $k_{xy}^* \leqslant s_x s_y$. **19.68.** $-0{,}6$. **19.69.** 0,55. **19.70.** $r \approx 0{,}535$,

$y = -1{,}43 + 0{,}43x$, $x = 5 + 1{,}5y$ (рис. 77). **19.71.** $r \approx 0{,}806$, $y = 0{,}5 + 0{,}5x$, $x = 2{,}5 + 1{,}3y$ (рис. 77). **19.72.** $r \approx 0{,}983$, $y = 0{,}5 + 0{,}77x$, $x = -0{,}5 + 1{,}3y$ (рис. 79). **19.73.** $r \approx 0{,}837$, $y = 8{,}96 + 0{,}44x$, $x = 10{,}96 + 1{,}6y$. **19.74.** $r \approx 0{,}743$, $y = 12{,}25 + 0{,}8x$. Диаграммы

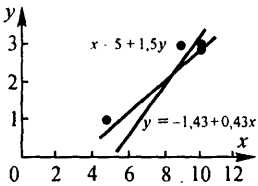


Рис. 77

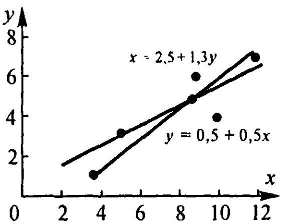


Рис. 78

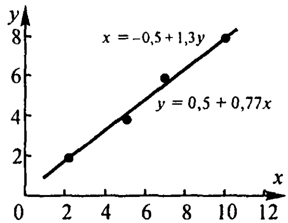


Рис. 79

рассеяния и графики приведены на рис. 80. Уравнение регрессии для всех наборов данных одно и то же, но точки на графиках расположены совершенно по-разному. Пример показывает необходимость предварительного анализа данных, в частности, построения диаграммы рассеяния. Уравнение регрессии: $y = 3 + 0{,}5x$. **19.75.** $\beta^*_{1uv} = \dfrac{c}{a}\beta^*_{1xy}$, $\beta^*_{0uv} = c\beta^*_{0xy} - \dfrac{ab}{c}\beta^*_{1xy} + d$. **19.76.** $v = ru$, $u = rv$. **19.77.** Выборочные коэффициенты корреляции равны соответственно 0,403 для первых девяти городов и 0,995 для десяти городов. Уравнение линейной регрессии $y = 2{,}716 + 0{,}43x$. Диаграмма рассеивания приведена на рис. 79 [2]). **19.80.** $-0{,}137$. **19.81.** $r \approx -0{,}44$, $y = 14 - 0{,}04x$, $x = 101{,}78 - 4{,}57y$. **19.82.** $r \approx 0{,}775$, $y = 10{,}1 + 0{,}72x$, $x = 5{,}87 + 0{,}83y$. **19.83.** $r \approx 0{,}921$,

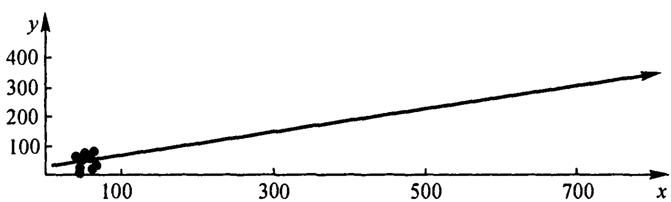


Рис. 80

$y = 25{,}3 + 0{,}92x$, $x = -8{,}5 + 0{,}98y$. **19.84.** $r \approx 0{,}848$, $y = -10{,}17 + 1{,}74x$, $x = 6{,}31 + 0{,}41y$.

[2]) Этот пример показывает, что одно наблюдение может существенно повлиять на величину коэффициента корреляции. Нью-Йорк входит в другую генеральную совокупность — городов с бо́льшим населением.

19.85.

```
      DIMENSION X(18), Y(18)
      READ (5, 2) Y
      DO 4 I=1, 18
    4 X(I)=.1*(I-1)
    2 FORMAT (18F4.3)
      N=18
      CALL STAXY1 (X,Y,N,SXR,SYR,QX,QY,QXY)
      CALL STAXY2 (QX,QY,QXY,N,DX,DY,KXY,RXY)
      CALL LINREG (SXR,SYR,QX,QY,QXY,BØYX,B1YX,BØXY,B1XY)
      PRINT 3,RXY,BØXY,B1YX,BØXY,B1XY
    3 FORMAT (//5X,'RXY=',2X,F5.2,2X,'BØXY=',2X,F5.2,2X,
     *'B1YX=',2X,F5.2,'BØXY=',2X,F5.2,'B1XY=',2X,F5.2)
      STOP
      END
```

$\overline{x} = 0{,}85$, $\overline{y} = 0{,}23$, $s_x^2 = 1{,}805$, $s_y^2 = 7{,}19 \cdot 10^{-3}$, $r = -0{,}503$, $y = 0{,}26 - 0{,}032x$, $x = 2{,}68 - 7{,}97y$. **19.86.** $\overline{x} = 5{,}5$, $\overline{y} = 2{,}94$, $s_x^2 = 7{,}92$, $s_y^2 = 6{,}80$, $r = 0{,}964$, $y = -1{,}97 + 0{,}89x$, $x = 2{,}44 + 1{,}04y$. **19.87.** $\overline{x} = 0{,}475$, $\overline{y} = 39{,}5$, $s_x^2 = 0{,}087$, $s_y^2 = 234{,}33$, $r = 0{,}981$, $y = 19{,}39 + 50{,}77x$, $x = 0{,}11 + 2{,}57y$. **19.88.** $\overline{x} = 8{,}5$, $\overline{y} = 0{,}79$, $s_x^2 = 28{,}5$, $s_y^2 = 0{,}32$, $r = 0{,}972$, $y = -0{,}11 + 0{,}1x$, $x = 15{,}44 - 6{,}94y$. **19.89.** $\overline{x} = 4{,}5$, $\overline{y} = 0{,}75$, $s_x^2 = 0{,}907$, $s_y^2 = 0{,}133$, $r = 0{,}932$, $y = -8{,}53 + 0{,}36x$, $x = 2{,}67 + 2{,}44y$. **19.90.** $\overline{x} = 0{,}85$, $\overline{y} = 2{,}07$, $s_x^2 = 0{,}23$, $s_y^2 = 1{,}64$, $r = 0{,}938$, $y = -7{,}67 + 2{,}52x$, $x = 0{,}29 + 0{,}35y$. **19.96.** $\tilde{\sigma}^2 = D_X^* = \frac{1}{n}\sum (x_i - \overline{x})^2$, $\tilde{h}_X = h_X^*$, $\tilde{a}_X = a_X^*$, $\tilde{e}_X = e_X^*$. **19.97.** $\tilde{\alpha}_l = \alpha_l^* = \frac{1}{n}\sum x_l^l$. Указание. Для доказательства состоятельности вычислить дисперсию $\tilde{\alpha}_l$ и использовать теорему 1. **19.98.** Смещение равно $-\sigma^2/n$. Указание. Выполнить преобразование $D_X^* = \frac{1}{n}\sum (x_i - \overline{x})^2 = \frac{1}{n}\sum ((x_i - m) - (\overline{x} - m))^2 = \frac{1}{n}\sum (x_i - m)^2 - (\overline{x} - m)^2$, затем найти математическое ожидание. **19.100.** ◁ Воспользоваться преобразованием D_X^* из указания к задаче 19.98. По теореме Чебышева (гл. 18, § 5) первое слагаемое $\frac{1}{n}\sum (x_i - m)^2$ при $n \to \infty$ сходится по вероятности к $\mathrm{M}[(x_i - m)^2] = \sigma^2$, а второе слагаемое $(\overline{x} - m)^2$ при $n \to \infty$ сходится по вероятности к нулю, так как $\overline{x} > \xrightarrow[n\to\infty]{\mathrm{P}} m$. Отсюда следует, что $D_X^* > \xrightarrow[n\to\infty]{\mathrm{P}} \sigma^2$, т.е. является состоятельной оценкой дисперсии. Состоятельность s^2 как оценки дисперсии следует из соотношения $s^2 = \frac{n}{n-1} D_X^*$. ▷ **19.102.** $\tilde{m}_1 = x_1$ является несмещенной и состоятельной оценкой m. **19.104.** $\tilde{a} = x^{(1)}$ является смещенной и состоятельной оценкой. Указание. Для нахождения плотности распределения $\tilde{a}$ воспользоваться результатом решения задачи 18.525. **19.105.** Указание. Использовать решение задачи 18.525. Вычислить $\mathrm{M}\,[x^{(1)}] = \dfrac{1}{n+1}$, $\mathrm{M}\,[x^{(n)}] = \dfrac{n}{n+1}$, $\mathrm{D}\,[x^{(1)}] = \mathrm{D}\,[x^{(n)}] = \dfrac{n}{(n+2)(n+1)^2}$. Показать, что $\mathrm{D}\left[\dfrac{1}{2}\left(x^{(1)} + x^{(n)}\right)\right] = O\left(\dfrac{1}{n^2}\right)$. **19.106.** $\tilde{k}_{XY} = k^*_{XY} = \dfrac{1}{n}\sum (x_i - \overline{x})(y_i - \overline{y})$. Несмещенная оценка имеет вид $\tilde{k}_{XY} = \dfrac{1}{n-1}\sum (x_i - \overline{x})(y_i - \overline{y})$. Указание. См. задачу 19.59. Использовать преобразование $x_i - \overline{x} = (x_i - m_X) + \dfrac{1}{n}\sum (x_i - m_X) = \mathring{x}_i + \dfrac{1}{n}\sum \mathring{x}_i$, $y_i - \overline{y} = \mathring{y}_i + \dfrac{1}{n}\sum \mathring{y}_i$. Тогда $\mathrm{M}\,[(x_i - \overline{x})(y_i - \overline{y})] = K_{XY} - \dfrac{2}{n} K_{XY} + \dfrac{1}{n} K_{XY} = \dfrac{n-1}{n} K_{XY}$, так как $K_{X_iY_j} = 0$ при $i \neq j$, $K_{X_iY_i} = K_{XY}$. $\mathrm{D}\,[K^*_{XY}] = \dfrac{1}{n^2}\sum \mathrm{D}\,[(X_i - \overline{x})(Y_i - \overline{y})] \to 0$ при $n \to \infty$. **19.107.** Смещение $\tilde{\theta}^2$ равно $\mathrm{D}\,[\tilde{\theta}]$. Указание. Рассмотреть тождество $\mathrm{D}\,[\tilde{\theta}] = \mathrm{M}\,[(\tilde{\theta} - \theta)^2]$. **19.109.** б) $p = 1/2$. **19.111.** Указание. Использовать указание к задаче 19.107. **19.112.** Указание. Использовать второе неравенство Чебышева (см. гл. 18, § 5). **19.116.** $\dfrac{n}{2\sigma^n}$. **19.117.** Указание. Использовать указание к задаче 19.105 и найти дисперсию выборочного среднего. **19.119.** Указание. Использовать результат решения задачи 19.107. **19.120.** x/n. **19.121.** x/n. **19.122.** $\overline{x}$. **19.123.** $\overline{x}$. **19.124.** $\tilde{p} = \dfrac{\sum x_i}{\sum n_i}$, $\mathrm{D}\,[\tilde{p}] = \dfrac{p(1-p)}{\sum n_i}$. **19.125.** $\tilde{a} = x^{(1)} = \min\limits_{1 \leqslant i \leqslant m} \{x_i\}$, $\tilde{b} = x^{(n)} = \max\limits_{1 \leqslant i \leqslant n} \{x_i\}$. Указание. Точка максимума $L\,(a, b)$ определяется графически: нужно построить область допустимых значений для a и b и линии уровня $L\,(a, b) = \dfrac{1}{(b-a)^n}$, $a \leqslant x \leqslant b$. **19.126.** $\tilde{m} = \dfrac{2}{3} \max\limits_i x_i$, $i = 1, 2, \ldots, n$. **19.127.** $\tilde{m} = \dfrac{\sum x_i/\sigma_i^2}{\sum l/\sigma_i^2}$; $\mathrm{D}\,[\tilde{m}] = \dfrac{1}{\sum 1/\sigma_i^2}$. **19.128.** $\tilde{p} = 1/k$, $\mathrm{M}\,[\tilde{p}] = -\dfrac{p \ln p}{1-p}$. **19.129.** $\tilde{p} = 1/k$, $\mathrm{M}\,[\tilde{p}] = -\dfrac{p \ln p}{1-p}$. **19.130.** $\tilde{p} = x/n$. Указание. Вероятность того, что событие A произойдет ровно x раз при n испытаниях, имеет распределение Паскаля: $\mathrm{P}\,[X = x] = C_{n-1}^{x-1} p^x (1-p)^{n-x}$, $n = x, x+1, \ldots$ Оценка $\tilde{p} = \dfrac{x-1}{n-1}$ является несмещенной:

$$\mathrm{M}\left[\frac{X-1}{n-1}\right]=$$

$$=\sum_{n=x}^{\infty}\frac{x-1}{n-1}\,C_{n-1}^{x-1}p^x(1-p)^{n-x}=p\sum_{n=x}^{\infty}C_{n-2}^{x-2}p^{x-1}(1-p)^{n-x};$$

при $l=x-1$, $m=n-1$ имеем $\mathrm{M}\left[\frac{X-1}{n-1}\right]=p\sum_{m=l}^{\infty}C_{m-1}^{l-1}p^l(1-p)^{m-l}=p$, следовательно, МП-оценка $\tilde{p}=x/n$ является смещенной. **19.131.** $\tilde{p}=x/n$. **19.132.** $\tilde{\lambda}=\frac{1}{n}\sum x_i=\overline{x}$. **19.133.** $\tilde{m}=\frac{1}{n}\sum x_i=\overline{x}$, $\tilde{\sigma}^2=D_x^*=\frac{1}{n}\sum(x_i-\overline{x})^2$. **19.134.** $\tilde{\lambda}=\frac{1}{\overline{x}}$. **19.135.** $\tilde{k}=\frac{1}{n}\sum x_i=\overline{x}$. **19.137.** $\overline{x}e^{-1/\overline{x}}$, где $\overline{x}=\frac{1}{n}\sum x_i$. Указание. $\mathrm{P}[X_1\geqslant 1]=\frac{1}{\lambda}\,e^{-\lambda}=\varphi(\lambda)$. Чтобы оценить функцию $\varphi(\lambda)$, необходимо представить $\varphi(\lambda)$ в виде непрерывной функции первых k моментов, т.е. в виде $\varphi(\alpha_1)=\alpha_1e^{-1/\alpha_1}$. **19.138.** 2,73; 169,64. **19.139.** −2,998; ≈ 1,659. **19.140.** 0,052; 9,025; 0,453. **19.141.** Указание. Использовать то, что $T(k)=U\Big/\sqrt{\frac{1}{k}\chi^2(k)}$ и $U^2=\chi^2(1)$. **19.143.** Статистика S_0^2 связана со случайной величиной $\chi^2(n)$ соотношением $S_0^2=\frac{1}{n}\sigma^2\chi^2(n)$. **19.144.** $\mathrm{M}[S^2]=\sigma^2$, $\mathrm{D}[S^2]=\frac{2}{n-1}\sigma^4$. Указание. Воспользоваться теоремой 2 и результатом задачи 19.116. **19.145.** $\mathrm{M}[T(n-1)]=0$, $\mathrm{D}[T(n-1)]=\frac{n-1}{n-3}$. **19.149.** Статистика S^2 связана со случайной величиной $\chi^2(n_1+n_2-2)$ соотношением

$$S^2=\frac{\sigma^2}{n_1+n_2-2}\chi^2(n_1+n_2-2),\qquad \mathrm{D}[S^2]=\frac{2\sigma^4}{n_1+n_2-2}.$$

