Н. А. МИКУЛИК Г. Н. РЕЙЗИНА

РЕШЕНИЕ ТЕХНИЧЕСКИХ ЗАДАЧ ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ

Н. А. МИКУЛИК Г. Н. РЕЙЗИНА

РЕШЕНИЕ ТЕХНИЧЕСКИХ ЗАДАЧ ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ

СПРАВОЧНОЕ ПОСОБИЕ

Минск
"Вышэйшая школа"
1991

ББК 22.171я2
М59
УДК 519.2(076.1) (035.5)

Р е ц е н з е н т: зав. кафедрой теории вероятностей и математической статистики БГУ им. В.И. Ленина, проф. *Г.А. Медведев*

Микулик Н.А., Рейзина Г.Н.
М59 Решение технических задач по теории вероятностей и математической статистике: Справ. пособие. – Мн.: Выш. шк., 1991. – 164 с.: ил.
ISBN 5-339-00561-5.

Содержатся задачи по всем основным разделам теории вероятностей и математической статистики. Большое внимание уделено контролю качества и основным вопросам планирования эксперимента. Большинство задач имеет техническое содержание, что важно для профессиональной ориентации студента втуза, молодого специалиста.

Для студентов технических вузов. Будет полезно инженерно-техническим работникам.

$М\frac{1602090000-094}{М304(03)-91}$ 61–91

ББК 22.171я2+22.172я2

ISBN 5-339-00561-5

ПРЕДИСЛОВИЕ

В настоящее время трудно назвать область науки и техники, где бы не использовались теоретико-вероятностные и статистические методы. Новые учебные планы и типовая программа по высшей математике, введенные в практику с 1988/89 учебного года, предусматривают усиление прикладной направленности курса высшей математики, а также обеспечение студентов соответствующими учебными пособиями.

Вниманию читателя предлагается справочное пособие, написанное на основе учебного пособия этих же авторов "Руководство к решению задач по теории вероятностей и математической статистике" (Мн.: Выш. шк., 1977), которое подверглось существенной переработке и дополнено новыми главами. В книге содержатся сведения из теории вероятностей, математической статистики, теории корреляции, теории случайных функций и планирования эксперимента.

Цель пособия – помочь изучающим теорию вероятностей и математическую статистику приобрести навыки в решении технических задач. Доказательство ряда теорем и выводы формул читатель может найти в учебных изданиях, указанных в списке рекомендуемой литературы.

В пособии приводятся краткие теоретические сведения и формулы, необходимые для решения задач, типовые примеры с решениями, а также задачи для самостоятельного решения. Все задачи снабжены ответами. Большинство из приведенных задач – с техническим содержанием (технология машиностроения и станкостроения, конструирование автомобилей). В книге использованы задачи, составленные авторами, а также задачи из отечественных источников.

В приложениях даны сведения, необходимые при решении задач.

Пособие предназначается в первую очередь для студентов технических специальностей вузов. Оно будет полезно инженерам и другим работникам, применяющим теорию вероятностей и математическую статистику.

Авторы выражают благодарность рецензенту – заведующему кафедрой теории вероятностей и математической статистики БГУ, профессору Г.А. Медведеву – за ценные советы и замечания, способствовавшие улучшению пособия.

Все отзывы и пожелания просьба присылать по адресу: 220048, Минск, проспект Машерова, 11, издательство "Вышэйшая школа".

Авторы

1. СЛУЧАЙНЫЕ СОБЫТИЯ

1.1. ЭЛЕМЕНТАРНЫЕ СОБЫТИЯ

Предметом изучения теории вероятностей являются математические модели экспериментов, результаты которых неоднозначно определяются условиями опыта. Задача теории вероятностей состоит в установлении закономерностей, которым подчиняются результаты таких экспериментов.

Любой результат эксперимента называется *элементарным событием* и обозначается ω. Произвольное множество всех элементарных событий, относящихся к одному и тому же эксперименту, называется *пространством элементарных событий* и обозначается Ω.

Событием или *случайным событием* называется любое подмножество пространства элементарных событий $\Omega(\omega)$. События обозначают прописными буквами латинского алфавита $A, B, C, \ldots$

Событие, состоящее из всех возможных элементарных событий ω, называется *достоверным* и обозначается Ω (так же, как и пространство элементарных событий). Событие, не содержащее ни одного элементарного события, называется *невозможным* и обозначается $\varnothing$ (пустое множество).

Рассмотрим соотношения между событиями.

1°. Если всякое элементарное событие ω из A принадлежит случайному событию B, то говорят, что *событие A влечет за собой событие B* (обозначается $A \subset B$ или $B \supset A$). Другими словами, если $A \subset B$, то случайное событие B происходит всегда, если происходит событие A.

2°. События A и B называются *равными* или *эквивалентными* (обозначается $A = B$), если они состоят из одних и тех же элементарных событий, т.е. всегда происходят или не происходят одновременно.

3°. *Суммой* (*объединением*) *событий* A и B (обозначается $A + B$ или $A \cup \cup B$) называется событие, которое состоит из элементарных событий, входящих хотя бы в одно из событий A и B, т.е. событие, происходящее тогда и только тогда, когда происходит хотя бы одно из событий A и B.

4°. *Произведением* (*пересечением*) *событий* A и B (обозначается AB или $A \cap B$) называется событие, которое состоит из элементарных событий, входящих и в событие A, и в событие B, т.е. событие, происходящее тогда и только тогда, когда происходит и событие A, и событие B.

5°. *Событием, противоположным событию* A (обозначается $\overline{A}$), называется событие, которое состоит из всех элементарных событий, не входящих в A.

6°. События A и B называются *несовместными*, если $AB = \varnothing$.

7°. Разностью событий A и B (обозначается $A - B$ или $A \setminus B$) называется событие, равное $A \setminus B = \overline{AB}$.

8°. *Симметрической разностью событий* A и B (обозначается $A \circ B$) называется событие, равное $A \circ B = A\overline{B} + B\overline{A} = \overline{AB}$.

9°. События $A_1, A_2, \ldots, A_n$ образуют *полную группу событий*, если $A_iA_j = \varnothing$, $i \neq j$, и $\bigcup\limits_{i=1}^{n} A_i = \Omega$.

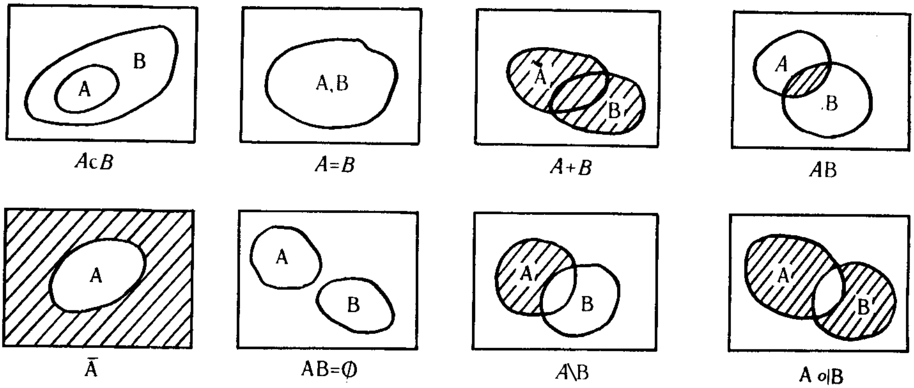


Рис. 1.1

Наглядной иллюстрацией всех этих понятий является *диаграмма Венна* (рис. 1.1).

Рассмотрим эксперимент, состоящий в том, что в квадрате случайным образом выбирается точка. Пространством элементарных событий Ω данного эксперимента является множество точек квадрата. Пусть событие A состоит из точек области A, событие B – из точек области B. Тогда можно рассмотреть следующие свойства (A, B, C – произвольные события):

1) $A \subset B \Leftrightarrow \overline{B} \subset \overline{A}$;

2) $A = B \Leftrightarrow A \subset B, B \subset A$;

3) $A = B \Leftrightarrow \overline{A} = \overline{B}$;

4) $A(B \quad = AB + AC$;

5) $\overline{A + B} = \overline{A}\,\overline{B}$;

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

З а м е ч а н и е. Равенство $(F_1 = F_2) \Leftrightarrow (F_1 \subset F_2, F_2 \subset F_1)$, где F_1, F_2 –любые выражения, полученные из случайных событий A и B с помощью операций сложения, умножения, дополнения.