19.150. $N\left(\frac{n_1m_1+n_2m_2}{n_1+n_2},\sqrt{\frac{n_1\sigma_1^2+n_2\sigma_2^2}{(n_1+n_2)^2}}\right)$. **19.157.** (18,35; 21,64), (17,42; 22,58). **19.158.** (498,35; 501,64), (497,42; 502,58). **19.159.** (28,14; 31,86), (26,64; 33,36). **19.160.** (16,63; 19,37), (15,76; 20,24). **19.164.** Указание. Найти $\mathrm{D}[\overline{X}]$ и установить, что $\frac{\overline{X}-m}{\sigma/\sqrt{n_1+n_2}}$ имеет распределение $N(0,\,1)$. Показать, что $S^2=\frac{\sigma^2}{n_1+n_2-2}\chi^2(n_1+n_2-2)$ и что $\dfrac{\overline{X} - m}{S/\sqrt{n_1 + n_2}}$ имеет распределение Стьюдента с n_1+n_2-2 степенями свободы. **19.165.** (17,96; 20,60), (17,22; 21,34). **19.166.** (490,91; 493,48), (490,18; 494,21). **19.167.** (28,52; 30,20), (27,98; 30,74). **19.168.** (17,01; 19,41), (16,27; 20,16). **19.169.** Указание. Использовать результат задачи 19.143. **19.171.** (2,70; 9,30), (2,45; 10,78). **19.172.** (7,27; 119,40), (10,32; 58,61). **19.173.** (96,81; 491,34), (9,84; 22,17). **19.174.** а) 0,68; б) $n \geqslant 385$. **19.175.** а) $n \geqslant 15$; б) $n \geqslant 4$. **19.176.** Указание. Найти распределение статистики $\overline{X}_1 - \overline{X}_2$. **19.177.** Указание. Использовать результаты решения задач 19.148, 19.149 и 19.151. **19.178.** (5,18; 6,82). **19.179.** (−11,85; 5,85); нет. **19.180.** а) (48,99; 63,41); б) (−15,1; 6,9), да. **19.181.** Указание. Использовать результат задачи 19.147 и формулу (17) § 2. **19.182.** (0,206; 5,757). **19.183.** (0,266; 0,454). **19.184.** (0,012; 0,038), $n \geqslant 88$. **19.185.** (0,39; 0,41), $n \geqslant 16\,231$. **19.186.** (0,088; 0,246). **19.187.** а) (0,225; 0,375); б) (0,210; 0,390). **19.188.** 0,026. **19.189.** Указание. Воспользоваться центральной предельной теоремой (гл. 18, § 5). **19.190.** Указание. Использовать результаты задач 19.189 и 19.108. **19.191.** (1,61; 2,39). **19.192.** (2,00; 2,60). **19.193.** (0,505; 0,810). **19.194.** (0,46; 0,86). **19.195.** (−0,93; 0,09); (−0,89; −0,12). **19.196.** (−0,01; 0,28), (−0,03; 0,25). **19.197.** (−0,14; 0,71), (0,02; 0,65). **19.199.** Гипотеза $H^{(1)}$ – $H^{(3)}$ — статистические; $H^{(4)}$ — не статистическая гипотеза, $H^{(1)}$ — простая, $H^{(2)}$ и $H^{(3)}$ — сложные гипотезы. **19.200.** Все гипотезы статистические; $H^{(1)}$ и $H^{(4)}$ — простые, $H^{(2)}$ и $H^{(3)}$ — сложные гипотезы. **19.201.** а) Число образцов с признаками коррозии имеет биномиальное распределение $B(20, p)$, $V = \{0, 1, 2, \ldots, 20\}$; б) $V_k = \{0\}$; в) ошибка первого рода: принимается решение, что антикоррозийное покрытие имеет эффективность 99 %, в то время как его эффективность составляет 90 %; ошибка второго рода: принимается решение, что антикоррозийное покрытие имеет эффективность 90 %, в то время как его эффективность составляет 99 %; $\alpha \approx$ $\approx 0{,}122$, $\beta \approx 0{,}182$. **19.202.** а) Ошибка первого рода; б) ошибка второго рода; в) и г) ошибки нет. **19.203.** H_0: объект является объектом противника; H_1: объект свой; ошибка второго рода — «ложная тревога». **19.204.** а) $V = \{0, 1, 2, \ldots, 5\}$, $V_k = \{0, 1, 2, 3, 4\}$ биномиальное $B(5, p)$; б) H_0: устройство функционирует правильно, $p = 0{,}99$; в) H_0: устройство функционирует неправильно, $p = 0{,}40$, ошибка второго рода — принятие неправильно функционирующего устройства; г) $\alpha \approx$ $\approx 0{,}05$, $\beta \approx 0{,}01$. **19.205.** а) Число дефектных изделий, $V = \{0, 1, \ldots, 10\}$, $V_k = \{2, 3, \ldots, 10\}$; б) биномиальное $B(10, p)$; в) H_0: $p = 0{,}05$, верно утверждение поставщика, H_1: $p = 0{,}10$, верно утверждение покупателя; г) ошибка первого рода: партия принята на условиях покупателя, в то время как верно утверждение поставщика; ошибка второго рода: партия принята на условиях поставщика; в то время как верно утверждение покупателя, $\alpha \approx 0{,}086$, $\beta \approx 0{,}736$. **19.206.** Да, $V_k = \{\overline{X} > 40{,}39\}$. **19.207.** $\alpha \approx 0{,}274$, $\beta \approx 0{,}107$. **19.208.** $\alpha \approx 0{,}548$, $\beta \approx 0{,}107$. **19.209.** $\alpha \approx 0{,}069$, $\beta \approx 0{,}910$. **19.210.** $n = 73$, $V_к = \{\overline{X} > 40{,}15\}$.

19.211. (39,912; ∞). **19.212.** а) Да; б) нет. **19.213.** а) Нет; б) нет. **19.214.** а) и б) Гипотеза принимается. **19.215.** а) Гипотеза отклоняется; б) гипотеза принимается. **19.216.** а) (8,647; 9,593), гипотеза отклоняется; б) (8,454; 10,146), гипотеза принимается. **19.217.** а) Нет; б) нет. **19.218.** а) $\beta = 0{,}198$; б) $\beta = 0{,}112$. **19.219.** Да. **19.220.** (0,1; ∞), да. **19.221.** Гипотеза принимается. **19.222.** (−1,11; 1,71), гипотеза принимается. **19.223.** $\lhd$ Доверительные интервалы для m_1 и m_2 будут $\bar{x}_1 \pm \dfrac{\sigma}{\sqrt{n_1}} u_{1-\alpha/2}$ и $\bar{x}_2 \pm \dfrac{\sigma}{\sqrt{n_2}} u_{1-\alpha/2}$. Эти интервалы не пересекаются, если

$$|\bar{x}_1 - \bar{x}_2| > u_{1-\alpha/2} \cdot \sigma \cdot \left(\frac{1}{\sqrt{n_1}} + \frac{1}{\sqrt{n_2}}\right). \qquad (*)$$

Критическая область (*) имеет размер меньше, чем критическая область

$$|\bar{x}_1 - \bar{x}_2| > u_{1-\alpha/2} \cdot \sigma \cdot \sqrt{1/n_1 + 1/n_2}. \qquad (**)$$

так как

$$\frac{1}{\sqrt{n_1}} + \frac{1}{\sqrt{n_2}} > \sqrt{1/n_1 + 1/n_2}. \quad \rhd$$

19.224. а) Нет; б) да. **19.225.** Нет. **19.226.** а) Гипотеза отклоняется; б) (0,132; 0,347). **19.227.** а) Гипотеза принимается; б) гипотеза отклоняется. **19.228.** Нет. **19.230.** а) Гипотеза принимается; б) гипотеза отклоняется. **19.231.** а) Гипотеза принимается; б) гипотеза отклоняется. **19.232.** а) Гипотеза принимается; б) гипотеза отклоняется. **19.233.** Нет. **19.234.** Да. **19.235.** Да. **19.236.** Нет. **19.238.** Гипотеза принимается. **19.239.** а) и б) Гипотеза принимается. **19.240.** Значения мощности критерия при $n = 36$ и $n = 100$ приведены в таблице, а кривые функции

m	$\mathcal{M}(V_к, m)$	
	$n = 36$	$n=100$
40,0	0,274	0,159
40,1	0,5	0,5
40,2	0,726	0,841
40,3	0,885	0,977
40,4	0,964	0,998
40,5	0,992	1,000

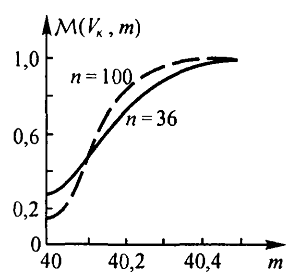


Рис. 81

мощности приведены на рис. 81. **19.241.** а) $1-\beta \approx 0{,}89$, $n \geqslant 27$; б) $1-\beta \approx$ $\approx 0{,}57$, $n \geqslant 53$. **19.242.** а) Значения мощности критерия приведены в таблице, кривые функции мощности приведены на рис. 82, $\beta = 0{,}94$; б) см. таблицу и рис. 82, $\beta = 0{,}74$. **19.243.** $1 - \beta = 0{,}323$, $n >$ > 294. **19.244.** Значения мощности критерия приведены в таблице, кривые приведены на рис. 83. Указание. Функция мощности симметрична относительно точки $m = 10$. **19.245.** Нет. **19.246.** а) Нет; б) да. **19.247.** а) $\alpha \approx 0{,}159$; б) $\beta \approx 0{,}159$. **19.248.** Нет. **19.249.** Да. **19.250.** Нет.

d_1	$\mathcal{M}(V_\kappa, d_1)$	
	$n = 100$	$n=15$
10,0	0	0,05
10,1	0,001	0,107
10,2	0,017	0,201
10,3	0,134	0,330
10,4	0,461	0,484
10,5	0,816	0,631

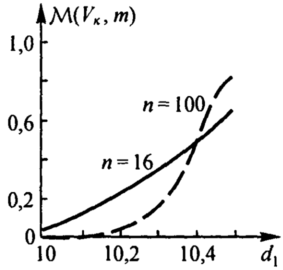


Рис. 82

19.251. Нет. **19.252.** Да. **19.253.** Да. **19.254.** а) и б) Незначим. а) **19.255.** а) Значим; б) незначим; в) значим. **19.256.** а) Нет; б) да; в) нет. **19.257.** а) Нет; б) (0,207; 0,91). **19.258.** а) Да; б) (−0,91; −0,814). **19.259.** а) Нет; б) (0,119; 0,653). **19.260.** а) Нет; б) (0,858; 1). **19.261.** $V_\kappa = (-\infty, -1) \cup (-x_\kappa, x_\kappa) \cup (1, \infty)$, где $x_\kappa = 0{,}05$; $1 - \beta = 0{,}126$.

m	$\mathcal{M}(V_\kappa, m)$	
	$n = 100$	$n=15$
9,0	0,851	0,999
9,1	0,770	0,994
9,2	0,670	0,979
9,3	0,556	0,938
9,4	0,440	0,851
9,5	0,323	0,705
9,6	0,225	0,516
9,7	0,147	0,323
9,8	0,092	0,171
9,9	0,061	0,079
10,0	0,050	0,050

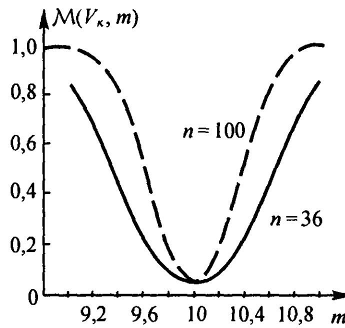


Рис. 83

Указание. Построить на одном графике плотности распределения $f(x)$ и нормального $N(0, 1)$. **19.262.** $V_\kappa = (-1, -x_\kappa) \cup (x_\kappa, 1)$, где $x_\kappa \approx 0{,}83$; $1 - \beta \approx 0{,}17$. **19.263.** $V_\kappa = (0; 0{,}1)$; $\mathcal{M}(V_\kappa, \theta) = 1 - e^{-0{,}1\theta}$, $\mathcal{M}(V_\kappa, 2) = 0{,}18$. **19.264.** а) $V_\kappa = \{\overline{x} > a_\alpha\}$; б) $V_\kappa = \{\overline{x} < b_\alpha\}$, где a_α и b_α — константы, зависящие от условия значимости α. **19.265.** При $p_0 < p_1$ имеем $V_\kappa = \{x > a_\alpha\}$; при $p_0 > p_1$ имеем $V_\kappa = \{x < b_\alpha\}$, где a_α и b_α — константы, зависящие от уровня значимости α, а x — наблюдаемое значение X. **19.266.** При $\sigma_0^2 > \sigma_1^2$ имеем $V_\kappa = \left\{\sum x_i^2 < b_\alpha\right\}$; при

$\sigma_0^2 < \sigma_1^2$ имеем $V_\kappa = \left\{\sum x_i^2 > a_\alpha\right\}$. **19.267.** $n \geqslant 9$. **19.268.** а) Да; б) нет. **19.270.** $|\overline{x}| > \overline{x}_\kappa$. **19.271.** $F_в = 0{,}115$, H_0 принимается, $\overline{x} = 8{,}42$, $s^2 = 13{,}32$. **19.272.** $F_в = 17{,}94$, H_0 отклоняется, $m_1 \neq m_2$, $m_1 \neq m_3$, $m_2 = m_3$. **19.273.** $F_в = 3{,}89$, H_0 принимается, $\overline{x} = 13{,}25$, $s^2 = 3{,}67$. **19.274.** $F_в = 2{,}86$, H_0 принимается, $\overline{x} = 22{,}08$, $s^2 = 33{,}23$. **19.275.** $F_в = 47{,}07$, H_0 отклоняется, $m_1 \neq m_4$, $m_1 \neq m_5$, $m_2 \neq m_5$. **19.276.** $F_в = 8{,}36$, H_0 отклоняется, $m_1 \neq m_2$, $m_1 = m_3$, $m_2 = m_3$. **19.277.** $F_в \approx 0$, H_0 принимается, $\overline{x} = 3{,}25$, $s^2 = 0{,}07$. **19.278.** $F_в = 1114{,}4$, H_0 отклоняется, $m_1 \neq m_2$, $m_1 \neq m_3$, $m_2 \neq m_3$. **19.279.** Указание. Обе части тождества $(X_{ik} - \overline{X}) \equiv (\overline{X}_k - \overline{X}) + (X_{ik} - \overline{X}_k)$ возвести в квадрат и просуммировать по i и по k; показать, что $\sum (\overline{X}_k - \overline{X})(X_{ik} - \overline{X}_k) = 0$. **19.281.** Да. ◁ Проверим гипотезу H_0 о том, что число появлений герба имеет биномиальное распределение с параметром $p = 1/2$:

$$\chi_в^2 = \frac{(20-25)^2}{25} + \frac{(30-25)^2}{25} = 2; \quad r - l - 1 = 2 - 0 - 1 = 1.$$

Так как $\chi_{0,90}^2(1) = 2{,}706$, гипотеза H_0 принимается. ▷ **19.282.** Нет. **19.284.** Да. **19.285.** Нет. **19.286.** Да. **19.287.** Да. **19.288.** H_0 отклоняется; $\chi_в^2 \approx 12{,}95$ (последние 6 интервалов объединяются). **19.289.** Да; $\chi_в^2 \approx 1{,}171$. **19.290.** H_0 принимается; $\chi_в^2 \approx 0{,}71$ (последние 5 интервалов объединяются). **19.291.** H_0 отклоняется; $\chi_в^2 \approx 4{,}9$ (последние 2 интервала объединяются). **19.292.** H_0 принимается; $\chi_в^2 \approx 7{,}03$ (первые 2 и последние 3 интервала объединяются). **19.293.** H_0 принимается; $\chi_в^2 \approx 3{,}26$ (первые 2 и последние 2 интервала объединяются). **19.294.** H_0 принимается; $\chi_в^2 \approx 0{,}476$. **19.295.** H_0 отклоняется; $\chi_в^2 \approx 6{,}22$ (первые 2 и последние 2 интервала объединяются). **19.296.** H_0 принимается; $\chi_в^2 \approx 0{,}517$ (первые 3 и последние 3 интервала объединяются). **19.297.** Не зависят; $\chi_в^2 \approx 0{,}73$. **19.298.** Нет; $\chi_в^2 \approx 4{,}78$. **19.299.** Нет; $\chi_в^2 \approx 13{,}27$. **19.300.** Да; $\chi_в^2 \approx 44{,}2$. **19.301.** Да; $\chi_в^2 \approx 36$. **19.302.** Нет; $\chi_в^2 \approx 0{,}51$. **19.303.** Гипотеза принимается; $\chi_в^2 \approx 0{,}61$. **19.304.** Нет; $\chi_в^2 \approx 0{,}17$. **19.305.** Да; $\chi_в^2 \approx 27{,}14$. **19.306.** Да, $\chi_в^2 \approx 38{,}57$. **19.307.** $y = 0{,}5 + 0{,}5x$; диаграмма рассеяния приведена на рис. 83. **19.309.** $\tilde{\beta}_0 = \overline{y} - \beta_1\overline{x}$. **19.310.** $\tilde{\beta}_1 = \dfrac{\sum x_i y_i - \beta_0 \sum x_i}{\sum x_i^2}$.

19.312. Указание. В (7) и (8) положить $y_i = Y_i$, воспользоваться преобразованием $\sum (x_i - \overline{x})(Y_i - \overline{Y}) = \sum (x_i - \overline{x}) Y_i$ и вычислить $\mathrm{M}[\tilde{\beta}_0]$ и $\mathrm{M}[\tilde{\beta}_1]$. **19.313.** Указание. Использовать указание к задаче 19.312, показать, что ковариация $K_{\overline{Y}\tilde{\beta}_1} = 0$ и вычислить $\mathrm{D}[\tilde{\beta}_1]$ и $\mathrm{D}[\tilde{\beta}_0]$. **19.314.** Указание. Использовать то, что $K_{(X+Y)Z} = K_{XZ} + K_{YZ}$;

в силу (8) $K_{\tilde\beta_0\tilde\beta_1} = K_{\overline{Y}\tilde\beta_1} - \overline{x}K_{\tilde\beta_1\tilde\beta_2} = -\overline{x}\mathrm{D}[\tilde\beta_1]$. **19.315.** $\mathrm{M}[Y_0] = \beta_0 + \beta_1 x_0$. Указание. $\mathrm{D}[Y_0] = \mathrm{D}[\tilde\beta_0] + x_0^2\mathrm{D}[\tilde\beta_1] + 2x_0 K_{\tilde\beta_0\tilde\beta_1}$; подставить в это выражение значения $\mathrm{D}[\tilde\beta_0]$, $\mathrm{D}[\tilde\beta_1]$ и $K_{\tilde\beta_0\tilde\beta_1}$ и упростить его, используя

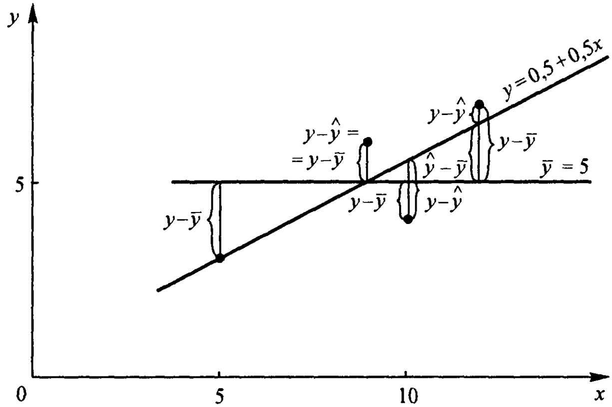


Рис. 84

соотношение $\sum x_i^2 = Q_x + n\overline{x}^2$. **19.317.** Нет. Указание. Модель необходимо преобразовать так: $y = \dfrac{1}{\beta_1}x - \dfrac{\beta_0}{\beta_1}$. **19.318.** $\tilde\beta_1 = \dfrac{Q_{xy}}{Q_x}$, $\tilde\beta_0 = \overline{y}$. **19.320.** $\widetilde\Theta = \dfrac{1}{n}\sum \theta_i = \overline\Theta$, $s^2 = \dfrac{1}{n-1}\sum(\theta_i - \overline\Theta)^2$. **19.321.** Нормальное распределение $N\left(\beta_0, \sqrt{\dfrac{\sigma^2\sum x_i^2}{nQ_x}}\right)$ и $N\left(\beta_1, \sqrt{\dfrac{\sigma^2}{Q_x}}\right)$. Указание. МНК-оценки являются линейными функциями случайных величин $Y_i \sim N(\beta_0 + \beta_1 x_i \sigma)$, $i = 1, 2, \ldots, n$. **19.322.** Показать, что $\dfrac{\tilde\beta_i - \beta_i}{\sqrt{\mathrm{D}[\tilde\beta_i]}} : \sqrt{\dfrac{s^2}{\sigma^2}} = T(n-2)$, $i = 0, 1$. **19.324.** Указание. Использовать то, что $\dfrac{(n-2)}{\sigma^2}s^2 = \chi^2(n-2)$. **19.325.** Указание. Применить метод максимального правдоподобия для вычисления оценок $\tilde\beta_1$ и $\tilde\beta_0$ по выборке независимых случайных величин $\varepsilon_i = y_i - \beta_0 - \beta_1 x_i$, $\varepsilon_i \sim N(0, \sigma)$, $i = 1, 2, \ldots, n$. **19.326.** Указание. Равенство $y_i - \overline{y} = (y_i - \overline{y}) + (y_i - \tilde y_i)$, $i = 1, 2, \ldots, n$, возвести в квадрат и просуммировать по i; $2\sum(\tilde y_i - \overline{y})(y_i - \tilde y_i) = 0$, так как $\tilde y_i - \overline{y} = \tilde\beta_1(x_i - \overline{x})$, $y_i - \tilde y_i = y_i - \overline{y} - \tilde\beta_1(x_i - \overline{x})$. **19.327.** Указание. Воспользоваться формулами $\tilde y_i - \overline{y} = \tilde\beta_1(x_i - \overline{x})$, $\tilde\beta_1 = Q_{xy}/Q_x$. **19.328.** Указание. Так как $\dfrac{\tilde\beta_1-\beta_1}{\sqrt{\mathrm{D}[\tilde\beta_1]}}=\dfrac{\tilde\beta_1-\beta_1}{\sqrt{\sigma^2/Q_x}}\sim N(0,\,1)$, то $\dfrac{(\tilde\beta_1-\beta_1)^2}{\sigma^2/Q_x}=\chi^2(1)$; использовать также (13). **19.329.** $y=12{,}37-0{,}44x$, $\beta_0\in(6{,}08;18{,}66)$, $\beta_1\in\in(-1{,}15;\,0{,}27)$, y_0: $\tilde y_0\pm 9{,}9\sqrt{0{,}09+5{,}11\cdot10^{-3}\,(x_0-7{,}82)^2}$, не согласуется, $r\approx-0{,}434$. **19.330.** $y=20{,}30-1{,}06x$, $\beta_0\in(18{,}99;\,21{,}60)$, $\beta_1\in(-1{,}21;-0{,}90)$, y_0: $\tilde y_0\pm1{,}79\sqrt{0{,}11+7{,}62\cdot10^{-3}\,(x_0-9{,}03)^2}$, согласуется, $r\approx-0{,}979$. **19.331.** $y=-1{,}49+2{,}03x$, $\beta_0\in(-5{,}78;\,2{,}80)$, $\beta_1\in(1{,}78;\,2{,}28)$, y_0: $\tilde y_0\pm3{,}65\sqrt{0{,}11+4{,}7\cdot10^{-3}(x_0-16{,}43)^2}$, согласуется, $r\approx0{,}986$. **19.332.** $y=-5{,}78+1{,}66x$, $\beta_0\in(-8{,}75;\,-2{,}81)$, $\beta_1\in(1{,}37;\,1{,}95)$, y_0 : $\tilde y_0\pm5{,}86\times\sqrt{0{,}06+2{,}14\cdot10^{-3}(x_0-9)^2}$, согласуется, $r\approx0{,}953$. **19.333.** $y=7{,}17+0{,}21x$; модель неадекватна,

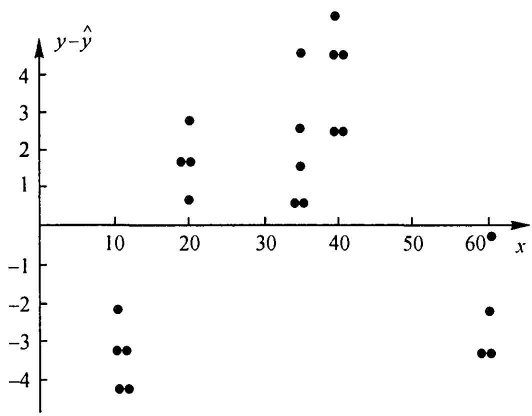


Рис. 85

график остатков приведен на рис. 85. **19.334.** $y=0{,}906+0{,}932x$, модель адекватна. **19.335.** $y=3{,}995-2{,}163x+0{,}268x^2$. **19.336.** $y=1{,}40-\\-1{,}23x-0{,}87x^2$. **19.337.** $y=0{,}14+0{,}088x+0{,}002x^2$. **19.338.** $y==-1{,}93-0{,}28x+1{,}54x^2$. **19.339.** $y=2+12/x$. **19.340.** $y=4{,}69+15{,}97/x$. **19.341.** $y=11{,}92+4{,}39/x$.

19.342. $A=\begin{pmatrix}1 & x_1 & x_1^2\\ 1 & x_2 & x_2^2\\ \dots & \dots & \dots\\ 1 & x_n & x_n^2\end{pmatrix}$. **19.343.** $\begin{pmatrix}1 & \sin\omega x_1 & \cos\omega x_1\\ 1 & \sin\omega x_2 & \cos\omega x_2\\ \dots & \dots & \dots\\ 1 & \sin\omega x_n & \cos\omega x_n\end{pmatrix}$.

19.344. $A=\begin{pmatrix}1 & e^{x_1}\\ 1 & e^{x_2}\\ \dots & \dots\\ 1 & e^{x_n}\end{pmatrix}$. **19.345.** $y=-1{,}48+2{,}45x+0{,}64x^2$, $s^2=6{,}77$,

$K=\begin{pmatrix}2{,}25 & 0 & -0{,}33\\ 0 & 0{,}24 & 0\\ -0{,}33 & 0 & 0{,}08\end{pmatrix}$. **19.346.** а) $y=-1{,}33+1{,}07x+1{,}12x^2$;

б) значима; в) $s^2 = 0{,}38$, $K = \begin{pmatrix} 0{,}128 & 0 & -0{,}018 \\ 0 & 0{,}014 & 0 \\ -0{,}018 & 0 & 0{,}005 \end{pmatrix}$; г) $\beta_0 \in (-2{,}095;\ -0{,}57)$, $\beta_1 \in (0{,}82;\ 1{,}32)$; $\beta_2 \in (1{,}01;\ 1{,}22)$, $\sigma^2 \in (0{,}161;\ 2{,}163)$.
19.347. а) $y = 2{,}49 + 1{,}8x + 0{,}86x^2$; б) не значима; в) $s^2 = 1{,}935$,

$K = \begin{pmatrix} 0{,}94 & 0 & -0{,}28 \\ 0 & 0{,}193 & 0 \\ -0{,}28 & 0 & 0{,}138 \end{pmatrix}$; г) $\beta_0 \in (-1{,}068;\ +6{,}56)$, $\beta_1 \in (-0{,}09;\ 3{,}69)$, $\beta_2 \in (-0{,}76;\ 2{,}45)$, $\sigma^2 \in (0{,}524;\ 76{,}48)$. **19.348** а) $y = -0{,}48 + 1{,}11x - 0{,}27x^2$; б) значима; в) $s^2 = 0{,}27$, $K = 10^{-2} \times \begin{pmatrix} 9{,}1 & 0 & -1{,}3 \\ 0 & 0{,}99 & 0 \\ -1{,}3 & 0 & 0{,}33 \end{pmatrix}$; г) $\beta_0 \in (-0{,}81;\ -0{,}14)$, $\beta_1 \in (1{,}00;\ 1{,}22)$, $\beta_2 \in (-0{,}34; -0{,}21)$, $\sigma^2 \in (0{,}12;\ 1{,}54)$. **19.349.** $y = 22{,}56 + 1{,}67x - 0{,}068x^2$; адекватна; нет. **19.350.** $y = -0{,}834 + 0{,}576x - 0{,}006x^2$; неадекватна; нет.