Алгеброй событий называется класс событий, замкнутый относительно операций объединения, пересечения и дополнения. Рассмотрим некоторый класс событий **A** из пространства Ω, в котором: 1) дополнение любого события из **A** также принадлежит **A**; 2) для любых двух событий из **A** их объединение и пересечение принадлежат **A**; 3) достоверное событие Ω принадлежит **A**. Тогда множество **A** с введенными на нем операциями объединения, пересечения и дополнения является алгеброй событий.

Пример 1.1. Монету подбрасывают один раз. Построить для данного случая пространство элементарных событий и алгебру событий.

Р е ш е н и е. В данном случае пространство элементарных событий состоит из двух элементарных событий: выпадения герба (Г) и выпадения цифры (Ц), т.е. $\Omega = \{Г, Ц\}$.

Составим множество событий **A** $= \{ \varnothing, \{Г\}, \{Ц\}, \Omega\}$. Это множество с определенными на нем операциями объединения, пересечения и дополнения образует алгебру событий. Составим таблицы сложения, умножения и дополнения элементов множества **A** (рис. 1.2).

∪	$\emptyset$	{Г}	{Ц}	Ω
$\emptyset$	$\emptyset$	{Г}	{Ц}	Ω
{Г}	{Г}	{Г}	Ω	Ω
{Ц}	{Ц}	Ω	{Ц}	Ω
Ω	Ω	Ω	Ω	Ω

∩	$\emptyset$	{Г}	{Ц}	Ω
$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
{Г}	$\emptyset$	{Г}	$\emptyset$	{Г}
{Ц}	$\emptyset$	$\emptyset$	{Ц}	{Ц}
Ω	$\emptyset$	{Г}	{Ц}	Ω

A	$\emptyset$	{Г}	{Ц}	Ω
$\bar{A}$	Ω	{Ц}	{Г}	$\emptyset$

Рис. 1.2

Из приведенных таблиц видно, что многократное применение операций объединения, пересечения и дополнения к событиям из множества **А** снова приводит к событиям из **А**. Таким образом, **А** – алгебра событий.

1.2. ОПРЕДЕЛЕНИЕ ВЕРОЯТНОСТИ И ЕЕ ОСНОВНЫЕ СВОЙСТВА

Определение вероятности как функции множества. Если задано множество Ω (пространство элементарных событий) и алгебра **А** его подмножеств, то говорят, что *построено измеримое пространство* (Ω, **А**).

Числовая функция $P(A)$, определенная на множествах $A \subset$ **А** и удовлетворяющая следующим аксиомам:

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

2) $P(\Omega) = 1$;

3) $P(\bigcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} P(A_i)$ при $A_i \cap A_j = \emptyset$, $i \neq j$,

называется *вероятностью* на (Ω, **А**) .

Из аксиом 1–3, предложенных А.Н. Колмогоровым, следует, что вероятность $P(A)$, определенная на алгебре **А**, соответствует любому событию $A \in$ **А** и вероятность $P(A)$ представляет собой неотрицательную количественную характеристику, называемую ***вероятностной мерой.***

Пространство элементарных событий Ω с выделенной в нем алгеброй **А** и определенной на измеримом пространстве (Ω, **А**) вероятностной мерой $P(A)$, $A \in$ **А**, называется ***вероятностным пространством*** и обозначается (Ω, **А**, P).

Классический (лапласовский) метод задания вероятности. Рассмотрим вероятностный эксперимент и его пространство элементарных событий $\Omega = \{\omega_i \mid i = 1, 2, ..., n\}$, где ω_i равновозможны. Пусть элементарные события ω_{i_k}, $k = 1,2,...,m$, входящие в событие $A = \{\omega_{i_1}, \omega_{i_2}, ..., \omega_{i_m}\}$, благоприятствуют событию A. Тогда вероятность наступления события определяется как отношение числа m элементарных событий, благоприятствующих событию A, к общему числу n равновозможных элементарных событий на Ω: $P(A) = m/n$. Функция $P(A)$ удовлетворяет аксиомам 1–3.

Статистический метод задания вероятности. Пусть проведен вероятностный эксперимент и пространство элементарных событий фиксировано. Повторим эксперимент N раз при тех же условиях. Если в этом случае событие $A \in$ $\in$ **A** произошло M раз, то отношение $M/N = W(A)$ числа наступления события к общему числу повторений вероятностного эксперимента называется *относительной частотой* или *частостью*. Так как при испытаниях число M появления события A может изменяться от 0 до N, то относительная частота $W(A)$ удовлетворяет условию $0 \leqslant W(A) \leqslant 1$.

Длительные наблюдения показали, что если в одинаковых условиях проводятся опыты, в каждом из которых число испытаний достаточно велико, то относительная частота изменяется мало, колеблясь около некоторого постоянного числа. В этом случае используют *статистическое определение вероятности*, принимая за вероятность события относительную частоту или число, близкое к ней.

Геометрическое определение вероятности. Пусть в результате испытания в некоторой области Ω появляется точка ω. Вероятность $P(A)$ того, что $\omega \in$ $\in A$, где A – область, принадлежащая Ω $(A \subset \Omega)$, определяется следующим образом: $P(A) = m(A)/m(\Omega)$ и называется *геометрической вероятностью;* $m(A)$ и $m(\Omega)$ – меры областей A и Ω соответственно. Под *мерой* понимается длина, площадь, объем в одно-, двух- и трехмерном пространствах.

Пример 1.2. Куб, все грани которого обработаны, распилен на 512 кубиков одинакового размера. Полученные кубики тщательно перемешаны. Определить вероятность того, что кубик, извлеченный наудачу, имеет две обработанные поверхности.

Р е ш е н и е. В рассматриваемом случае пространство элементарных событий состоит из $n = 512$ элементов: $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$, а алгебру событий **A** образует множество всех подмножеств Ω с операциями объединения, пересечения и дополнения. Интересующее нас событие **A** состоит из m элементарных событий: $m = 12 \cdot 6 = 72$, и принадлежит алгебре событий **A**. Считаем, что вероятности всех элементарных событий одинаковы, а значит,

$$P(A) = m/n = 72/512 \cong 0,14.$$

Пример 1.3. Найти вероятность того, что сумма выбранных наугад положительных чисел, каждое из которых не больше 1, не превзойдет 1, а их произведение будет не больше 1/3.

Р е ш е н и е. Пусть взяты числа x и y. Согласно условию, $0 < x < 1$, $0 < y < 1$, $x + y < 1$, $xy < 1/3$. Область всевозможных значений ограничена квадратом $OABC$ (рис. 1.3), площадь которого $S = 1$. Область благоприятствующих значений ограничена заштрихованной областью и площадь ее равна сумме площадей

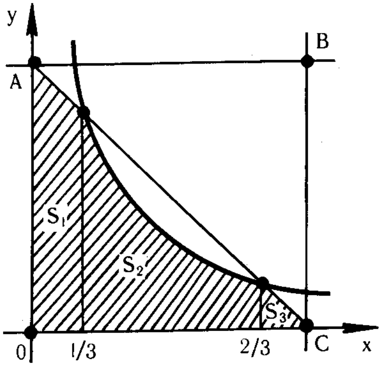


Рис. 1.3

$$S_1 + S_2 + S_3 = \int_0^{1/3} (1-x)\,dx + \int_{1/3}^{2/3} \frac{1}{3x}\,dx + \int_{2/3}^{1} (1-x)\,dx =$$

$$= \left(x - \frac{x^2}{2}\right)\Big|_0^{1/3} + \frac{1}{3}\ln x\Big|_{1/3}^{2/3} + \left(x - \frac{x^2}{2}\right)\Big|_{2/3}^{1} = \frac{1}{3} + \frac{1}{3}\ln 2.$$

Искомая вероятность

$$P(A) = \frac{1}{3}(1 + \ln 2)/1 = \ln(2e)/3.$$

ЗАДАЧИ

1.1. Доказать, что:

1) $(A + \overline{B})\ (\overline{A} + B)\ (A + B) = AB$;
2) $(A + B)\ (A + \overline{B})\ (\overline{A} + \overline{B})\ (\overline{A} + \overline{B}) = \emptyset$;
3) $(A + \overline{B})\ (A + \overline{B}) + A = A + \overline{B}$;
4) $(A + \overline{B})\ (\overline{A} + B) = \overline{A}B = \overline{A}\,\overline{B}$.