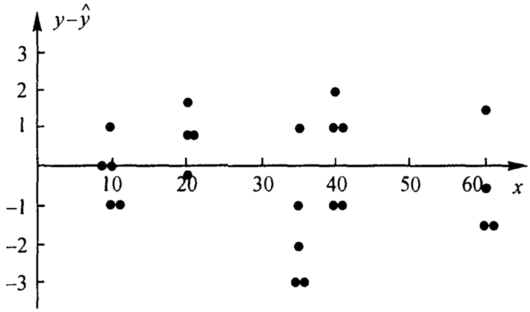


Рис. 86

19.351. $y = -2{,}26 + 0{,}93x - 0{,}01x^2$; адекватна; нет; график остатков приведен на рис. 86. **19.352.** $y = -0{,}89 + 2{,}21x - 0{,}19x^2$; адекватна; да.
19.353. Программы для решения задач 19.353–19.362.

```
      EXTERNAL F1
      DIMENSION COV (3, 3)
      DIMENSION A(16, 3), AT(3, 16), X(16), Y(16),
    * BET (3), ATB (3, 3) AB (3, 16)
      DIMENSION B(16, 3), ATB1 (3), ATB2 (3)
      N=16
      K=3
      DO 1 I=1, 16
    1 X(I) =1*.5
      READ 3,(Y(I), I=1, N)
    3 FORMAT (16F3.1)
      PRINT 1Ø, (Y(I), I=1, N)
   1Ø FORMAT (12F7.3)
      CALL FORM (N, K, B, A, X, F1)
      CALL MNK (A, A, N, K, ATB, AT, AB, Y, BET)
    6 FORMAT (5X, 'BETA=')
```

```
      PRINT 6
      PRINT 2, BET
    2 FORMAT (//5X, 3(2X, F7.3))
      CALL DISP (N, K, Y, A, BET, ATB, Q1, Q2, QY, EVD, FB)
      PRINT 78, EVD
   78 FORMAT (//10X, 'EVD=', E12.5)
      CALL MTRCOV (ATB, COV, EVD, K, K)
   19 FORMAT (//10X, 3E12.5)
      PRINT 19,((COV(I, J), J=1, K), I=1, K)
      STOP
      END
```

Здесь:

```
      FUNCTION F1 (X, K)
      F1=X**K
      RETURN
      END
```

$\tilde{\beta}_0 = 0{,}265, \quad \tilde{\beta}_1 = -0{,}173, \quad \tilde{\beta}_2 = 0{,}396, \quad \tilde{\sigma}^2 = 4{,}357 \cdot 10^{-2},$

$$\widetilde{K} = \begin{pmatrix} 0{,}31821 \cdot 10^{-1} & -0{,}15405 \cdot 10^{-1} & 0{,}15560 \cdot 10^{-2} \\ -0{,}15405 \cdot 10^{-1} & 0{,}93302 \cdot 10^{-2} & -0{,}10374 \cdot 10^{-2} \\ 0{,}15560 \cdot 10^{-2} & -0{,}10374 \cdot 10^{-2} & 0{,}12204 \cdot 10^{-3} \end{pmatrix}.$$

19.354. $\tilde{\beta}_0 = 1{,}128, \ \tilde{\beta}_1 = -0{,}869, \ \tilde{\beta}_2 = 1{,}019, \ \tilde{\sigma}^2 = 0{,}11663,$

$$\widetilde{K} = \begin{pmatrix} -0{,}7889 \cdot 10^{-1} & -0{,}45021 \cdot 10^{-1} & 0{,}53597 \cdot 10^{-2} \\ -0{,}45021 \cdot 10^{-1} & 0{,}32252 \cdot 10^{-1} & -0{,}42314 \cdot 10^{-2} \\ 0{,}53597 \cdot 10^{-2} & -0{,}42314 \cdot 10^{-2} & 0{,}58769 \cdot 10^{-3} \end{pmatrix}.$$

19.355. $\tilde{\beta}_0 = 16{,}887, \ \tilde{\beta}_1 = 9{,}791, \ \tilde{\beta}_2 = -6{,}904, \ \tilde{\beta}_2 = 0{,}13511 \cdot 10^{-3},$

$$\widetilde{K} = \begin{pmatrix} 0{,}74784 \cdot 10^{2} & -0{,}29155 \cdot 10^{3} & 0{,}23703 \cdot 10^{3} \\ -0{,}29155 \cdot 10^{3} & 0{,}14388 \cdot 10^{4} & -0{,}12929 \cdot 10^{4} \\ 0{,}23703 \cdot 10^{3} & -0{,}12929 \cdot 10^{4} & 0{,}12313 \cdot 10^{4} \end{pmatrix}.$$

19.356. $\tilde{\beta}_0 = 0{,}226, \ \tilde{\beta}_1 = -0{,}106, \ \tilde{\beta}_2 = 0{,}100, \ \tilde{\sigma}^2 = 0{,}14252 \cdot 10^{-1},$

$$\widetilde{K} = \begin{pmatrix} 0{,}83979 \cdot 10^{-2} & -0{,}34415 \cdot 10^{-2} & 0{,}29414 \cdot 10^{-3} \\ -0{,}34415 \cdot 10^{-2} & 0{,}17808 \cdot 10^{-2} & -0{,}16808 \cdot 10^{-3} \\ 0{,}29414 \cdot 10^{-3} & -0{,}16808 \cdot 10^{-3} & 0{,}16808 \cdot 10^{-4} \end{pmatrix}.$$

19.357. $\tilde{\beta}_0 = 2{,}209, \ \tilde{\beta}_1 = 0{,}350, \ \tilde{\sigma}^2 = 0{,}44583, \ \widetilde{K} = \begin{pmatrix} 0{,}49234 & -0{,}24866 \\ -0{,}24866 & 0{,}13685 \end{pmatrix}.$

19.358. $\tilde{\beta}_0 = 4{,}973, \ \tilde{\beta}_1 = 0{,}466, \ \tilde{\sigma}^2 = 0{,}14038 \cdot 10^{-2},$

$$\widetilde{K} = \begin{pmatrix} 0{,}16415 \cdot 10^{-3} & -0{,}10729 \cdot 10^{-3} \\ -0{,}10729 \cdot 10^{-3} & 0{,}18020 \cdot 10^{-3} \end{pmatrix}.$$

19.359. $\tilde\beta_0 = -0{,}863$, $\tilde\beta_1 = 0{,}913$, $\tilde\sigma^2 = 0{,}52277\cdot 10^{-2}$,

$$\widetilde{K} = \begin{pmatrix} 0{,}41818\cdot 10^{-2} & -0{,}18439\cdot 10^{-2} \\ -0{,}18439\cdot 10^{-2} & -0{,}90754\cdot 10^{-3} \end{pmatrix}.$$

19.360. $\tilde\beta_0 = 3{,}018$, $\tilde\beta_1 = 0{,}992$, $\tilde\sigma^2 = 0{,}84339\cdot 10^{-3}$,

$$\widetilde{K} = \begin{pmatrix} 0{,}19319\cdot 10^{-2} & -0{,}12299\cdot 10^{-2} \\ -0{,}12299\cdot 10^{-2} & 0{,}81017\cdot 10^{-3} \end{pmatrix}.$$

19.361. $\tilde\beta_0 = 2{,}217$, $\tilde\beta_1 = 0{,}836$, $\tilde\beta_2 = -0{,}450$, $\tilde\sigma^2 = 0{,}75299$,

$$\widetilde{K} = \begin{pmatrix} 0{,}63048\cdot 10^{-1} & 0{,}68163\cdot 10^{-3} & 0{,}62628\cdot 10^{-2} \\ 0{,}68163\cdot 10^{-3} & 0{,}11984 & 0{,}13818\cdot 10^{-2} \\ 0{,}62628\cdot 10^{-2} & 0{,}13818\cdot 10^{-2} & 0{,}13238 \end{pmatrix}.$$

19.362. $\tilde\beta_0 = 0{,}960$, $\tilde\beta_1 = -0{,}047$, $\tilde\beta_2 = 0{,}969$, $\tilde\sigma^2 = 0{,}13599\cdot 10^{-1}$,

$$\widetilde{K} = \begin{pmatrix} 0{,}10483\cdot 10^{-2} & -0{,}68179\cdot 10^{-4} & 0{,}29972\cdot 10^{-6} \\ -0{,}68179\cdot 10^{-4} & 0{,}20267\cdot 10^{-2} & 0{,}61506\cdot 10^{-6} \\ 0{,}29972\cdot 10^{-6} & 0{,}61506\cdot 10^{-6} & 0{,}21669\cdot 10^{-2} \end{pmatrix}.$$

19.363. $y_i = 23{,}3 + 2{,}59P_1(i)$. **19.364** $y_i = 0{,}455 + 0{,}081P_1(i)$. **19.365.** $y_i = 23{,}662 + 2{,}984P_1(i) + 0{,}317P_2(i) + 0{,}087P_3(i)$.

19.366. $\mathrm{D}[\tilde\beta_j] = \dfrac{\sigma^2}{\sum\limits_{i=1}^{n}\psi_j^2(x_i)}$, $j = 0, 1, \ldots, k-1$, где σ^2 — дисперсия ошибок наблюдений. **19.368** $\tilde\alpha_0 = \overline{y}$, $\tilde\alpha_j = \dfrac{2}{n}\sum\limits_{k=0}^{n-1} y_k\cos jx_k$, $\tilde\beta_j = \dfrac{2}{n}\sum\limits_{k=0}^{n-1} y_k\sin jx_k$, $\mathrm{D}[\tilde\alpha_0] = \dfrac{\sigma^2}{n}$, $\mathrm{D}[\tilde\alpha_j] = \mathrm{D}[\tilde\beta_j] = \dfrac{2}{n}\sigma^2$, $s^2 = \dfrac{Q_e}{n-2m-1}$,

где

$$Q_e = \sum_{k=0}^{n-1} y_k^2 - n\tilde\alpha_0^2 - \frac{n}{2}\sum_{j=1}^{m}(\tilde\alpha_j^2 + \tilde\beta_j^2).$$

Указание. $\sum\limits_{k=0}^{n-1}\cos^2 jx_k = \sum\limits_{k=0}^{n-1}\sin^2 jx_k = \dfrac{n}{2}$, для определения s^2 использовать формулу (26). **19.369.** $\tilde\alpha_0 = 3{,}156$, $\tilde\alpha_1 = 0{,}802$, $\tilde\beta_1 = 0{,}765$, $s^2 = 0{,}565$. **19.370.** $\tilde\alpha_0 = 0{,}044$, $\tilde\alpha_1 = 0{,}911$,

$\tilde{\beta}_1 = 1{,}85$, $s^2 = 5{,}457$. **19.371.** $z = \beta_0 - \beta_1 x$, где $z = \ln y$. **19.372.** $w = \beta_0 + \theta z$, где $w = 1/y$, $z = 1/x$, $\theta = 1/\beta_1$. **19.373.** Модель линейна. **19.374.** $y = \theta_0 + \beta_1 \sin x$, где $\theta_0 = \sin \beta_0$. **19.375.** $w = \theta_0 + \beta_1 \ln x$, где $w = \ln y$, $\theta_0 = \ln \beta_0$. **19.376.** $w = \beta_0 + \beta_1 x$, где $w = \dfrac{x}{y}$. **19.377.** $\tilde{p}_0 = 3053{,}365$, $\tilde{k}/T = 0{,}0015$. **19.378.** $\tilde{\beta} = 1{,}73$, $\tilde{\alpha} = -1{,}066$. **19.379.** $\tilde{a} = 4480{,}4$, $\tilde{b} = 3{,}304$. **19.380.** $\tilde{A} = 2$, $\tilde{\gamma} = 2{,}6$. **19.381.** $\tilde{\beta}_0 = 0{,}01$, $\tilde{\beta}_1 = 5{,}7 \cdot 10^{-3}$. **19.382.** а) $y = 1{,}8 + 1{,}07x_1 + 0{,}59x_2$, $\beta_1 \in (0{,}51;\, 1{,}64)$, $\beta_2 \in (0{,}18;\, 1{,}0)$; б) и в) гипотезы отклоняются; г) $R \approx 0{,}815$. **19.383.** а) $y = 2{,}75 - 1{,}42x_1 - 0{,}26x_2$, $\beta_1 \in (-4{,}07;\, 1{,}232)$, $\beta_2 \in (-1{,}534;\, 1{,}012)$; б) и в) гипотезы принимаются; г) $R \approx 0{,}739$. **19.384.** а) $y = -94{,}55 + 2{,}80x_1 + 1{,}07x_2$, $\beta_1 \in (2{,}03;\, 3{,}57)$, $\beta_2 \in (0{,}83;\, 1{,}31)$; б) и в) гипотезы отклоняются; г) $R \approx 0{,}99$. **19.385.** а) $y = 2{,}21 + 0{,}91x_1 - 0{,}059x_2$, $\beta_1 \in (0{,}6;\, 1{,}23)$, $\beta_2 \in (-0{,}48;\, 0{,}36)$; б) гипотеза отклоняется; в) гипотеза $H_0 : \beta_1 = 0$ отклоняется; гипотеза $H_0 : \beta_2 = 0$ принимается; г) $R \approx 0{,}933$.

19.386.
$$\begin{aligned}
&\beta_0 n + \beta_1 \sum x_{1i} + \beta_2 \sum x_{2i} = \sum y_i,\\
&\beta_0 \sum x_{1i} + \beta_1 \sum x_{1i}^2 + \beta_2 \sum x_{1i}x_{2i} = \sum y_i x_{1i},\\
&\beta_0 \sum x_{2i} + \beta_1 \sum x_{1i}x_{2i} + \beta_2 \sum x_{2i}^2 = \sum y_i x_{2i}.
\end{aligned}$$

19.387. $\tilde{\beta}_0 = -1{,}59$, $\tilde{\beta}_1 = 2{,}61$, $\tilde{\beta}_2 = 0{,}66$, $\tilde{\sigma}^2 = 10{,}429$,
$$\tilde{K} = \begin{pmatrix} 3{,}16 & 0 & -0{,}42 \\ 0 & 0{,}21 & 0 \\ -0{,}42 & 0 & 8{,}4 \cdot 10^{-2} \end{pmatrix}.$$

19.388. $\tilde{m} = \dfrac{\sum w_i m_i}{\sum w_i}$, $\mathrm{D}[\tilde{m}] = \dfrac{\sigma^2}{\sum w_i}$. **19.389.** $\tilde{\beta} = \dfrac{\sum y_i}{\sum x_i}$, $\mathrm{D}[\tilde{\beta}] = \dfrac{a}{\sum x_i}$, где a — коэффициент пропорциональности. **19.390.** $\tilde{\beta} = \dfrac{1}{n}\sum \dfrac{y_i}{x_i}$, $\mathrm{D}[\tilde{\beta}] = \dfrac{a}{n}$, где a — коэффициент пропорциональности. **19.391.** а) $\tilde{\beta} = 0{,}2605$, $\mathrm{D}[\tilde{\beta}] = 2{,}26 \cdot 10^{-3}$; б) $\tilde{\beta} = 0{,}2618$, $\mathrm{D}[\tilde{\beta}] = 1{,}82 \cdot 10^{-2}$; в) $\tilde{\beta} = 0{,}2914$, $\mathrm{D}[\tilde{\beta}] = 0{,}1$. **19.392.** ◁ Пусть $(x_i,\, y_i)$, $i = 1, 2, \ldots, n$, — попарно связанные значения двух выборок $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_n$. Предположим, что разности $d_i = x_i - y_i$ получены из нормально распределенной генеральной совокупности $N(m_d,\, \sigma_d)$. Проверяется гипотеза $H_0 : m_d = 0$ при альтернативной гипотезе $H_1 : m_d \neq 0$. Статистика Z критерия: $\dfrac{\bar{d} - m_d}{s_d}$, где $\bar{d}$ — выборочное среднее, а s_d^2 — выборочная дисперсия последовательности разностей $d_1, d_2, \ldots, d_n$. Z имеет (при условии, что верна гипотеза H_0) распределение Стьюдента с $(n-1)$ степенями свободы. ▷ **19.393.** Нет. Распределение знаков не будет биномиальным.

19.394. Да. **19.395.** а) и б) Нет. **19.396.** Нет. **19.397.** Предположение отклоняется. **19.398.** Предположение принимается. **19.399.** Нет. **19.400.** Нет. **19.401.** Нет. **19.402.** Да. **19.403.** Да. **19.404.** Нет. **19.405.** Предположение отклоняется. **19.406.** Да. **19.407.** Да. **19.408.** H_0 принимается. **19.409.** H_0 отклоняется. **19.410.** Да. **19.411.** H_0 принимается. **19.412.** Да. **19.413.** Нет. **19.414.** Да. **19.415.** Да. **19.416.** Да. **19.417.** Нет. **19.418.** $r_s \approx 0{,}455$. **19.419.** $r_s \approx 0{,}789$. **19.420.** $r_s \approx 0{,}944$. **19.421.** $-0{,}009$; $-0{,}22$, оба коэффициента незначимы. **19.422.** а) $r_s \approx 0{,}86$; б) $r_s \approx 0{,}889$, оба коэффициента значимы. **19.423.** а) $r_s \approx 0{,}706$; б) $r \approx 0{,}636$. **19.426.** Нет.