1.2. При движении автомобиля его левые и правые колеса наезжают на препятствия (выступы и впадины дорожного полотна). Пусть A – событие, заключающееся в наезде на препятствие левым колесом, B – правым колесом. В чем состоят следующие события: 1) $\overline{A}$; 2) $\overline{B}$; 3) $A + B$; 4) $\overline{A + B}$; 5) $\overline{AB}$?

1.3. Упростить выражение $A = (A + \overline{C})\ (B + C)\ (\overline{B} + C)$.

1.4. Когда возможны равенства: 1) $A + \overline{B} = A$; 2) $AB = \overline{A}$; 3) $A + B = AB$?

1.5. Найти случайное событие X, если $\overline{X + A} + \overline{X + \overline{A}} = B$.

1.6. Совместны ли события A и $\overline{A + B}$?

1.7. Доказать, что события A, $\overline{A}B$, $\overline{A + B}$ образуют полную группу. (*Указание.* Воспользоваться равенством $\overline{A + B} = \overline{A}\overline{B}$.)

1.8. Прибор состоит из двух блоков первого типа и трех блоков второго типа. Событие A_k $(k = 1, 2)$ означает, что исправен k-й блок первого типа, B_j $(j = 1, 2, 3)$ – исправен j-й блок второго типа. Прибор исправен, если исправны хотя бы один блок первого типа и не менее двух блоков второго типа. Выразить событие C, означающее исправность прибора, через A_k и B_j.

1.9. Из партии изготовленных шестерен, среди которых 20 годных и 5 бракованных, для контроля взято наудачу 8 штук. В результате контроля оказалось, что первые 3 шестерни из 8 годные. Определить вероятность того, что четвертая деталь будет годной.

1.10. Определить вероятность того, что номер первой встретившейся автомашины: 1) не содержит одинаковых цифр; 2) имеет две одинаковые цифры; 3) имеет три одинаковые цифры; 4) содержит две пары одинаковых цифр; 5) состоит из одинаковых цифр. Известно, что все номера четырехзначные (начиная с 0001), неповторяющиеся и равновозможные.

1.11. На сборку автомобилей поданы 10 дифференциалов, собранные двумя рабочими. Определить вероятность того, что 3 дифференциала, поставленные последовательно на машины, окажутся изготовленными первым рабочим, если известно, что первый рабочий собрал 6 дифференциалов, а второй – 4.

1.12. Из десяти подсобранных узлов карданной передачи два получили высокую оценку при контроле. Определить вероятность того, что среди взятых наудачу пяти узлов: 1) один высокого качества; 2) два высокого качества; 3) хотя бы один высокого качества.

1.13. В партии из 50 деталей 5 нестандартных. Определить вероятность того, что среди выбранных наудачу для проверки 6 деталей 2 окажутся нестандартными.

1.14. В точке C, положение которой на телефонной линии AB длиной L равновозможно, произошел обрыв. Определить вероятность того, что точка C удалена от точки A на расстояние, не меньшее l.

1.15. Стержень длиной l случайным образом разломан на три части. Определить вероятность того, что из этих частей можно построить треугольник.

1.3. УСЛОВНАЯ ВЕРОЯТНОСТЬ. НЕЗАВИСИМОСТЬ СОБЫТИЙ

Условной вероятностью наступления события A при условии, что произошло событие B, называется величина

$$\boxed{P(A/B) = P(AB)/P(B),} \tag{1.1}$$

где $P(B) > 0$. Если $A = B$, то $P(B/B) = 1$ и $P(A)$ в (1.1) называют *безусловной вероятностью*.

Из определения условной вероятности следует, что

$$P(AB) = P(B)P(A/B),$$

т.е. вероятность совместного наступления двух событий равна произведению вероятности одного из них на условную вероятность другого при условии, что первое наступило.

События A и B называются ***независимыми***, если

$$P(AB) = P(A)\ P(B)$$

или, что то же самое, $P(A/B) = P(A)$, т.е. условная вероятность события A при условии, что произошло событие B, равна безусловной вероятности этого события.

События $A_1, A_2, \ldots, A_n$ называются ***независимыми в совокупности***, если для любых k из них выполняется соотношение

$$P(A_{i_1}, A_{i_2}, \ldots, A_{i_k}) = \prod_{j=1}^{k} P(A_{i_j}). \tag{1.2}$$

Если соотношение (1.2) выполняется только при $k = 2$, то события A_1, $A_2, \ldots, A_n$ называются *попарно независимыми*. Отметим, что из попарной независимости событий не следует независимость их в совокупности.

Пример 1.4. Рабочий обслуживает два независимо работающих станка. Вероятности того, что в течение некоторого времени станки не потребуют внимания рабочего, соответственно равны: $P(A) = 0{,}7$, $P(B) = 0{,}8$. Найти вероятность того, что за время отсутствия рабочего ни один станок не потребует его внимания.

Р е ш е н и е. Требуется найти вероятность произведения двух событий AB, состоящего в том, что оба станка не потребуют внимания рабочего. Эти события независимы (так как станки работают независимо друг от друга), поэтому получим

$$P(AB) = P(A)\,P(B) = 0{,}7 \cdot 0{,}8 = 0{,}56.$$

Пример 1.5. На стенде независимо друг от друга проходят испытание в определенных условиях 250 приборов. Вероятность отказа в течение часа одинакова для всех приборов и равна 0,004. Найти вероятность того, что в течение часа откажет хотя бы один из испытываемых приборов.

Р е ш е н и е. Для каждого прибора вероятность того, что он не откажет, равна $1 - 0{,}004 = 0{,}996$. Вероятность того, что не откажет ни один из испытываемых приборов, равна $(0{,}996)^{250}$ (независимые события в совокупности). А вероятность того, что откажет по меньшей мере один из приборов, равна $1 - (0{,}996)^{250} \cong 5/8$.

Пример 1.6. Какова вероятность того, что выбранное наудачу изделие окажется первосортным, если известно, что 5 % всей продукции составляют нестандартные изделия, а 75 % стандартных изделий удовлетворяют требованиям, предъявляемым к первому сорту?

Р е ш е н и е. Пусть B – событие, заключающееся в том, что выбранное изделие стандартное, A – что изделие первосортное. Тогда

$$P(B) = 1 - 0{,}05 = 0{,}95, \quad P(A/B) = 0{,}75.$$

Искомая вероятность

$$P(A) = P(B)\,P(A/B) = 0{,}95 \cdot 0{,}75 \cong 0{,}71.$$

ЗАДАЧИ

1.16. Прибор, работающий в течение времени t, состоит из трех узлов, каждый из которых независимо от других может отказать в течение этого времени. Отказ хотя бы одного узла приводит к отказу прибора в целом. Вероятность безотказной работы первого узла равна 0,7, второго – 0,8 и третьего – 0,9. Найти вероятность безотказной работы прибора в целом.

1.17. Достаточным условием сдачи коллоквиума является ответ на один из двух вопросов, предлагаемых студенту преподавателем. Студент не знает ответов на 8 вопросов из тех 40, которые могут быть заданы ему. Какова вероятность того, что студент сдаст коллоквиум?

1.18. На сборку передано шесть узлов, каждый из которых с вероятностью 0,2 оказывается высокого качества. Найти вероятности следующих событий: A – два узла из шести высокого качества; B – не менее двух узлов из шести высокого качества.

1.19. Техническое устройство состоит из 5 узлов, работающих независимо друг от друга. Каждый узел во время эксплуатации отказывает независимо от других: если откажет более 3 узлов, устройство не сможет работать; если же откажет 1 или 2 узла, оно работает, но с меньшей эффективностью. Найти вероятность следующих событий: A – в устройстве не отказал ни один узел; B – устройство может работать; C – устройство работает с меньшей эффективностью.

1.20. Электрическая схема состоит из трех блоков, работающих независимо друг от друга. Вероятности того, что каждый из них работает исправно, соответственно равны: $P_1 = 0{,}8$, $P_2 = 0{,}4$, $P_3 = 0{,}7$. Схема годна к эксплуатации при наличии двух исправных блоков из трех. Определить вероятность того, что схема будет работать.