19.427. Указание. Так как $\overline{x'} = \dfrac{n+1}{2}$, а $\sum\limits_{k=1}^{n} k^2 = \dfrac{n(n+1)(2n+1)}{6}$, то

$$\sum_{i=1}^{n} (x'_i - \overline{x'})^2 = \sum_{k=1}^{n} \left[k - \frac{n+1}{2}\right]^2 = \sum_{k=1}^{n} k^2 - \frac{1}{4} n (n+1)^2.$$

Аналогично преобразуется сумма $\sum\limits_{i=1}^{n} (y'_i - \overline{y'})^2$. Сумма

$$\sum_{i=1}^{n} (x'_i - x')(y'_i - \overline{y}) = \sum_{i=1}^{n} x'_i y'_i - \frac{n}{4} (n+1)^2 =$$

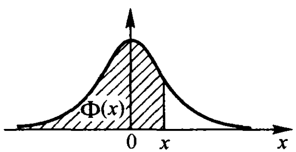
$$= -\frac{1}{2} \sum_{i=1}^{n} (x'_i - y'_i)^2 + \sum_{k=1}^{n} k^2 - \frac{n}{4} (n+1)^2 =$$

$$= \frac{n}{12} (n^2 - 1) - \frac{1}{2} \sum_{i=1}^{n} (x'_i - y'_i)^2.$$

ПРИЛОЖЕНИЯ

Таблица П1. **Функция распределения $\Phi(x)$ нормального закона $N(0, 1)$;**

$$\Phi(-x) = 1 - \Phi(x),\ \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$$



x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441

Таблица П1 (*продолжение*)

x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

Квантили u_p нормального распределения $N(0,\ 1)$:

p	0,90	0,95	0,975	0,99	0,995	0,999	0,9995
u_p	1,282	1,645	1,960	2,326	2,576	3,090	3,291

Таблица П2. **Значения функции плотности нормального распределения $N(0, 1)$ $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$**

x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	0,3970	0,3965	0,3961	0,3956	0,3951	0,3945	0,3939	0,3932	0,3925	0,3918
0,2	0,3910	0,3902	0,3894	0,3885	0,3876	0,3867	0,3857	0,3847	0,3836	0,3825
0,3	0,3814	0,3802	0,3790	0,3778	0,3765	0,3752	0,3739	0,3725	0,3712	0,3697
0,4	0,3683	0,3668	0,3653	0,3637	0,3621	0,3605	0,3589	0,3572	0,3555	0,3538
0,5	0,3521	0,3503	0,3485	0,3467	0,3448	0,3429	0,3410	0,3391	0,3372	0,3352
0,6	0,3332	0,3312	0,3292	0,3271	0,3251	0,3230	0,3209	0,3187	0,3166	0,3144
0,7	0,3123	0,3101	0,3079	0,3056	0,3034	0,3011	0,2989	0,2966	0,2943	0,2920
0,8	0,2897	0,2874	0,2850	0,2827	0,2803	0,2780	0,2756	0,2732	0,2709	0,2685
0,9	0,2661	0,2637	0,2613	0,2589	0,2565	0,2541	0,2516	0,2492	0,2468	0,2444
1,0	0,2420	0,2396	0,2371	0,2347	0,2323	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	0,2179	0,2155	0,2131	0,2107	0,2083	0,2059	0,2036	0,2012	0,1989	0,1965
1,2	0,1942	0,1919	0,1895	0,1872	0,1849	0,1826	0,1804	0,1781	0,1758	0,1736
1,3	0,1714	0,1691	0,1669	0,1647	0,1626	0,1604	0,1582	0,1561	0,1539	0,1518
1,4	0,1497	0,1476	0,1456	0,1435	0,1415	0,1394	0,1374	0,1354	0,1334	0,1315
1,5	0,1295	0,1276	0,1257	0,1238	0,1219	0,1200	0,1182	0,1163	0,1145	0,1127
1,6	0,1109	0,1092	0,1074	0,1057	0,1040	0,1023	0,1006	0,0989	0,0973	0,0957
1,7	0,0940	0,0925	0,0909	0,0893	0,0878	0,0863	0,0848	0,0833	0,0818	0,0804
1,8	0,0790	0,0775	0,0761	0,0748	0,0734	0,0721	0,0707	0,0694	0,0681	0,0669
1,9	0,0656	0,0644	0,0632	0,0620	0,0608	0,0596	0,0584	0,0573	0,0562	0,0551
x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09

Таблица П2 (*продолжение*)

x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
2,0	0,0540	0,0529	0,0519	0,0508	0,0498	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0,0440	0,0431	0,0422	0,0413	0,0404	0,0396	0,0387	0,0379	0,0371	0,0363
2,2	0,0355	0,0347	0,0339	0,0332	0,0325	0,0317	0,0310	0,0303	0,0297	0,0290
2,3	0,0283	0,0277	0,0270	0,0264	0,0258	0,0252	0,0246	0,0241	0,0235	0,0229
2,4	0,0224	0,0219	0,0213	0,0208	0,0203	0,0198	0,0193	0,0189	0,0184	0,0180
2,5	0,0175	0,0171	0,0167	0,0163	0,0158	0,0154	0,0151	0,0147	0,0143	0,0139
2,6	0,0136	0,0132	0,0129	0,0126	0,0122	0,0119	0,0116	0,0113	0,0110	0,0107
2,7	0,0104	0,0101	0,0099	0,0096	0,0093	0,0091	0,0088	0,0086	0,0084	0,0081
2,8	0,0079	0,0077	0,0075	0,0073	0,0071	0,0069	0,0067	0,0065	0,0063	0,0061
2,9	0,0060	0,0058	0,0056	0,0055	0,0053	0,0051	0,0050	0,0048	0,0047	0,0046
3,0	0,0044	0,0043	0,0042	0,0040	0,0039	0,0038	0,0037	0,0036	0,0035	0,0034
3,1	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026	0,0025	0,0025
3,2	0,0024	0,0023	0,0022	0,0022	0,0021	0,0020	0,0020	0,0019	0,0018	0,0018
3,3	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014	0,0013	0,0013
3,4	0,0012	0,0012	0,0012	0,0011	0,0011	0,0010	0,0010	0,0010	0,0009	0,0009
3,5	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007	0,0007	0,0007	0,0006
3,6	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005	0,0005	0,0005	0,0005	0,0004
3,7	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003	0,0003	0,0003	0,0003
3,8	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002	0,0002	0,0002	0,0002	0,0002
3,9	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0001	0,0001
4,0	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09

Таблица П3. **Распределение Пуассона** $\mathbf{P}[X=k]=\frac{\lambda^k}{k!}e^{-\lambda}$

k	λ = 0,1	0,2	0,3	0,4	0,5
0	0,90484	0,81873	0,74082	0,67032	0,60653
1	0,09048	0,16375	0,22223	0,26813	0,30327
2	0,00452	0,01638	0,03334	0,05363	0,07582
3	0,00015	0,00109	0,00333	0,00715	0,01204
4		0,00006	0,00025	0,00072	0,00158
5			0,00002	0,00006	0,00016
6					0,00001

k	λ = 0,6	0,7	0,8	0,9	
0	0,54881	0,49659	0,44933	0,40657	
1	0,32929	0,34761	0,25946	0,36591	
2	0,09879	0,12166	0,14379	0,16466	
3	0,01976	0,02839	0,03834	0,04940	
4	0,00296	0,00497	0,00767	0,01112	
5	0,00036	0,00070	0,00123	0,00200	
6	0,00004	0,00008	0,00016	0,00030	
7		0,00001	0,00002	0,00004	

k	λ = 1,0	2,0	3,0	4,0	5,0
0	0,36788	0,13534	0,04979	0,01832	0,00674
1	0,36788	0,27067	0,14936	0,07326	0,03369
2	0,18394	0,27067	0,22404	0,14653	0,08422
3	0,06131	0,18045	0,22404	0,19537	0,14037
4	0,01533	0,09022	0,16803	0,19537	0,17547
5	0,00307	0,03609	0,10082	0,15629	0,17547
6	0,00051	0,01203	0,05041	0,10419	0,14622
7	0,00007	0,00344	0,02160	0,05954	0,10445
8	0,00001	0,00086	0,00810	0,02977	0,06528
9		0,00019	0,00270	0,01323	0,03627
10		0,00004	0,00081	0,00529	0,01813
11		0,00001	0,00022	0,00193	0,00824
12			0,00006	0,00064	0,00343
13			0,00001	0,00020	0,00132
14				0,00006	0,00047
15				0,00002	0,00016
16					0,00005
17					0,00001

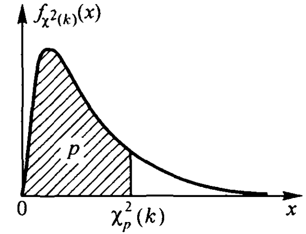
Таблица П4. **Суммарные вероятности для распределения Пуассона:**

$$\mathbf{P}[X \geqslant x] = \sum_{k=x}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda}$$

k	λ						
	0,1	0,2	0,3	0,4	0,5	0,6	0,7
0	1,00000	1,00000	1,00000	1,00000	1,00000	1,00000	1,00000
1	0,09516	0,18127	0,25918	0,32968	0,39347	0,45119	0,50341
2	00468	01752	03694	06155	09020	12190	15580
3	00016	00115	00360	00793	01439	02312	03414
4	0	00006	00027	00078	00175	00396	00575
5		0	00002	00006	00017	00039	00079
6					00001	00004	00009
7							00001

k	λ						
	0,8	0,9	1,0	2,0	4,0	6,0	8,0
0	1,00000	1,00000	1,00000	1,00000	1,00000	1,00000	1,00000
1	0,55067	0,59343	0,63212	0,86466	0,98168	0,99752	0,99966
2	19121	22752	26424	59399	90842	98265	99698
3	04742	06286	08030	32332	76190	93803	98625
4	00908	01346	01899	14288	56653	84880	95762
5	00141	00234	00366	05265	37116	71494	90037
6	00018	00034	00059	01656	21487	55432	80876
7	00002	00004	00008	00453	11067	39370	68663
8		00001	00002	00110	05113	25602	54704
9				00024	02136	15276	40745
10				00005	00823	08392	28338
11				00001	00284	04262	18411
12					00092	02009	11192
13					00027	00883	06380
14					00008	00363	03418
15					00002	00140	01726
16					00001	00051	00823
17						00017	00372
18						00006	00159
19						00002	00065
20						00001	00025
21							00009
22							00003
23							00001
24							

Таблица П5. **Квантили χ-квадрат распределения $\chi_p^2(k)$**

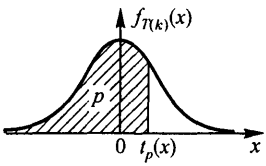


k	p														
	0.005	0.10	0.025	0.05	0.10	0.20	0.30	0.70	0.80	0.90	0.95	0.975	0.990	0.995	0.999
1	0.0^4393	0.0^3157	0.0^3982	0.0^2393	0.0158	0.0642	0.148	0.07	1.64	2.71	3.84	5.02	6.63	7.88	10.8
2	0.0108	0.201	0.0506	0.103	0.211	0.446	0.713	2.41	3.22	4.61	5.99	7.38	9.21	10.6	13.8
3	0.0717	0.115	0.216	0.352	0.584	1.00	1.42	3.67	4.64	6.25	7.81	9.35	11.3	12.8	16.3
4	0.207	0.297	0.484	0.711	1.06	1.65	2.19	4.88	5.99	7.78	9.49	11.1	13.3	14.9	18.5
5	0.412	0.554	0.831	1.15	1.61	2.34	3.00	6.06	7.29	9.24	11.1	12.8	15.1	16.7	20.5
6	0.676	0.872	1.24	1.64	2.20	3.07	3.83	7.23	8.56	10.6	12.6	14.4	16.8	18.5	22.5
7	0.989	1.24	1.69	2.17	2.83	3.82	4.67	8.38	9.80	12.0	14.1	16.0	18.5	20.3	24.3
8	1.34	1.65	2.18	2.73	3.49	4.59	5.53	9.52	11.0	13.4	15.5	17.5	20.1	22.0	26.1
9	1.73	2.09	2.70	3.33	4.17	5.38	6.39	10.7	12.2	14.7	16.9	19.0	21.7	23.6	27.9
10	2.16	2.56	3.25	3.94	4.87	6.18	7.27	11.8	13.4	16.0	18.3	20.5	23.2	25.2	29.6
11	2.60	3.05	3.82	4.57	5.58	6.99	8.15	12.9	14.6	17.3	19.7	21.9	24.7	26.8	31.3
12	3.07	3.57	4.40	5.23	6.30	7.81	9.03	14.0	15.8	18.5	21.0	23.3	26.2	28.3	32.9
13	3.57	4.11	5.01	5.89	7.04	8.63	9.93	15.1	17.0	19.08	22.4	24.7	27.7	29.8	34.5
14	4.07	4.66	5.63	6.57	7.79	9.47	10.8	16.2	18.2	21.1	23.7	26.1	29.1	31.3	36.1
15	4.60	5.23	6.26	7.26	8.55	10.3	11.7	17.3	19.3	22.3	25.0	27.5	30.6	32.8	37.7

Таблица П5 (*продолжение*)

k	p														
	0.005	0.10	0.025	0.05	0.10	0.20	0.30	0.70	0.80	0.90	0.95	0.975	0.990	0.995	0.999
16	5.14	5.81	6.91	7.96	9.31	11.2	12.6	18.4	20.5	23.5	26.3	28.8	32.0	34.3	39.3
17	5.70	6.41	7.56	8.67	10.1	12.0	13.5	19.5	21.6	24.8	27.6	30.2	33.4	35.7	40.8
18	6.26	7.01	8.23	9.39	10.9	12.9	14.4	20.6	22.8	26.0	28.9	31.5	34.8	37.2	42.3
19	6.84	7.63	8.91	10.1	11.7	13.7	15.4	21.7	23.9	27.2	30.1	32.9	36.2	38.6	43.8
20	7.43	8.26	9.59	10.9	12.4	14.6	16.3	22.8	25.0	28.4	31.4	34.2	37.6	40.0	45.3
21	8.03	8.90	10.3	11.6	13.2	15.4	17.2	23.9	26.9	29.6	32.7	35.5	38.9	41.4	46.8
22	8.64	9.54	11.0	12.3	14.0	16.3	18.1	24.9	27.3	30.8	33.9	36.8	40.3	42.8	48.3
23	9.26	10.2	11.7	13.1	14.8	17.2	19.0	26.0	28.4	32.0	35.2	38.1	41.6	44.2	49.7
24	9.89	10.9	12.4	13.8	15.7	18.1	19.9	27.1	29.6	33.2	36.4	39.4	43.0	45.6	51.2
25	10.5	11.5	13.1	14.6	16.5	18.9	20.9	28.2	30.7	34.4	37.7	40.6	44.3	46.9	52.6
26	11.2	12.2	13.8	15.4	17.3	19.8	21.8	29.2	31.8	35.6	38.9	41.9	45.6	48.3	54.1
27	11.8	12.9	14.6	16.2	18.1	20.7	22.7	30.3	32.9	36.7	40.1	43.2	47.0	49.6	55.5
28	12.5	13.6	15.3	16.9	18.9	21.6	23.6	31.4	34.0	37.9	41.3	44.5	48.3	51.0	56.9
29	13.1	14.3	16.0	17.7	19.8	22.5	24.6	32.5	35.1	39.1	42.6	45.7	49.6	52.3	58.3
30	13.8	15.0	16.8	18.5	20.6	23.4	25.5	33.5	36.3	40.3	43.8	47.0	50.9	53.7	59.7
35	17.2	18.5	20.6	22.5	24.8	27.8	30.2	38.9	41.8	46.1	49.8	53.2	57.3	60.3	66.6
40	20.7	22.2	24.4	26.5	29.1	32.3	34.9	44.2	47.3	51.8	55.8	59.3	63.7	66.8	73.4
45	24.3	25.9	28.4	30.6	33.4	36.9	39.6	49.5	52.7	57.5	61.7	65.4	70.0	73.2	80.1
50	28.0	29.7	32.4	34.8	37.7	41.4	44.3	54.7	58.2	63.2	67.5	71.4	76.2	79.5	86.7
75	47.2	49.5	52.9	56.1	59.8	64.5	68.1	80.9	85.1	91.1	96.2	100.8	106.4	110.3	118.6
100	67.3	70.1	74.2	77.9	82.4	87.9	92.1	106.9	111.7	118.5	124.3	129.6	135.6	140.2	149.4

Таблица П6. **Квантили распределения Стьюдента $t_p(k)$**



k	p						
	0,750	0,900	0,950	0,975	0,990	0,995	0,999
1	1,000	3,078	6,314	12,796	31.821	63,657	318
2	0,816	1,886	2,920	4,303	6,965	9,925	22,3
3	0,765	1,638	2,353	3,182	4,541	5,841	10,2
4	0,741	1,533	2,132	2,776	3,747	4,604	7,173
5	0,727	1,476	2,015	2,571	3,365	4,032	5,893
6	0,718	1,440	1,943	2,447	3,143	3,707	5,208
7	0,711	1,415	1,895	2,365	2,998	3,499	4,785
8	0,706	1,397	1,860	2,306	2,896	3,355	4,501
9	0,703	1,373	1,833	2,262	2,821	3,250	4,297
10	0,700	1,372	1,812	2,228	2,764	3,169	4,144
11	0,697	1,363	1,796	2,201	2,718	3,106	4,025
12	0,695	1,356	1,782	2,179	2,681	3,055	3,930
13	0,694	1,350	1,771	2,160	2,650	3,012	3,852
14	0,692	1,345	1,761	2,145	2,624	2,977	3,787
15	0,691	1,341	1,753	2,131	2,602	2,947	3,733
16	0,690	1,337	1,746	2,120	2,583	2,921	3,686
17	0,689	1,333	1,740	2,110	2,567	2,898	3,646
18	0,688	1,330	1,734	2,101	2,552	2,878	3,610
19	0,688	1,328	1,729	2,093	2,539	2,861	3,579
20	0,687	1,325	1,725	2,086	2,528	2,845	3,552
21	0,686	1,323	1,721	2,080	2,518	2,831	3,527
22	0,686	1,321	1,717	2,074	2,508	2,819	3,505
23	0,685	1,319	1,714	2,069	2,500	2,807	3,485
24	0,685	1,318	1,711	2,064	2,492	2,797	3,467
25	0,684	1,316	1,708	2,060	2,485	2,787	3,450
26	0,684	1,315	1,706	2,056	2,479	2,779	3,435
27	0,684	1,314	1,703	2,052	2,473	2,771	3,421
28	0,683	1,313	1,701	2,048	2,467	2,763	3,408
29	0,683	1,311	1,699	2,045	2,462	2,756	3,398
30	0,683	1,310	1,697	2,042	2,457	2,750	3,385
40	0,681	1,303	1,684	2,021	2,423	2,704	3,307
60	0,679	1,296	1,671	2,000	2,390	2,660	3,232
120	0,677	1,289	1,658	1,980	2,358	2,617	3,160
∞	0,674	1,282	1,645	1,960	2,326	2,576	3,090

Таблица П7. **Квантили распределения Фишера $F_p(k_1, k_2)$**

k_2	k_1																	
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
	$p = 0,9$																	
1	39.86	49,50	53.59	55.83	57,24	58.20	58.91	59.44	59.86	60,19	60.71	61.22	61.74	62.00	62,26	62.53	62.79	63.06
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9,37	9.38	9,39	9,41	9,42	9.44	9.45	9.46	9.47	9,47	9.48
3	5,54	5,46	5.39	5,34	5.31	5.28	5,27	5,25	5.24	5.23	5,22	5.20	5.18	5.18	5,17	5.16	5.15	5.14
4	4.54	4.32	4.19	4,11	4,05	4,01	3.98	3,95	3,94	3,92	3.90	3,87	3.84	3,83	3.82	3.80	3.79	3.78
5	4,06	3.78	3,62	3,52	3,45	3,40	3.37	3.34	3,32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12
6	3.78	3,46	3.29	3,18	3.11	3.05	3,01	2,98	2.96	2,94	2.90	2,87	2.84	2.82	2.80	2.78	2.76	2.74
7	3,59	3.26	3.07	2.96	2,88	2.83	2,78	2,75	2.72	2.70	2.67	2,63	2.59	2.58	2,56	2.54	2.51	2.49
8	3,46	3,11	2,92	2,81	2.73	2,67	2,62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2,38	2.36	2.34	2.32
9	3,36	3,01	2.81	2,69	2,61	2,55	2,51	2.47	2.44	2.42	2,38	2,34	2.30	2.28	2.25	2.23	2.21	2.18
10	3,29	2,92	2,73	2,61	2.52	2,46	2.41	2,38	2,35	2,32	2.28	2.24	2.20	2,18	2,16	2.13	2.11	2.08
11	3.23	2.86	2,66	2,54	2.45	2,39	2,34	2,30	2.27	2,25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00
12	3.18	2,81	2,61	2.48	2,39	2,33	2,28	2,24	2,21	2,19	2,15	2.10	2.06	2,04	2.01	1.99	1.96	1.93
13	3,14	2,76	2,56	2,43	2,35	2.28	2.23	2,20	2.16	2,14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1,88
14	3.10	2,73	2.52	2.39	2,31	2.24	2.19	2,15	2.12	2.10	2,05	2.01	1,96	1.94	1.91	1.89	1.86	1.83
15	3.07	2,70	2,49	2,36	2,27	2.21	2,16	2.12	2.09	2,06	2.02	1.97	1.92	1,90	1.87	1.85	1.82	1.79
16	3,05	2,67	2,46	2,33	2,24	2,18	2.13	2,09	2,06	2.03	1.99	1,94	1,89	1.87	1,84	1.81	1.78	1.75
17	3,03	2,64	2.44	2,31	2,22	2,15	2.10	2.06	2.03	2,00	1.96	1.91	1.86	1,84	1.81	1,78	1.75	1,72
18	3,01	2,62	2,42	2,29	2,20	2,13	2,08	2,04	2,00	1,98	1.93	1,89	1.84	1.81	1,78	1.75	1,72	1.69
19	2,99	2,61	2.40	2.27	2,18	2,11	2,06	2,02	1.98	1.96	1.91	1.86	1,81	1,79	1,76	1.73	1.70	1.67
20	2,97	2,59	2.38	2,25	2.16	2,09	2.04	2.00	1.96	1,94	1.89	1,84	1,79	1.77	1,74	1.71	1,68	1.64
21	2.96	2,57	2,36	2.23	2.14	2,08	2.02	1,98	1.95	1,92	1,87	1.83	1,78	1.75	1,72	1.69	1.66	1.62
22	2.95	2,56	2,35	2,22	2.13	2.06	2.01	1,97	1.93	1.90	1.86	1.81	1.76	1,73	1.70	1.67	1.64	1.60
23	2,94	2.55	2.34	2.21	2,11	2.05	1.99	1.95	1.92	1,89	1.84	1.81	1.74	1.72	1.69	1.66	1.62	1.59
24	2.93	2.54	2.33	2,19	2,10	2,04	1,98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1,67	1.64	1,61	1,57
25	2.92	2,53	2.32	2,18	2,09	2.02	1,97	1.93	1,89	1,87	1.82	1.77	1.72	1.69	1,66	1.63	1.59	1,56
26	2,91	2,52	2.31	2,17	2.08	2,01	1.96	1,92	1.88	1,86	1.81	1.76	1.71	1,68	1,65	1.61	1,58	1.54
27	2,90	2,51	2,30	2,17	2.07	2.00	1,95	1.91	1.87	1.85	1,80	1.75	1,70	1.67	1.64	1.60	1.57	1.53
28	2.89	2,50	2,29	2,16	2,06	2,00	1,94	1.90	1.87	1.84	1,79	1.74	1.69	1.66	1,63	1.59	1,56	1.52
29	2,89	2.50	2,28	2,15	2.06	1.99	1.93	1.89	1,86	1.83	1,78	1.73	1.68	1.65	1.62	1.58	1.55	1,51
30	2,88	2,49	2.28	2.14	2,05	1.98	1,93	1.88	1.85	1,82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50
40	2,84	2,44	2,23	2,09	2.00	1,93	1,87	1,83	1,79	1,76	1,71	1,66	1.61	1,57	1,54	1.51	1,47	1,42
60	2,79	2,39	2,18	2.04	1,95	1,87	1,82	1,77	1,74	1,71	1.66	1,60	1,54	1,51	1,48	1.44	1,40	1.35
120	2.75	2,35	2,13	1.99	1.90	1,82	1.77	1,72	1.68	1.65	1.60	1,55	1,48	1.45	1.41	1.37	1.32	1.36
∞	2.71	2.30	2,08	1.94	1.85	1.77	1,72	1.67	1.63	1,60	1.55	1,49	1.42	1.38	1,34	1.30	1.24	1.17

Таблица П7 (*продолжение*)

k_2	k_1																	
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
	$p = 0,95$																	
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19,46	19.47	19.48	19.49
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55
4	7,71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5,69	5.66
5	6.61	5,79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40
6	5,99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27
8	5.32	4.46	4.07	3.84	3,69	3,58	3,50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3,04	3.01	2.97
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75
10	4.96	4,10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2,70	2.66	2.62	2.58
11	4.84	3.98	3.59	3.36	3.20	3.09	3,01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45
12	4.75	3.89	3.49	3.26	3.11	3,00	2.91	2.85	2,80	2.75	2.69	2.62	2.54	2.51	2,47	2.43	2.38	2.34
13	4.67	3.81	3.41	3.18	3.03	2,92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2,42	2.38	2.34	2.30	2.25
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2,46	2.39	2.35	2.31	2.27	2.22	2.18
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2,29	2.25	2.20	2.16	2.11
16	4.49	3.63	3.24	3.01	2.85	2,74	2,65	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06
17	4.45	3.59	3.20	2.96	2.81	2.70	2,61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01
18	4.41	3.55	3.16	2.93	2,77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93
20	4.35	3.49	3,10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90
21	4.32	3,47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	3.18	2,10	2.05	2.01	1.96	1.92	1.87
22	4.30	3,44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84
23	4.28	3,42	3,03	2.80	2,64	2.53	2.44	2.37	2.32	2,27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1,94	1.89	1.84	1.79
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77
26	4,23	3.37	2.98	2.74	2.59	2.47	2,39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73
28	4.20	3.34	2,95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70
30	4.17	3.32	2.92	2.69	2.53	2,42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47
120	3.92	3.07	2.68	2,45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35
∞	3.84	3.00	2,60	2.37	2.21	2.10	3.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22

Таблица П7 (*продолжение*)

k_2	k_1 = 1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
	$p = 0,975$																	
1	647.8	799.5	864.6	899.6	921.8	917.1	948.2	956.7	963,3	968.6	976.7	984.9	993.1	997.2	1001	1006	1010	1014
2	38.51	39.00	39.17	39.25	39.30	39,33	39.36	39,37	39.39	39.40	39,41	39.43	39.45	39.46	39.46	39.47	39.48	39.49
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31
5	10,01	8.43	7.76	7,39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6,43	6.33	6.28	6.23	6.18	6.12	6.07
6	8.81	7.26	6.60	6.23	5,99	5.82	5.70	5,60	5.52	5.46	4.37	5.27	5.17	5,12	5.07	5.01	4.96	4.90
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20
8	7,57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73
9	7,21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3,96	3.87	3,77	3,67	3.61	3.56	3.51	3.45	3.39
10	6.94	5.46	4.83	4,47	4.24	4.07	3.95	3,85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3,20	3.14
11	6,72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2,85	2.79
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2,78	2.72	2.66
14	6.30	4,86	4.24	3.89	3,66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55
15	6.20	4.77	4,15	3.80	3.58	3,41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.70	2.64	2.59	2,52	2.46
16	6.12	4.69	4.08	3.73	3,50	3.34	3.22	3.12	3.05	2,99	2,89	2,79	2.68	2.63	2.57	2.51	2.45	2.38
17	6,04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.56	2.50	2.44	2,38	2.32
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.20
20	5.87	4.46	3,86	3.51	3.29	3,13	3.01	2.91	2.84	2.77	2.68	2.57	2,46	2.41	2.35	2.29	2.22	2,16
21	5,83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2,73	2.64	2.53	2.42	2.37	2.31	2.25	2.18	2.11
22	5,79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08
23	5.75	4.35	3.75	3.41	3.18	3,02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2,30	2.24	2.18	2.11	2.04
24	5,72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01
25	5.69	4.29	3.69	3,35	3,13	2,97	2.85	2,75	2.68	2.61	2.51	2.41	2.30	2.24	2.18	2.12	2.05	1.98
26	5.66	4.27	3.67	3.33	3.10	2,94	2.82	2.73	2.65	2.59	2.49	2.39	2.28	2.22	2.16	2.09	2.03	1.95
27	5.63	4.24	3.65	3.31	3.08	2,92	2.80	2.71	2.63	2.57	2.47	2.36	2.25	2.19	2.13	2.07	2.00	1.93
28	5,61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91
29	5,59	4.20	3,61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.96	1.89
30	5.57	4.18	3,59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58
120	5.15	3.80	3.23	2,89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43
∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1,71	1.64	1.57	1.48	1.39	1.27

Таблица П7 (*продолжение*)

k_2	k_1																	
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
	$p = 0,99$																	
1	4052	4999.5	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339
2	98.50	99.00	99.17	99,25	99,30	99,33	99.36	99.37	99.39	89.40	99.42	99.43	99.45	99.46	99.47	99.47	99,48	99.49
3	34.12	30.82	29,46	28.71	28.24	27.91	27,67	27.49	27.35	27,23	27.05	26.87	26,69	26,60	26.50	26.41	26.32	26.22
4	21,20	18.00	16.69	15.98	15,52	15,21	14.98	14.80	14,66	14.55	14,37	14.20	14,02	13.93	13.84	13.75	13.65	13.56
5	16,26	13.27	12,06	11.39	10,97	10,67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11
6	13,75	10,92	9,78	9,15	8.75	8.47	8,26	8.10	7.98	7.87	7.72	7,56	7.40	7,31	7.23	7.14	7.06	6.97
7	12.25	9.55	8,45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6,31	6,16	6.07	5.99	5.91	5,82	5.74
8	11.26	8,65	7.59	7,01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5,36	5.28	5.20	5.12	5,03	4.95
9	10.56	8.02	6,99	6.42	6.06	5.80	5,61	5.47	5.35	5.26	5.11	4.96	4.81	4,73	4.65	4.57	4.48	4.40
10	10.04	7,56	6.55	5.99	5.64	5,39	5.20	5.06	4,94	4.85	4.71	4.56	4,41	4.33	4.25	4.17	4,08	4,00
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4,63	4.54	4.40	4,25	4.10	4.02	3,94	3.86	3,78	2.69
12	9.33	6.93	5,95	5.41	5.06	4.82	4,64	4,50	4.39	4.30	4.16	4.01	3.86	3,78	3,70	3.62	3.54	3.45
13	9,07	6.70	5,74	5.21	4.86	4.62	4.44	4.30	4,19	4.10	3.96	3.82	3.66	3,59	3.51	3,43	3.34	3.25
14	8.86	6.51	5.56	5,04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3,66	3.51	3.43	3.35	3,27	3.18	3.09
15	8.68	6.36	5.42	4,89	4.56	4.32	4.14	4.00	3,89	3,80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96
16	8.53	6.23	5.29	4.77	4.44	4,20	4.03	3.89	3.78	3.69	3.55	3,41	3.26	3.18	3,10	3.02	2.93	2.84
17	8.40	6,11	5.18	4.67	4.34	4,10	3.93	3.79	3.68	3,59	3.46	3.31	3,16	3.08	3.00	2.92	2.83	2.75
18	8,29	6,01	5,09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3,08	3,00	2.92	2,84	2.75	2.66
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2,76	2.67	2.58
20	8,10	5.85	4.94	4,43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2,55	2.46
22	7.95	5,72	4.82	4.31	3,99	3.76	3.59	3.45	3.35	3,26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2,40
23	7.88	5.66	4.76	4.26	3.94	3,71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2,70	2.62	2.54	2.45	2.15
24	7.82	5,61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31
25	7.77	5.57	4.68	4,18	3.85	3.63	3.46	3.32	3.22	3,13	2.99	2.85	2.70	2.62	2.54	2,45	2.36	2.27
26	7,72	5.53	4,64	4.14	3.82	3,59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23
27	7.68	5.49	4,60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2,78	2,63	2.55	2.47	2.38	2.29	2.20
28	7.64	5.45	4.57	4.07	3,75	3,53	3,36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2,35	2.26	2.17
29	7,60	5.42	4.54	4.04	3.73	3,50	3.33	3.20	3.09	3.00	2,87	2.73	2.57	2.49	2.41	2.33	2,23	2.14
30	7.56	5.39	4.51	4,02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2,70	2,55	2.47	2.39	2.30	2.21	2.11
40	7,31	5,18	4.31	3,83	3.51	3,29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92
60	7.08	4.98	4.13	3,65	3.34	3.12	2.95	2,82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73
120	6,85	4,79	3.95	3.48	3.17	2.96	2.79	2.66	2,56	2,47	2.34	2,19	2.03	1.95	1.86	1,76	1.66	1.53
∞	6.63	4.61	3.78	3.32	3.02	2,80	2.64	2.51	2.41	2.32	2,18	2.04	1.88	1.79	1.70	1.59	1.47	1.32