1.4. ФОРМУЛА ПОЛНОЙ ВЕРОЯТНОСТИ. ФОРМУЛА БЕЙЕСА

Пусть пространство Ω представлено в виде

$$\Omega = H_1 + H_2 + \ldots + H_n ,$$

где H_i – попарно несовместные (непересекающиеся) события ($H_i H_j = \varnothing$ при $i \neq j$). События H_i называются *гипотезами.*

Для произвольного события A имеет место соотношение

$$A = AH_1 + AH_2 + \ldots + AH_n ,$$

причем $(AH_i)(AH_j) = \varnothing$ для любых $i \neq j$. Согласно второй аксиоме вероятности,

$$P(A) = P(AH_1) + P(AH_2) + \ldots + P(AH_n),$$

откуда

$$\boxed{P(A) = \sum_{k=1}^{n} P(AH_k) = \sum_{k=1}^{n} P(H_k)P(A/H_k),} \qquad (1.3)$$

где $P(H_k) > 0$. Формулу (1.3) принято называть *формулой полной вероятности.*

Пусть для оценки правдоподобия гипотез H_i ставится эксперимент. Вероятности $P(H_i)$ гипотез до проведения опыта называются *априорными.* Полезность эксперимента определяется возможностью переоценить вероятности гипотез по известному результату A эксперимента, т.е. найти *апостериорные* (послеопытные) вероятности. Переоценить вероятность гипотез позволяет *формула Бейеса:*

$$\boxed{P(H_i /A) = P(H_i)P(A/H_i)/P(A).}$$

Эта формула является следствием записи условной вероятности в виде

$$P(H_i /A) = P(H_i A)/P(A),$$

где числитель $P(H_i A) = P(H_i)P(A/H_i)$. Заметим, что если $P(A/H_i) = P(A)$, то $P(H_i /A) = P(H_i)$, т.е. опыт бесполезен.

Пример 1.7. Прибор состоит из двух блоков, исправность каждого из которых необходима для функционирования прибора. Вероятность безотказной работы в течение заданного промежутка времени T для первого блока равна p_1, а для второго p_2. Прибор проходил испытание в течение времени T и вышел из строя. Найти вероятность того, что отказал: первый блок; второй блок; отказали оба блока, если работа одного из них не оказывает влияния на работу другого.

Р е ш е н и е. Пространство элементарных событий для первого блока имеет вид

$\Omega_1 = \{U_1, O_1\}$, где U_1 – исправность блока, а O_1 – отказ этого блока за время T. Так как по условию $P(U_1) = p_1$, то $P(O_1) = 1 - p_1$. Аналогично строится пространство элементарных событий $\Omega_2 = \{U_2, O_2\}$ с вероятностями $P(U_2) = p_2$, $P(O_2) = 1 - p_2$. Работу всего прибора можно описать с помощью прямого произведения пространств Ω_1 и Ω_2:

$$\Omega = \Omega_1\Omega_2 = \{U_1U_2, U_1O_2, U_2O_1, O_1O_2\},$$

где

$$P(U_1U_2) = P(U_1)P(U_2) = p_1p_2,$$
$$P(U_1O_2) = P(U_1)P(O_2) = p_1(1-p_2),$$
$$P(O_1U_2) = P(O_1)P(U_2) = (1-p_1)p_2,$$
$$P(O_1O_2) = P(O_1)P(O_2) = (1-p_1)(1-p_2).$$

Событие – выход прибора из строя – обозначим через A. Очевидно, что:

$$A = \{O_1U_2, O_2U_2, O_1O_2\}$$

и

$$P(A) = P(O_1U_2) + P(U_1O_2) + P(O_1O_2) =$$
$$= (1-p_1)p_2 + p_1(1-p_2) + (1-p_1)(1-p_2) = 1 - p_1p_2.$$

Введем следующие обозначения: $\{O_1U_2\} = H_1$, $\{U_1O_2\} = H_2$, $\{O_1O_2\} = H_{1,2}$ и определим вероятность того, что выход прибора из строя обусловлен неисправностью первого блока, т.е. найдем величину $P(H_1/A)$. По формуле Бейеса

$$P(H_i/A) = P(A/H_i)P(H_i)/P(A),$$

но $P(A/H_1) = 1$ (так как при неисправности первого блока прибор обязательно выходит из строя), $P(H_1) = P(O_1U_2) = (1-p_1)p_2$ и $P(A) = 1 - p_1p_2$. Подставив эти величины в формулу Бейеса, получим

$$P(H_1/A) = \frac{(1-p_1)p_2}{1-p_1p_2}.$$

Аналогично

$$P(H_2/A) = \frac{(1-p_2)p_1}{1-p_1p_2}, \quad P(H_{1,2}/A) = \frac{(1-p_1)(1-p_2)}{1-p_1p_2}$$

Пример 1.8. Станок обычно работает в двух режимах: рентабельном и нерентабельном. Рентабельный режим наблюдается в 90 % всех случаев работы, нерентабельный – в 10 %. Вероятность выхода станка из строя за время t работы в рентабельном режиме равна 0,1, в нерентабельном – 0,7. Найти вероятность выхода станка из строя за время t.

Р е ш е н и е. Событие – выход станка из строя – обозначим через A; H_1, H_2 – работа станка в рентабельном и нерентабельном режимах соответственно. Искомая вероятность

$$P(A) = P(H_1)P(A/H_1) + P(H_2)P(A/H_2) = 0{,}9 \cdot 0{,}1 + 0{,}1 \cdot 0{,}7 = 0{,}16.$$

ЗАДАЧИ

1.21. На сборку поступают шестерни с трех автоматов: первый дает 25 %, второй – 30 % и третий 45 % общего количества шестерен. Первый автомат допускает 0,1 % брака, второй – 0,2 %, третий – 0,3 %. Найти вероятность поступления на сборку бракованной шестерни и вероятность того, что оказавшаяся бракованной шестерня изготовлена первым автоматом.

1.22. Для контроля продукции из трех партий деталей взята одна деталь. Сколь велика вероятность обнаружения бракованной детали, если в одной партии 2/3 деталей бракованные, а в двух других – все годные?

1.23. Имеются две партии деталей по 12 и 10 штук, причем в каждой партии одно изделие бракованное. Изделие, взятое наудачу из первой партии, перекладывается во вторую, после чего выбирается наудачу изделие из второй партии. Определить вероятность извлечения бракованного изделия из второй партии.

1.24. Вероятность того, что изготовленная деталь окажется годной, равна 0,96. Деталь подвергается упрощенной системе контроля, которая для годных деталей дает положительный результат с вероятностью 0,98, а для деталей с отклонениями – с вероятностью лишь 0,05. Какова вероятность того, что изделие, выдержавшее упрощенный контроль, является годным?

1.25. Исследование на химический состав рудной породы указывает на присутствие в ней одного из трех элементов. Вероятности содержания каждого из них в породе соответственно равны: $p_1 = 1/2, p_2 = 1/6, p_3 = 1/3$. Для уточнения химического состава породы назначается очередной анализ, дающий положительный результат с вероятностью 0,1 в случае присутствия первого элемента, с вероятностью 0,2 – второго и с вероятностью 0,9 – третьего элемента. Анализ был произведен пять раз и дал четыре раза положительный результат и один раз отрицательный. Найти вероятность содержания в породе каждого элемента после проведения анализа.

1.26. Определить вероятность того, что 100 гаек, взятых наудачу из 1000, окажутся годными, если известно, что число негодных гаек на 1000 штук равновозможно от 0 до 5.

1.27. Определить вероятность того, что среди 1000 гаек нет ни одной неисправной, если из взятых наудачу 100 гаек все оказались исправными. Предполагается, что число неисправных гаек из 1000 равновозможно от 0 до 5.

1.28. Вероятности правильного определения химического состава проверяемых деталей на промежуточном контроле для каждого из трех контролеров соответственно равны: 4/5, 3/4, 2/3. При одновременном контроле тремя контролерами химический состав трех деталей оказался правильно определенным для двух деталей (что подтвердилось во время окончательного контроля). Найти вероятность того, что ошибку допустил третий контролер.

1.29. Деталь, одновременно подвергавшаяся обработке тремя инструментами, была признана негодной. Определить вероятность того, что деталь была признана негодной в результате обработки: первым; вторым; третьим инструментами, если вероятности неисправностей для них соответственно равны: 0,2, 0,4, 0,6.

1.5. ЭЛЕМЕНТЫ КОМБИНАТОРИКИ

Для подсчета числа возможных исходов того или иного случайного эксперимента необходимо применение основ комбинаторного анализа.

Пусть дано k групп элементов: в первой группе содержится n_1 элементов, во второй – n_2 элементов и т.д. Тогда можно образовать $n_1 n_2 \ldots n_k$ различных комбинаций, содержащих по одному элементу из каждой группы.

Пример 1.9. Из пункта A в пункт B ведут 4 дороги, из B в C – 7 дорог, а из C в D – 2 дороги. Сколько существует различных маршрутов из A в D?

Р е ш е н и е. Любой маршрут определяется выбором одной из дорог между каждыми двумя пунктами. Таким образом, имеем три группы дорог ($k = 3$), причем в первой группе 4 дороги ($n_1 = 4$), во второй 7 дорог ($n_2 = 7$), а в третьей 2 дороги ($n_3 = 2$). Общее число маршрутов определится формулой $n = n_1 n_2 n_3$, откуда $n = 4 \cdot 7 \cdot 2 = 56$.

В некоторых случаях роль каждой из k ($k < n$) групп может играть одно и то же множество, состоящее из n элементов (т.е. мы как бы располагаем k экземплярами одного и того же множества). Полное число комбинаций из k групп элементов в этом случае

$$\underbrace{n\,n \cdots n}_{k \text{ раз}} = n^k.$$

Эту же задачу можно рассматривать и с другой точки зрения: каждую комбинацию из k элементов (k-элементное подмножество) строить последовательно, выбирая элементы из исходного множества, содержащего n элементов, отмечая выбранные элементы и возвращая их в исходное множество. Такая процедура называется *выбором с возвращением*.

Итак, при выборе с возвращением из множества, состоящего из n элементов, можно извлечь ровно n^k различных подмножеств объемом k.

Существует n различных способов выбрать первый элемент из заданного множества. Если первый элемент выбран, то второй элемент можно выбрать уже из совокупности $n - 1$ элементов, так как рассматриваются подмножества, не содержащие повторяющихся элементов. При следующем выборе располагаем $n - 2$ элементами и т.д. Таким образом, число способов (комбинаций)

$$n_k = n(n-1)\,(n-2) \cdots (n-k+1).$$

Произвольное k-элементное подмножество n-элементного множества называется *сочетанием* из n элементов по k и обозначается C_n^k. Порядок элементов в подмножестве не имеет значения. Отличаются подмножества самими элементами. Число сочетаний (комбинаций) из n элементов по k

$$C_n^k = \frac{n!}{k!\,(n-k)!}.$$

Запись $n!$ (читается: "n-факториал") обозначает произведение всех натуральных чисел от 1 до n включительно. Принято считать, что $0! = 1$.

Пример 1.10. Сколькими способами читатель может выбрать три книги из пяти?

Р е ш е н и е. Искомое число способов равно числу трехэлементных подмножеств множества из 5 элементов: $C_5^3 = 5!/(3!\,2!) = 10$.

Пример 1.11. Сколькими способами из трех элементов множества можно составить подмножества, содержащие по два элемента?

Р е ш е н и е. Подмножества имеют вид: (a, b), (b, c), (a, c). Подмножества (a, b) и (b, c) определяют одно сочетание.

Множество называется *упорядоченным*, если каждому элементу этого множества поставлено в соответствие некоторое число (номер элемента) от 1 до n. Упорядоченные множества считаются различными, если они отличаются друг от друга либо своими элементами, либо их порядком. Различные упорядоченные множества, которые различаются лишь порядком элементов, называют *перестановками* этого множества и обозначают P_n: $P_n = n!$

Пример 1.12. Сколькими способами можно упорядочить множество, состоящее из трех элементов $\{a, b, c\}$?

Р е ш е н и е. Упорядоченные множества имеют вид: (a, b, c), (b, c, a), (c, b, a), (a, c, b), (b, a, c), (c, a, b). Число способов $P_3 = 3! = 1 \cdot 2 \cdot 3 = 6$.

Пример 1.13. Сколькими способами можно разместить на полке четыре книги?

Р е ш е н и е. Искомое число способов равно числу способов упорядочения множества, состоящего из 4 элементов, т.е. $P_4 = 1 \cdot 2 \cdot 3 \cdot 4 = 24$.

Упорядоченные k-элементные подмножества множества из n элементов называются ***размещениями*** из n элементов по k. Различные размещения из n по k, отличаются друг от друга самими элементами либо их порядком. Обозначается число размещений A_n^k:

$$A_n^k = \frac{n!}{(n-k)!} = k!\ C_n^k .$$

Пример 1.14. Сколькими способами из 3 элементов множества $\{a, b, c\}$ можно составить упорядоченное 2-элементное подмножество, характеризующееся как самими элементами, так и их порядком?

Р е ш е н и е. Подмножества имеют вид: (a, b), (b, c), (c, a), (b, a), (c, b), (a, c).

Пример 1.15. Имеется 25 мест. Сколькими способами можно разместить на них 4 учащихся?

Р е ш е н и е. Искомое число способов $A_{25}^4 = 25 \cdot 24 \cdot 23 \cdot 22 = 303\,600$.

1.6. ПОСЛЕДОВАТЕЛЬНОСТЬ НЕЗАВИСИМЫХ ИСПЫТАНИЙ В СХЕМЕ БЕРНУЛЛИ

Пусть произведено n независимых экспериментов (испытаний). В результате каждого из них может произойти либо не произойти событие A. Пусть событие A наступило в k испытаниях, а в $n-k$ наступило событие $\bar{A}$ (т.е. A не произошло). Тогда по теореме умножения независимых событий вероятность этого результата в n независимых испытаниях

$$P(\underbrace{AA\cdots A}_{k\text{ раз}} \cdot \underbrace{\bar{A}\cdot\bar{A}\cdots\bar{A}}_{n-k\text{ раз}}) = p^k q^{n-k} .$$

Но событие A может произойти в любом независимом эксперименте, и число способов наступления события A k раз в n испытаниях будет равно числу сочетаний из n элементов по k, т.е. C_n^k и все результаты несовместны. Тогда по теореме сложения вероятностей для несовместных событий вероятность наступления события A k раз в n испытаниях

$$\boxed{P_n(k) = C_n^k p^k q^{n-k},} \qquad (1.4)$$

где $q = 1 - p$; $C_n^k = \dfrac{n!}{k!(n-k)!}$ – биномиальные коэффициенты.

Пример 1.16. В скольких партиях с равным по силе противником выигрыш более вероятен:

1) в 3 партиях из 4 или в 5 из 8;

2) не менее, чем в 3 партиях из 4, и не менее, чем в 5 партиях из 8 (ничьи не учитывать)?

Р е ш е н и е. Назовем успехом в k-м испытании событие, состоящее в выигрыше k-й партии у противника. По условию вероятность успеха равна 1/2 (противник "равносильный").

1. Требуется сравнить следующие вероятности: $P_4(3)$ – три успеха в 4 испытаниях и $P_8(5)$ – пять успехов в 8 испытаниях. По формуле вероятностей заданного числа успехов в испытаниях Бернулли получаем:

$$P_4(3) = C_4^3(1/2)^3\ (1/2)^1 = C_4^3(1/2^4) = 8/32,$$

$$P_8(5) = C_8^5(1/2)^5\ (1/3)^3 = C_8^5(1/2^8) = 7/32,$$

$$P_4(3) > P_8(5).$$

2. Требуется сравнить следующие вероятности: $\sum\limits_{i=3}^{4} P_4(i)$ и $\sum\limits_{i=5}^{8} P_8(i)$. Имеем:

$$\sum_{i=3}^{4} = P_4(3) + P_4(4) = C_4^3\left(\frac{1}{2}\right)^4 + C_4^4\left(\frac{1}{2}\right)^4 = \frac{5}{16},$$

$$\sum_{i=5}^{8} = P_8(5) + P_8(6) + P_8(7) + P_8(8) = C_8^5\left(\frac{1}{2}\right)^8 +$$

$$+ C_8^6\left(\frac{1}{2}\right)^8 + C_8^7\left(\frac{1}{2}\right)^8 + C_8^8\left(\frac{1}{2}\right)^8 = \frac{93}{256}.$$

Следовательно, $\sum\limits_{i=5}^{8} P_8(i) > \sum\limits_{i=3}^{4} P_4(i)$.