Таблица П7 (*продолжение*)

k_2	k_1 = 1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
	$p = 0,995$																	
1	16211	20000	21615	22500	23056	23437	23715	23925	24091	24224	24426	24630	24836	24940	25044	25148	25253	25359
2	198,5	199,0	199.2	199.2	199,3	199,3	199,4	199,4	199,4	199.40	199.4	199.4	199.4	199.5	199,5	199,5	199,5	199.5
3	55,55	49,80	47.47	46.19	45.39	44.84	44.43	44,13	43.88	43,69	43.39	43.08	42.78	42.62	42,47	42.31	42.15	41.99
4	31.33	26.28	24.26	23.15	22.46	21.97	21.62	21,35	21,14	20.97	20,70	20.44	20,17	20.03	19.89	19.75	19.61	19.47
5	22.78	18,31	16.53	15.56	14.94	14,51	14.20	13.96	13.77	13.62	13.38	13.15	12.90	12.78	12.66	12,53	12.40	12.27
6	18.64	14.54	12.92	12.03	11.46	11,07	10,79	10,57	10.39	10.25	10.03	9.81	9,59	9.47	9.36	9,24	9.12	9.00
7	16.24	12,40	10.88	10.05	9.52	9.16	8,89	8.68	8,51	8,38	8.18	7.97	7.75	7.65	7.53	7.42	7.31	7.19
8	14,69	11.04	9,60	8.81	8.30	7.95	7,69	7.50	7.34	7.21	7.01	6.81	6,61	6.50	6.40	6.29	6.18	6.06
9	12,61	10.11	8,72	7.96	7.47	7,13	6.88	6.69	6.54	6.42	6.23	6.03	5.83	5,73	5.62	5.52	5,41	5.30
10	12,83	9.43	8.08	7.34	6,87	6.54	6.30	6.12	5.97	5,85	5.66	5,47	5.27	5.17	5.07	4.97	4.86	4.73
11	12,23	8.91	7.60	6,88	6.42	6.10	5.86	5,68	5,54	5.42	5.24	5.05	4.86	4.76	4.65	4,55	4.44	4.34
12	11.75	8,51	7.23	6.52	6.07	5.76	5.52	5.35	5.20	5,09	4.91	4.72	4.53	4.43	4.33	4.23	4.12	4.01
13	11,37	8,19	6.93	6.23	5.79	5.48	5.25	5,08	4.94	4,82	4,64	4,46	4,27	4.17	4.07	3.97	3.87	3.26
14	11.06	7.92	6.68	6.00	5.56	5.26	5.03	4.86	4.72	4,60	4.43	4.25	4,06	3.96	3.86	3.76	3,66	3.55
15	10.80	7,70	6.48	5.80	5.37	5.07	4.85	4.67	4.54	4.42	4,25	4.07	3.88	3.79	3.69	3.58	3.48	3.37
16	10,58	7,51	6.30	5.64	5.21	4.91	4.69	4.52	4.38	4.27	4.10	3.92	3.73	3.64	3.54	3,44	3.33	3.22
17	10,38	7.35	6.16	5.50	5.07	4.78	4.56	4.39	4,25	4.14	3.97	3.79	3.61	3.51	3,41	3.31	3.21	3.10
18	10,22	7.21	6.03	5,37	4.96	4.66	4,44	4.28	4,14	4.03	3.86	3.68	3.50	3.40	3.30	3.20	3.10	2.99
19	10,07	7.09	5.92	5.27	4.85	4.56	4,34	4.18	4,04	3.93	3,76	3.59	3.40	3.31	3.21	3,11	3.00	2.89
20	9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3,96	3,85	3,68	3.50	3,32	3.22	3.12	3,02	2.92	2.81
21	9,83	6.89	5,73	5.09	4.68	4.39	4.18	4.01	3,88	3.77	3.60	3.43	3.24	3.15	3.05	2,95	2.84	2.73
22	9,73	6,81	5,65	5.02	4.61	4.32	4.11	3.94	3,81	3,70	3,54	3.36	3,18	3,08	2.98	2.88	2.77	2.66
23	9,63	6.73	5.58	4.95	4.54	4.26	4.05	3,88	3,75	3.64	3.47	3.30	3.12	3,02	2.92	2.82	2,71	2.60
24	9,55	6,66	5.52	4.89	4.49	4.20	3.99	3.83	3,69	3,59	3.42	3.25	3.06	2,97	2.87	2.77	2.66	2.55
25	9,48	6.60	5.46	4.84	4.43	4,15	3.94	3,78	3.64	3.54	3.37	3.20	3.01	2.92	2,82	2.72	2.61	2.50
26	9,41	6.54	5.41	4.79	4.38	4.10	3,89	3,73	3,60	3.49	3,33	3.15	2.97	2.87	2.77	2,67	2.56	2.45
27	9,34	6.49	5,36	4.74	4.34	4.06	3,85	3.69	3.56	3.45	3,28	3.11	2.93	2.83	2.73	2.63	2.52	2.41
28	9,28	6,44	5.32	4,70	4.30	4.02	3.81	3.65	3,52	3.41	3.25	3.07	2,89	2,79	2.69	2.59	2.48	2.37
29	9.23	6.40	5.28	4.66	4.26	3.98	3.77	3,61	3.48	3.38	3.21	3.04	2.86	2.76	2,66	2.56	2.45	2.33
30	9,18	6.35	5.24	4.62	4.23	3,95	3.74	3.58	3.45	3.34	3.18	3.01	2,82	2.73	2.63	2.52	2.42	2.30
40	8.83	6,07	4.98	4.37	3.99	3.71	3,51	3.35	3,22	3.12	2.95	2.78	2,60	2,50	2.40	2,30	2,18	2.06
60	8,49	5.79	4,73	4.14	3.76	3.49	3.29	3.13	3,01	2.90	2.74	2.57	2.39	2.29	2.19	2.08	1,96	1.83
120	8.18	5.54	4,50	3.92	3.55	3.28	3.09	2.93	2.81	2.71	2.54	2.37	2.19	2.09	1.98	1.87	1.75	1.61
∞	7,88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62	2.52	2.36	2,19	2.00	1.90	1.79	1.67	1.53	1.36

Таблица П7 (*окончание*)

k_2	k_1																	
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
	$p = 0,999$																	
1	4053+	5000+	5404+	5625+	5764+	5859+	5929+	25981+	6023+	6056+	6107+	6158+	6209+	6235+	6261+	6287+	6313+	6340+
2	998,5	999.0	999.2	999,2	999.3	999,3	999,4	999.4	999.4	999.40	999.4	999.4	999.4	999.5	999.5	999.5	999.5	999.5
3	167,0	148.5	141.1	137.1	134.6	132.8	131.6	130.6	129.9	129.2	128.3	127.4	126.4	125.9	125.4	125.0	124.5	124,0
4	74,14	61.25	56,18	53.44	51.71	50.53	49.66	49.00	48.47	48.05	47,41	46.76	46,10	45.77	45.43	45.09	44.75	44.40
5	47.18	37.12	33.20	31,09	29.75	28,84	28.16	27.64	27,24	26.92	26.42	25.91	25,39	25.14	24,87	24,60	24.33	24.06
6	35,51	27,00	23.20	21,92	20.81	20.03	19.46	19.03	18.69	18.41	17.99	17.56	17.12	16,89	16.67	16.44	16.21	15.99
7	29.25	21.69	18.77	17.19	16.21	15.52	15.02	14.63	14.33	14.08	13.71	13,32	12,93	12.73	12.53	12.33	12,12	11,91
8	25.42	18,49	15.83	14.39	13,49	12.86	12.40	12,04	11,77	11.54	11.19	10.84	10.48	10.30	10.11	9.92	9.73	9.53
9	22.86	16.29	13.90	12,56	11.71	11.13	10.70	10.37	10.11	9,89	9.57	9.24	8.90	8.72	8.55	8.37	8.19	8.00
10	21,04	14.91	12,55	11.28	10.48	9,92	9.52	9.20	8.96	8.75	8.45	8.13	7.80	7.64	7.47	7.30	7.12	6.94
11	19,69	13.81	11.56	10.35	9.58	9.05	8.66	8.35	8,12	7.92	7,61	7.32	7.01	6.85	6.68	6.52	6.35	6.17
12	18.64	12.97	10.80	9.63	8.89	8.38	8.00	7,71	7.48	7.29	7.00	6.71	6.40	6.25	6.09	5.93	5.76	5.59
13	17.81	12.31	10.21	9,07	8.35	7.86	7.49	7.21	6,98	6.80	6.52	6.23	5.93	5,78	5.63	5.47	5.30	5.14
14	17,14	11.78	9.73	8.62	7.92	7.43	7.08	6.80	6,58	6.40	6,13	5.85	5.56	5.41	5.25	5.10	4.94	4.77
15	16,50	11,34	9.34	8.25	7.57	7.09	6.74	6,47	6.26	6.08	5.81	5.54	5.25	5.10	4.95	4.80	4.64	4.47
16	16,12	10.97	9.00	7.94	7.27	6,81	6.46	6,19	5,98	5,81	5.55	5.27	4.99	4.85	4.70	4.54	4.39	4.21
17	15,72	10,66	8.73	7.68	7,02	6.56	6.22	5.96	5.75	5.58	5,32	5,05	4.78	4,63	4.48	4.33	4.18	4.02
18	15.38	10.39	8.49	7.46	6.81	6.35	6.02	5.76	5.56	5.39	5.13	4.87	4.59	4,45	4.30	4,15	4.00	3.84
19	15,08	10,16	8,28	7.26	6.62	6.18	5.85	5.59	5,39	5.22	4.97	4.70	4.43	4.29	4.14	3.99	3.84	3,68
20	14.82	9.95	8,10	7.10	6,46	6.02	5.69	5.44	5.24	5.08	4.82	4.56	4,29	4.15	4.00	3,86	3.70	3,54
21	14,59	9,77	7.94	6.95	6.32	5.88	5.56	5.31	5.11	4.95	4.70	4.44	4.17	4.03	3.88	3.74	3,58	3.42
22	14,38	9.61	7.80	6.81	6.19	5.76	5.44	5.19	4.99	4.83	4.58	4.33	4,06	3.92	3.78	3.63	3.48	3.32
23	14.19	9.47	7,67	6.69	6.08	5.65	5.33	5,09	4.99	4,71	4.48	4.23	3.96	3.82	3.68	3.53	3.38	3.22
24	14.03	9,34	7,55	6.59	5.98	5.55	5.23	4.99	4,80	4.64	4,39	4.14	3.87	3.74	3.59	3.45	3.29	3.14
25	13.88	9,22	7.45	6.49	5.88	5.46	5.15	4.91	4.71	4.56	4.31	4.06	3.79	3.66	3.52	3.37	3.22	3.06
26	13,74	9.12	7.36	6,41	5,80	5,38	5.07	4.83	4.64	4.48	4.24	3.99	3.72	3,59	3.44	3.30	3.15	2,99
27	13.61	9.02	7.27	6.33	5.73	5.31	5.00	4.76	4,57	4.41	4.17	3.92	3.66	3.52	3.38	3.23	3.08	2.92
28	13.50	8.93	7,19	6.25	5,66	5.24	4.93	4.69	4,50	4.35	4.11	3.86	3,60	3,46	3.32	3.18	3.02	2,86
29	13,39	8.85	7,12	6.19	5.59	5.18	4.87	4,64	4.45	4.29	4.05	3.80	3.54	3.41	3.27	3.12	2.97	2.81
30	13,29	8.77	7,05	6.12	5.53	5,12	4.82	4,58	,39	4.24	4.00	3,75	3.49	3.36	3.22	3.07	2.92	2.76
40	12.61	8.25	6.60	5.70	5.13	4.73	4.44	4.21	4,02	3.87	3,64	3.40	3.15	3.01	2.87	2.73	2,57	2.41
60	11.97	7.76	6.17	5.31	4.76	4.37	4.09	3.87	3.69	2.54	3.31	3,08	2,83	2.69	2.55	2.41	2.25	2.08
120	11.38	7.32	5.79	4.95	4.42	4.04	3.77	3.55	3,38	3.24	3.02	2.78	2.53	2.40	2.26	2.11	1.95	1.76
∞	10.83	6.91	5,42	4.62	4.10	3.74	3.47	3.27	3,10	2,96	2.74	2.51	2.27	2.13	1.99	1.84	1.66	1,45

4053+ означает $4053 \cdot 10^2$

Таблица П8. **Значения функции $r = \operatorname{th} z$**

z	0,00	0,02	0,04	0,06	0,08
0,0	0,000	0,020	0,040	0,060	0,080
0,1	0,100	0,119	0,139	0,159	0,178
0,2	0,197	0,217	0,236	0,254	0,273
0,3	0,291	0,310	0,328	0,345	0,363
0,4	0,380	0,397	0,414	0,430	0,446
0,5	0,462	0,478	0,493	0,508	0,523
0,6	0,537	0,551	0,565	0,578	0,592
0,7	0,604	0,617	0,629	0,641	0,653
0,8	0,664	0,675	0,686	0,696	0,706
0,9	0,716	0,726	0,735	0,744	0,753
1,0	0,762	0,770	0,778	0,786	0,793
1,1	0,801	0,808	0,814	0,821	0,828
1,2	0,834	0,840	0,846	0,851	0,851
1,3	0,862	0,867	0,872	0,876	0,881
1,4	0,885	0,890	0,894	0,898	0,902
1,5	0,905	0,909	0,912	0,915	0,919
1,6	0,922	0,925	0,928	0,930	0,933
1,7	0,936	0,938	0,940	0,943	0,945
1,8	0,947	0,949	0,951	0,953	0,955
1,9	0,956	0,958	0,960	0,961	0,963