Пример 1.17. Пусть вероятность попадания в "десятку" при одном выстреле равна 0,2. Определить наименьшее число независимых выстрелов, которые надо произвести, чтобы с вероятностью, не меньшей 0,9, попасть хотя бы один раз в "десятку".

Р е ш е н и е. По условию вероятность успеха $P = 0{,}2$. Пусть $P_n(0)$ – вероятность события, состоящего в том, что при n выстрелах не произошло ни одного попадания в "десятку". Тогда $1 - P_n(0) > 0{,}9$ или $1 - C_n^0(0{,}2)^0\ (0{,}8)^n > 0{,}9$; $(0{,}8)^n < 0{,}1$. Логарифмируя, получаем

$$n > \frac{1}{1 - \lg 8} = \frac{1}{1 - 0{,}903} \cong 10{,}3.$$

Итак, достаточно произвести 11 выстрелов.

Если испытания независимы, но вероятности наступления события A различны, т.е. в i-м испытании событие A появляется с вероятностью p_i, то вероятность $P_n(k)$ того, что событие A появится в n опытах ровно k раз, равна коэффициенту при z^k в разложении по степеням производящей функции $\varphi(z)$:

$$\varphi(z) = P_n(z) = \prod_{i=1}^{n} (q_i + p_i z), \tag{1.5}$$

где $q_i = 1 - p_i$. Если $p_1 = p_2 = ... = p_k = p$, то по формуле (1.5) получим

$$\prod_{i=1}^{n} (q_i + p_i z) = (q + pz)^n = \sum_{i=0}^{n} P_i(z) z^i ,$$

где выражение $P_i(z)$ определяется формулой (1.4).

Формула (1.5) применяется при нахождении работы в течение фиксированного интервала времени T некоторой физической системы, состоящей из k определенным образом соединенных элементов (деталей, узлов и т.п.). Надежность этой системы в узком смысле количественно определяется вероятностью безотказной работы в течение интервала времени T. Возникновение отказов системы вызывается отказами ее элементов, и, следовательно, надежность всей системы зависит от надежности ее элементов.

Различают два основных типа соединения элементов системы: последовательное и параллельное. Последовательное соединение приводит к тому, что отказ хотя бы одного элемента вызывает отказ всей системы. При параллельном соединении отказ системы происходит лишь в случае отказа всех параллельно соединенных элементов. Если события A_i и B_i означают безотказность работы i-го элемента в течение заданного промежутка времени соответственно при последовательном и параллельном соединениях, то вероятности безотказной работы системы будут равны:

$$P = \prod_{i=1}^{k} P(A_i),$$

$$P = \prod_{i=1}^{k} P(\overline{B}_i) = 1 - \prod_{i=1}^{k} (1 - P(B_i)),$$

где $\overline{B}_i$ — отказ в работе i-го элемента.

Пример 1.18. Основные неисправности коробки передач автомобиля таковы: произвольный выход шестерен из зацепления, затрудненное включение передач, шум шестерен при работе, поломка зубьев. Вероятности этих неисправностей различны и равны соответственно: $p_1 = 0{,}1$, $p_2 = 0{,}2$, $p_3 = 0{,}3$, $p_4 = 0{,}4$. Найти вероятности $P_4(0)$, $P_4(1)$, $P_4(2)$, $P_4(3)$, $P_4(4)$ (ни одной неисправности, одна, две, три, четыре неисправности), $R_4(2)$ (не менее двух неисправностей).

Р е ш е н и е. Согласно формуле (1.5), производящая функция

$$\varphi_4(z) = (0{,}9 + 0{,}1z)(0{,}8 + 0{,}2z)(0{,}7 + 0{,}3z)(0{,}6 + 0{,}4z).$$

Раскрывая скобки и приводя подобные члены, имеем

$$\varphi_4(z) = 0{,}302 + 0{,}460z + 0{,}205z^2 + 0{,}031z^3 + 0{,}002z^4,$$

откуда

$$P_4(0) = 0{,}302, P_4(1) = 0{,}460, P_4(2) = 0{,}205, P_4(3) = 0{,}031,$$
$$P_4(4) = 0{,}002, R_4(2) = 1 - P_4(0) - P_4(1) = 0{,}238.$$

1.7. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ В СХЕМЕ БЕРНУЛЛИ

Часто приходится вычислять вероятности различных событий при малых p, q или при больших n. В последнем случае удается заменить формулу (1.4) асимптотической *формулой Пуассона*. Если $n \to \infty$, $p \to 0$, $np = \lambda$, то вероятность наступления события m раз в n испытаниях стремится к величине

$\dfrac{\lambda^m}{m!} e^{-\lambda}$, т.е.

$$\boxed{P_n(m) = C_n^m p^m q^{n-m} \to P_m(\lambda) = \frac{\lambda^m}{m!} e^{-\lambda}.} \tag{1.6}$$

При малых q пуассоновским приближением можно воспользоваться для определения числа неудач.

Справедливы две следующие теоремы.

Локальная теорема Муавра–Лапласа. Пусть функции $\varphi(x)$ и $\Phi(x)$ определены следующими равенствами:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad \Phi(x) = \int_{-\infty}^{x} \varphi(u)\,du,$$

где $\Phi(x)$ – функция Лапласа. Если $n \to \infty$, а p $(0 < p < 1)$ постоянно, то вероятность того, что в n испытаниях событие наступит ровно m раз, определяется по формуле

$$P_n(m) = \frac{1}{\sqrt{npq}} \varphi(x), \tag{1.7}$$

где $\varphi(x) = \varphi((m - np)/\sqrt{npq})$.

Таблица значений функции $\varphi(x)$ для положительных x приведена в прил. 1. Функция $\varphi(x)$ – четная: $\varphi(-x) = \varphi(x)$.

Интегральная теорема Муавра–Лапласа. Если $n \to \infty$, а p $(0<p<1)$ постоянно, то вероятность того, что в n испытаниях событие произойдет от m_1 до m_2 раз, определяется по формуле

$$P_n(m_1, m_2) \cong \Phi(x_2) - \Phi(x_1), \tag{1.8}$$

где

$$x_1 = \frac{m_1 - np}{\sqrt{npq}}; \quad x_2 = \frac{m_2 - np}{\sqrt{npq}}.$$

Функция Лапласа – нечетная: $\Phi(-x) = -\Phi(x)$. Таблица значений функции $\Phi(x)$ для положительных x $(0 \leqslant x \leqslant 5)$ приведена в прил. 2; для $x > 5$ полагают $\Phi(x) = 0{,}5$.

Пример 1.19. В массовом производстве шестерен вероятность брака при штамповке равна 0,1. Какова вероятность того, что из 400 наугад взятых шестерен 50 будут бракованными?

Р е ш е н и е. Применив к данной задаче схему Бернулли, мы должны были бы подсчитать вероятность $P_{400}(50) = C_{400}^{50}p^{50}q^{350}$, где $p = 0,1$; $q = 0,9$. Однако это связано с громоздкими вычислениями. Поэтому используем локальную теорему Муавра–Лапласа. По формуле (1.6)

$$P_{400}(50) = \frac{1}{\sqrt{400\cdot 0,1\cdot 0,9}}\ \varphi\left(\frac{50 - 400\cdot 0,1}{\sqrt{400\cdot 0,1\cdot 0,9}}\right) = \frac{1}{6}\varphi\left(\frac{5}{3}\right) \cong 0,0989.$$

Пример 1.20. Найти вероятность того, что на 243-километровой трассе переключение передач (событие A) произойдет 70 раз, если вероятность такого переключения на каждом километре этой трассы равна 0,25.

Р е ш е н и е. По условию $n = 243$, $m = 70$, $p = 0,25$, $q = 0,75$. Так как $n = 243$ – достаточно большое число, воспользуемся локальной теоремой Муавра–Лапласа (см. формулу (1.7)). Тогда

$$P_n(m) = \frac{1}{\sqrt{npq}}\varphi(x),$$

где $x = \dfrac{m - np}{\sqrt{npq}}$. Вычислим значение x:

$$x = \frac{m - np}{\sqrt{npq}} = \frac{70 - 243\cdot 0,25}{\sqrt{243\cdot 0,25\cdot 0,75}} = \frac{9,25}{6,75} = 1,37.$$

По таблице (см. прил. 1) найдем $\varphi(1,37) = 0,1561$. Искомая вероятность

$$P_{243}(70) = \frac{1}{6,75}\cdot 0,1561 = 0,0231.$$

Пример 1.21. Вероятность появления события за время испытаний $p = 0,8$. Найти вероятность того, что событие появится: 1) не менее 75 раз и не более 90 раз; 2) не менее 75 раз; 3) не более 74 раз при 100 испытаниях.

Р е ш е н и е. 1. Воспользуемся интегральной теоремой Муавра–Лапласа. По формуле (1.7)

$$P_n(m_1, m_2) = \Phi(x_2) - \Phi(x_1),$$

где $\Phi(x)$ – функция Лапласа. Вычислим:

$$x_1 = \frac{75 - 100\cdot 0,8}{100\cdot 0,8\cdot 0,2} = -1,25,\quad x_2 = \frac{90 - 100\cdot 0,8}{100\cdot 0,8\cdot 0,2} = 2,5.$$

Учитывая, что функция Лапласа – нечетная, т.е. $\Phi(-x) = -\Phi(x)$, получаем

$$P_{100}(75;\ 90) = \Phi(5) - \Phi(-1,25) = \Phi(5) + \Phi(1,25).$$

По таблице (см. прил. 2) найдем: $\Phi(5) = 0,499$, $\Phi(1,25) = 0,3944$. Искомая вероятность

$$P_{100}(75;\ 90) = 0,4998 + 0,3944 = 0,888.$$

2. Требование, чтобы событие появилось не менее 75 раз, означает, что число появлений события может быть равно 75, 76,..., 100. Таким образом, в рассматриваемом случае следует принять $m_1 = 75, m_2 = 100$. Тогда:

$$x_1 = \frac{75 - 100 \cdot 0{,}8}{100 \cdot 0{,}8 \cdot 0{,}2} = -1{,}25, \quad x_2 = \frac{100 - 100 \cdot 0{,}8}{100 \cdot 0{,}8 \cdot 0{,}2} = 5.$$

По таблице (см. прил. 2) находим: $\Phi(125) = 0{,}394$, $\Phi(5) = 0{,}5$. Искомая вероятность

$$P_{100}(75;\ 100) = \Phi(5) - \Phi(-1{,}25) = 0{,}5 + 0{,}394 = 0{,}894.$$

3. События "A наступило не менее 75 раз" и "A наступило не более 74 раз" противоположны, поэтому сумма вероятностей этих событий равна единице. Следовательно, искомая вероятность

$$P_{100}(0;\ 74) = 1 - P_{100}(75;\ 100) = 1 - 0{,}894 = 0{,}106.$$

ЗАДАЧИ

1.30. Имеется 100 станков равной мощности, работающих независимо друг от друга в одинаковом режиме при включенном приводе в течение 0,8 всего рабочего времени. Какова вероятность того, что в произвольный момент времени окажутся включенными от 70 до 86 станков?

1.31. Вероятность выхода из строя за время T одного конденсатора равна 0,2. Определить вероятность того, что за время T из 100 конденсаторов, работающих независимо, выйдут из строя: 1) не менее 20 конденсаторов; 2) менее 28 конденсаторов; 3) от 14 до 26 конденсаторов.

1.32. Вероятность изготовления деталей номинальных размеров равна 0,51. Найти вероятность того, что среди 100 деталей окажется 50 деталей номинальных размеров.

1.33. Вероятность появления события в каждом из 21 независимых испытаний равна 0,7. Найти вероятность того, что событие появится в большинстве испытаний.

1.34. Вероятность появления события в каждом из независимых испытаний равна 0,8. Сколько испытаний нужно провести, чтобы событие появилось не менее 75 раз с вероятностью 0,9?

1.35. Вероятность изготовления годной детали на токарном станке равна 0,9. Сколько деталей нужно обработать, чтобы с вероятностью 0,98 можно было ожидать, что не менее 150 деталей окажутся годными?

2. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

2.1. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ. РАСПРЕДЕЛЕНИЕ СЛУЧАЙНЫХ ВЕЛИЧИН

Случайной величиной (СВ) называется величина X, которая в результате опыта может принимать то или иное значение, причем заранее (до опыта) неизвестное. В теоретико-вероятностной схеме случайной величиной называют числовую измеримую функцию от элементарных исходов: $X = X(\omega)$, $\omega \in \Omega$. Следовательно, можно говорить о числовой функции $X(\omega)$, описывающей результаты эксперимента и определенной на пространстве элементарных событий. Такие числовые функции называют случайными величинами и обозначают прописными буквами латинского алфавита $X, Y, Z, \ldots$, а их значения – строчными $x, y, z, \ldots$

Дискретной СВ называется величина $X = X(\omega)$, которая в зависимости от элементарных исходов ω принимает конечное или бесконечное счетное множество значений.

Случайная величина считается заданной, если известен закон распределения, который может принимать различные формы.

Законом распределения дискретной СВ называется всякое соотношение, устанавливающее связь между возможными значениями СВ и соответствующими им вероятностями.

Закон распределения СВ можно задать таблицей, графиком или аналитическим выражением.

1. При *табличном задании дискретной СВ* имеем ряд распределения. *Рядом распределения дискретной СВ* X называется таблица, в которой перечислены всевозможные значения $x_1, x_2, \ldots, x_n, \ldots$ СВ X и соответствующие им вероятности $p_1, p_2, \ldots, p_n, \ldots$:

$$p_i = P(X = x_i), \quad \sum_{i=1}^{n} p_i = 1 \quad \left(\sum_{i=1}^{\infty} p_i = 1\right).$$

Ряд распределения в случае конечного пространства Ω записывается следующим образом:

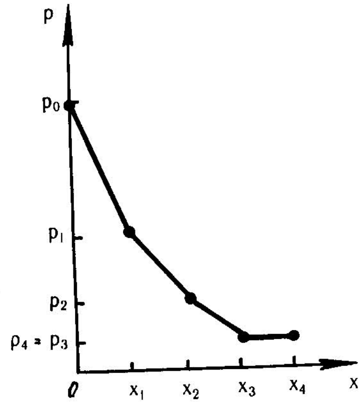
x_i	x_1	x_2	...	x_n
p_i	p_1	p_2	...	p_n

2. В случае *графического задания дискретной СВ* по оси Ox откладываются значения этой СВ, а по оси Oy – соответствующие им вероятности; ломаная линия, соединяющая точки (x_i, p_i), называется *многоугольником распределения.*

3. При *аналитическом задании дискретной СВ*

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

где $F(x)$ – *функция распределения дискретной СВ.*



Р и с. 2.1

Непрерывной СВ называется такая СВ X, для которой существует неотрицательная функция $f(x)$, удовлетворяющая при любом x равенству

$$F(x) = \int_{-\infty}^{x} f(x)\,dx .$$

Непрерывную СВ можно задать графически или аналитически. График непрерывной СВ называют *кривой распределения.*

Пример 2.1. На пути движения автомашины четыре перекрестка, регулируемых светофорами. Каждый из светофоров с вероятностью 0,5 либо разрешает, либо запрещает автомашине дальнейшее движение. Построить многоугольник распределения вероятностей СВ, выражающей число перекрестков, пройденных автомашиной без остановки.

Р е ш е н и е. Пусть СВ X – число перекрестков, пройденных автомашиной без остановки. Эта СВ может принимать следующие значения: $x_0 = 0$, $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $x_4 = 4$. Вероятности $p_i = P(X = x_i)$ того, что число пройденных перекрестков X будет равно данному частному значению, определяются по формуле

$$p_i = P(X=x_i) = \begin{cases} p(1-p)^i & \text{при } i = 0,1,2,3, \\ (1-p)^4 & \text{при } i = 4 , \end{cases}$$

где p – вероятность того, что машина будет задержана данным светофором ($p = 0,5$). В результате вычисления найдем: $p_0 = 0,5$, $p_1 = 0,25$, $p_2 = 0,125$, $p_3 = 0,0625$, $p_4 = 0,0625$. По полученным данным построим многоугольник распределения вероятностей (рис. 2.1).

Пример 2.2. Производится ряд независимых опытов, в результате каждого из которых возможно наступление некоторого события A. Вероятность появления события A в каждом опыте равна p. Опыты производятся до первого появления события A, после чего прекращаются. Случайная величина X – число произведенных опытов. Построить ряд распределения этой СВ.