Таблица П9. **Значения ортогональных полиномов $P_k(i)$**

$n = 8$						$n = 10$					
i	P_1	P_2	P_3	P_4	P_5	i	P_1	P_2	P_3	P_4	P_5
1	−7	+7	−7	+7	−7	1	−9	+6	−42	+18	−6
2	−5	+1	+5	−13	+23	2	−7	+2	+14	−22	+14
3	−3	−3	+7	−3	−17	3	−5	−1	+35	−17	−1
4	−1	−5	+3	+9	−15	4	−3	−3	+31	+3	−11
5	+1	−5	−3	+9	+15	5	−1	−4	+12	+18	−6
6	+3	−3	−7	−3	+17	6	+1	−4	−12	+18	+6
7	+5	+1	−5	−13	−23	7	+3	−3	−31	+3	+11
8	+7	+7	+7	+7	+7	8	+5	−1	−35	−17	+1
						9	+7	+2	−14	−22	−14
						10	+9	+6	+42	+18	+6
$\sum P_k^2$	168	168	264	616	2184	$\sum P_k^2$	330	132	8580	2860	780
λ_k	2	1	2/3	7/12	7/10	λ_k	2	1/2	5/3	5/12	1/10

$n = 12$						$n = 13$					
i	P_1	P_2	P_3	P_4	P_5	i	P_1	P_2	P_3	P_4	P_5
1	−11	+55	−33	+33	−33	1	−6	+22	−11	+99	−22
2	−9	+25	+3	−27	+57	2	−5	+11	0	−66	+33
3	−7	+1	+21	−33	+21	3	−4	+2	+6	−96	+18
4	−5	−17	+25	−13	−29	4	−3	−5	+8	−54	−11
5	−3	−29	+19	+12	−44	5	−2	−10	+7	+11	−26
6	−1	−35	+7	+28	−20	6	−1	−13	+4	+64	−20
7	+1	−35	−7	+28	+20	7	0	−14	0	+84	0
8	+3	−29	−19	+12	+44	8	+1	−13	−4	+64	+20
9	+5	−17	−25	−13	+29	9	+2	−10	−7	+11	+26
10	+7	+1	−21	−33	−21	10	+3	−5	−8	−54	+11
11	+9	+25	−3	−27	−57	11	+4	+2	−6	−96	−18
12	+11	+55	+33	+33	+33	12	+5	+11	0	−66	−33
						13	+6	+22	+11	+99	+22
$\sum P_k^2$	572	12012	5148	8008	15912	$\sum P_k^2$	182	2002	572	68068	6188
λ_k	2	3	2/3	7/24	3/20	λ_k	1	1	1/6	7/12	7/120

Таблица П10. **Критерий Вилкоксона. Манна и Уитни. Вероятности $p = \mathbf{P}[W < w_в]$ для выборок объема n_1 и n_2 $(n_1 \geqslant n_2)$**

$n_1 = 3$

$w_в$	n_2		
	1	2	3
0	0,250	0.100	0,050
1	0,500	0.200	0.100
2	0,750	0,400	0,200
3		0,600	0,350
4			0,500
5			0,650

$n_1 = 4$

$w_в$	n_2			
	1	2	3	4
0	0,200	0,067	0,028	0,014
1	0,400	0,133	0,057	0,029
2	0,600	0.267	0,114	0.057
3		0,400	0,200	0,100
4		0.600	0,314	0.171
5			0.429	0,243
6			0,571	0.343
7				0,344
8				0,557

$n_1 = 5$

$w_в$	n_2				
	1	2	3	4	5
0	0.167	0,047	0,018	0,008	0,004
1	0.333	0,095	0,036	0,016	0,008
2	0,500	0.190	0,071	0,032	0,016
3	0.667	0,286	0,125	0,056	0.028
4		0,429	0,196	0,095	0,048
5		0,571	0,286	0,143	0.075
6			0,393	0,206	0,111
7			0,500	0,278	0,155
8			0.607	0,365	0,210
9				0,452	0,274
10				0,548	0,345
11					0,421
12					0,500
13					0,579

$n_1 = 6$

$w_в$	n_2					
	1	2	3	4	5	6
0	0,143	0,036	0,012	0,005	0,002	0,001
1	0,286	0,071	0,024	0,010	0,004	0,002
2	0,428	0,143	0,048	0,019	0,009	0.004
3	0,571	0,214	0,083	0,033	0,015	0,008
4		0,321	0.131	0,057	0.026	0,013
5		0,429	0,190	0,086	0,041	0,021
6		0,571	0.274	0,129	0,063	0,032
7			0,357	0,176	0,089	0,047
8			0,452	0,238	0,123	0,066
9			0,548	0,305	0.165	0.090
10				0,381	0.214	0,120
11				0,457	0,268	0,155
12				0,545	0,331	0,197
13					0.396	0,242
14					0,465	0,294
15					0,535	0,350
16						0.409
17						0,469
18						0,531

Таблица П10 (*подолжение*)

$n_1 = 7$

$w_в$	n_2						
	1	2	3	4	5	6	7
0	0,125	0,028	0,008	0,003	0,001	0,001	0,000
1	0,250	0,056	0,017	0,006	0,003	0,001	0,001
2	0,375	0,111	0,033	0,012	0,005	0,002	0,001
3	0,500	0,167	0,058	0,021	0,009	0,004	0,002
4	0,625	0,250	0,092	0,036	0,015	0,007	0,003
5		0,333	0,133	0,055	0,024	0,011	0,006
6		0,444	0,192	0,082	0,037	0,017	0,009
7		0,556	0,258	0,115	0,053	0,026	0,013
8			0,333	0,158	0,074	0,037	0,019
9			0,417	0,206	0,101	0,051	0,027
10			0,500	0,264	0,134	0,069	0,036
11			0,583	0,324	0,172	0,090	0,049
12				0,394	0,216	0,117	0,064
13				0,464	0,265	0,147	0,082
14				0,538	0,319	0,183	0,104
15					0,378	0,223	0,130
16					0,438	0,267	0,159
17					0,500	0,314	0,191
18					0,562	0,365	0,228
19						0,418	0,267
20						0,473	0,310
21						0,527	0,355
22							0,402
23							0,451
24							0,500
25							0,549

$n_1 = 8$

$w_в$	n_2							
	1	2	3	4	5	6	7	8
0	0,111	0,022	0,006	0,002	0,001	0,000	0,000	0,000
1	0,222	0,044	0,012	0,004	0,002	0,001	0,000	0,000
2	0,333	0,089	0,024	0,008	0,003	0,001	0,001	0,000
3	0,444	0,133	0,042	0,014	0,005	0,002	0,001	0,001
4	0,556	0,200	0,067	0,024	0,009	0,004	0,002	0,001
5		0,267	0,097	0,036	0,015	0,006	0,003	0,001
6		0,356	0,139	0,055	0,023	0,010	0,005	0,002
7		0,444	0,188	0,077	0,033	0,015	0,007	0,003
8		0,556	0,248	0,107	0,047	0,021	0,010	0,005
9			0,315	0,141	0,064	0,030	0,014	0,007
10			0,387	0,184	0,085	0,041	0,020	0,010
11			0,461	0,230	0,111	0,054	0,027	0,014
12			0,539	0,285	0,142	0,071	0,036	0,019
13				0,341	0,177	0,091	0,047	0,025
14				0,404	0,217	0,114	0,060	0,032
15				0,467	0,262	0,141	0,075	0,041
16				0,533	0,311	0,172	0,095	0,052
17					0,362	0,207	0,116	0,063
18					0,416	0,245	0,140	0,080
19					0,472	0,286	0,168	0,097
20					0,528	0,331	0,198	0,117
21						0,377	0,232	0,139
22						0,426	0,268	0,164
23						0,475	0,306	0,191
24						0,525	0,347	0,221
25							0,389	0,253
26							0,433	0,287
27							0,478	0,323
28							0,522	0,360
29								0,399
30								0,439
31								0,480
32								0,522

Таблица П11.

Критические значения N_1 и N_2 для критерия серий при уровне значимости $\alpha = 0{,}05$. В заголовке столбца стоит наибольшее из чисел n_1 и n_2, которые равны количествам одинаковых знаков в последовательности знаков. Номер строки соответствует меньшему из чисел n_1 и n_2.

	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	– –	– –	– –	– –	– –	– –	– –	2 –	2 –	2 –	2 –	2 –	2 –	2 –	2 –	2 –
3	– –	2 –	2 –	2 –	2 –	2 –	2 –	2 –	2 –	2 –	3 –	3 –	3 –	3 –	3 –	3 –
4	2 9	2 9	2 –	3 –	3 –	3 –	3 –	3 –	3 –	3 –	3 –	4 –	4 –	4 –	4 –	4 –
5	2 10	3 10	3 11	3 11	3 –	3 –	4 –	4 –	4 –	4 –	4 –	4 –	4 –	5 –	5 –	5 –
6		3 11	3 12	3 12	4 13	4 13	4 13	4 13	5 –	5 –	5 –	5 –	5 –	5 –	6 –	6 –
7			3 13	4 13	4 14	5 14	5 14	5 14	5 15	5 15	6 15	6 –	6 –	6 –	6 –	6 –
8				4 14	5 14	5 15	5 15	6 16	6 16	6 16	6 16	6 17	7 17	7 17	7 17	7 17
9					5 15	5 16	6 16	6 16	6 17	7 17	7 18	7 18	7 18	8 18	8 18	8 18
10						6 16	6 17	7 17	7 18	7 18	7 18	8 19	8 19	8 19	8 20	9 20
11							7 17	7 18	7 19	8 19	8 19	8 20	9 20	9 20	9 21	9 21
12								7 19	8 19	8 20	8 20	9 21	9 21	9 21	10 22	10 22
13									8 20	9 20	9 21	9 21	10 22	10 22	10 23	10 23
14										9 21	9 22	10 22	10 23	10 23	11 23	11 24
15											10 22	10 23	11 23	11 24	11 24	12 25
16												11 23	11 24	11 25	12 25	12 25
17													11 25	12 25	12 26	13 26
18														12 26	13 26	13 27
19															13 27	13 27
20																14 28

Таблица П12. **Равномерно распределенные случайные числа**

10	09	73	25	33	76	52	01	35	86	34	67	35	48	76
37	54	20	48	05	64	89	47	42	96	24	80	52	40	37
08	42	26	89	53	19	64	50	93	03	23	20	90	25	60
99	01	90	25	29	09	37	67	07	15	38	31	13	11	65
12	80	79	99	70	80	15	73	61	47	64	03	23	66	53
80	95	90	91	17	39	29	27	49	45	66	06	57	47	17
20	63	61	04	02	00	82	29	16	65	31	06	01	08	05
15	95	33	47	64	35	08	03	36	06	85	26	97	76	02
88	67	67	43	97	04	43	62	76	59	63	57	33	21	35
98	95	11	68	77	12	17	17	68	33	73	79	64	57	53
34	07	27	68	50	36	69	73	61	70	65	81	33	98	85
45	57	18	24	06	35	30	34	26	14	86	79	90	74	39
02	05	16	56	92	68	66	57	48	18	73	05	38	52	47
05	32	54	70	48	90	55	35	75	48	28	46	82	87	09
03	52	96	47	78	35	80	83	42	82	60	93	52	03	44
11	19	92	91	70	98	52	01	77	67	14	90	56	86	07
23	40	30	97	32	11	80	50	54	31	39	80	82	77	32
18	62	38	85	79	83	45	29	96	34	06	28	89	80	83
83	49	12	56	24	88	68	54	02	00	86	50	75	84	01
35	27	38	84	35	99	59	46	73	48	87	51	76	49	69
22	10	94	05	58	60	97	09	34	33	50	50	07	39	98
0	72	56	82	48	29	40	52	42	01	52	77	56	78	51
13	74	67	00	78	18	47	54	06	10	68	71	17	78	17
36	76	66	79	51	90	36	47	64	93	29	60	91	10	62
91	82	60	89	28	93	78	56	13	68	23	47	83	41	13
65	48	11	76	74	17	46	85	09	50	58	04	77	69	74
80	12	43	56	35	17	72	70	80	15	45	31	82	23	74
74	35	09	98	17	77	40	27	72	14	43	23	60	02	10
69	91	62	68	03	66	25	22	91	48	36	93	68	72	03
09	89	32	05	05	14	22	56	85	14	46	42	75	67	88
73	03	95	71	86	40	21	81	65	44	91	49	91	45	23
21	11	57	82	53	14	38	55	37	63	80	33	69	45	98
45	52	16	42	37	96	28	60	26	55	44	10	48	19	49
76	62	11	39	90	94	40	05	64	18	12	55	07	37	42
96	29	77	88	22	54	38	21	45	98	63	60	64	93	29
68	47	92	76	86	46	16	28	35	54	94	75	08	99	23
26	94	03	68	58	70	29	73	41	35	53	14	03	33	40
85	15	74	79	54	32	97	92	65	75	57	60	04	08	81
11	10	00	20	40	12	86	07	46	97	96	64	48	94	39
16	50	53	44	84	40	21	95	25	63	43	65	17	70	82

СПИСОК ЛИТЕРАТУРЫ

1. Чистяков В.П. Курс теории вероятностей. — 3-е изд. — М.: Наука, 1987.
2. Гнеденко Б.В. Курс теории вероятностей. — 6-е изд. — М.: Наука, 1988.
3. Коваленко И.Н., Филиппова А.А. Теория вероятностей и математическая статистика. — 2-е изд. — М.: Высшая школа, 1982.
4. Вентцель Е.С. Теория вероятностей. — М.: Наука, 1964.
5. Лавренченко А.С. Лекции по математической статистике и теории случайных процессов. — М.: Изд-во МАИ, 1974.
6. Сборник задач по теории вероятностей, математической статистике и теории случайных функций / Под общ. ред. А.А.Свешникова. — 2-е изд. — М.: Наука, 1970.
7. Емельянов Г.В., Скитович В.П. Задачник по теории вероятностей и математической статистике. — Л.: Изд-во ЛГУ, 1967.
8. Севастьянов Б.А., Чистяков В.П., Зубков А.М. Сборник задач по теории вероятностей. — М.: Наука, 1980.
9. Смирнов Н.В., Дунин-Барковский И.В. Курс теории вероятностей и математической статистики для технических приложений. — 3-е изд. — М.: Наука, 1969.
10. Браунли К.А. Статистическая теория и методология в науке и технике. — М., 1977.
11. Химмельблау Д. Анализ процессов статистическими методами. — М.: Мир, 1973.
12. Дрейпер М., Смит Г. Прикладной регрессионный анализ. Кн. 1. — М.: Финансы и статистика, 1986.
13. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере. — М.: Финансы и статистика, 1995.
14. Боровиков В.П., Боровиков И.П. Statistica. Статистический анализ и обработка данных в среде Windows. — М.: Филинъ, 1997.
15. Дюк В. Обработка данных на ПК в примерах. — Спб. Питер, 1997.
16. Потемкин В.Г. Система Matlab. — М.: «Диалог–МИФИ», 1997.

СБОРНИК ЗАДАЧ ПО МАТЕМАТИКЕ ДЛЯ ВТУЗОВ.
В 4 частях / Под общ. ред. А.В.Ефимова, А.С.Поспелова

Содержание частей 1–4

СБОРНИК ЗАДАЧ ПО МАТЕМАТИКЕ

ДЛЯ ВТУЗОВ

4

- ТЕОРИЯ ВЕРОЯТНОСТЕЙ
- МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

ФМ