Р е ш е н и е. Возможные значения СВ X: $x_1 = 1, x_2 = 2, \ldots, x_k = k$ ... Найдем их вероятности. Пусть p_1 – вероятность того, что при первом опыте наступит событие A: $p_1 = p$. Пусть, далее, p_2 – вероятность того, что будет произведено два опыта; для этого нужно, чтобы в первом опыте событие A не появилось, во втором – появилось. Вероятность этого события $p_2 = (1-p)p$. Аналогично $p_3 = (1-p)^2 p, \ldots, p_k = (1-p)^{k-1} p$. Запишем ряд распределения СВ X:

x_i	1	2	3	...	k	...
p_i	p	qp	$q^2 p$	...	$q^{k-1} p$	...

Заметим, что вероятности p_k образуют геометрическую прогрессию со знаменателем $1 - p = q$, $q < 1$, поэтому такое распределение называется *геометрическим.*

ЗАДАЧИ

2.1. Контролируемая деталь подвергается измерениям в трех плоскостях, причем каждое измерение выявляет ее годность с вероятностью 0,5. Случайная величина X – количество годных деталей. Построить: 1) ряд распределения X; 2) многоугольник распределения X.

2.2. Производятся последовательные независимые испытания пяти приборов на надежность. Каждый следующий прибор испытывается в том случае, если предыдущий оказался надежным. Построить ряд распределения числа испытанных приборов, если вероятность выдержать испытание для каждого из них равна 0,9.

2.3. Независимые опыты продолжаются до первого положительного исхода. Для случайного числа опытов m найти (вероятность положительного исхода при каждом опыте считать равной 0,5): 1) ряд распределения; 2) многоугольник распределения; 3) наивероятнейшее число опытов.

2.4. Производятся испытания n изделий на надежность, причем вероятность выдержать испытания для каждого изделия равна p. Найти ряд распределения случайного числа изделий, выдержавших испытания.

2.5. В партии из 50 деталей имеется 10 нестандартных. Наудачу отобрано 4 детали. Построить ряд распределения числа стандартных деталей среди отобранных.

2.2. ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ И ПЛОТНОСТЬ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ НЕПРЕРЫВНОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Функцией распределения (интегральной функцией распределения) СВ X называется функция $F(x)$, значения которой равны вероятностям $P(X < x)$ того, что СВ X примет значение, меньшее произвольно выбранного значения x, т.е. $X \in \,]-\infty; x[$:

$$F(x) = P(X < x) = P(-\infty < X < x).$$

Для дискретных СВ $F(x)$ вычисляется по формуле

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

где суммирование ведется по всем значениям i, для которых $x_i < x$.

Приведем *свойства функции распределения.*

1. ***Функция распределения $F(x)$ является неотрицательной, заключенной между нулем и единицей:***

$$0 \leqslant F(x) \leqslant 1.$$

2. *Функция распределения $F(x)$ является неубывающей, т.е. при $b > a$*

$F(b) \geqslant F(a)$.

3. *Вероятность появления непрерывной СВ в интервале $[a; b[$ равна приращению функции распределения $F(x)$ на этом интервале:*

$$P(a \leqslant x < b) = F(b) - F(a).$$

4. *Если все возможные значения СВ X принадлежат интервалу $[a; b[$, то*

$$F(x) = \begin{cases} 0 \text{ при } x < a, \\ 1 \text{ при } x > b. \end{cases}$$

Если же все возможные значения СВ X принадлежат всей числовой оси, то

$$\lim_{x \to -\infty} F(x) = 0, \qquad \lim_{x \to +\infty} F(x) = 1.$$

Плотностью распределения вероятностей $f(x)$ (дифференциальной функцией распределения) СВ X называют первую производную от функции распределения, т.е.

$$f(x) = F'(x).$$

График плотности распределения вероятностей $f(x)$ называют *кривой плотности распределения вероятностей.*

Перечислим *свойства плотности распределения вероятностей* СВ X.

1. *Плотность распределения вероятностей неотрицательна:*

$$f(x) \geqslant 0.$$

2. *Интеграл в бесконечных пределах от плотности распределения вероятностей равен единице:*

$$\int_{-\infty}^{\infty} f(x)\,dx = 1.$$

Вероятность того, что непрерывная СВ X принадлежит интервалу $]a; b[$

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

3. *Функция распределения $F(x)$ выражается через плотность распределения вероятностей $f(x)$ формулой*

$$F(x) = \int_{-\infty}^{x} f(u)\,du.$$

Пример 2.3. Плотность распределения вероятностей СВ $f(x) = ax^2 e^{-kx}$ ($k > 0$, $0 < x < \infty$). Найти коэффициент a и функцию распределения СВ X. Вычислить вероятность попадания СВ X в интервал $]0; 1/k[$.

Р е ш е н и е. Коэффициент a определяется с помощью равенства

$$\int_0^{\infty} a x^2 e^{-kx} dx = 1,$$

откуда

$$a = \frac{1}{\int_0^{\infty} x^2 e^{-kx} dx}$$

Интегрируя по частям, получаем

$$\int_0^{\infty} x^2 e^{-kx} dx = \frac{2}{k^3}.$$

Следовательно, $a = k^3/2$ и плотность распределения вероятностей имеет вид

$$f(x) = \frac{k^3}{2} x^2 e^{-kx}$$

Функция распределения $F(x)$ СВ X определяется так:

$$F(x) = \int_0^{x} \frac{k^3}{2} x^2 e^{-kx} dx = 1 - \frac{k^2 x^2 + 2kx + 2}{2} e^{-kx}$$

Вероятность $P(0 < X < 1/k)$ попадания СВ X в заданный промежуток

$$P(0 < X < 1/k) = F(1/k) - F(0) = 1 - 5/(2e) = 0{,}086.$$

Пример 2.4. Известно, что СВ X имеет следующую функцию распределения:

$$F(x) = \begin{cases} 0 & \text{при } x < 0, \\ x^2/16 & \text{при } 0 < x < 2, \\ x - 7/4 & \text{при } 2 < x < 11/4, \\ 1 & \text{при } x > 11/4. \end{cases}$$

Вычислить плотность распределения вероятностей $f(x)$ СВ X. Найти вероятность попадания СВ X на отрезок $[1; 1{,}5]$.

Р е ш е н и е. Для отыскания $f(x)$ воспользуемся равенством $f(x) = F(x)$:

$$f(x) = \begin{cases} 0 & \text{при } x < 0, \\ x/8 & \text{при } 0 < x < 2, \\ 1 & \text{при } 2 < x < 11/4, \\ 0 & \text{при } x > 11/4. \end{cases}$$

Искомая вероятность

$$P(1 < X < 1{,}5) = F(1{,}5) - F(1) = 0{,}4332 - 0{,}3413 = 0{,}0919.$$

Пример 2.5. Функция распределения СВ X имеет вид $F(x) = A - B\,\mathrm{arctg}\,x$. Определить постоянные A, B и найти плотность распределения вероятностей $f(x)$.

Р е ш е н и е. Воспользуемся свойством 4 функции распределения:

$$\lim_{x \to -\infty} (A + B\operatorname{arctg} x) = 0, \quad \lim_{x \to +\infty} (A + B\operatorname{arctg} x) = 1.$$

Получили $A + B \cdot (-\frac{\pi}{2}) = 0$, $A + B \cdot \frac{\pi}{2} = 1$, откуда $A = 1/2, B = 1/\pi$ и, значит,

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} x .$$

Теперь нетрудно найти $f(x)$:

$$f(x) = F'(x) = \frac{1}{\pi(1 + x^2)} .$$

Пример 2.6. Известно, что СВ X имеет следующую плотность распределения вероятностей:

$$f(x) = \begin{cases} 0 & \text{при } x < 1, \\ A/x^2 & \text{при } x \geqslant 1. \end{cases}$$

Определить: 1) коэффициент A; 2) функцию распределения $F(x)$; 3) вероятность $P(2 < X < 3)$ попадания СВ X на отрезок $[2; 3]$; 4) вероятность того, что при четырех независимых испытаниях СВ X ни разу не попадет на отрезок $[2; 3]$.

Р е ш е н и е. 1. Для нахождения коэффициента A воспользуемся свойством 2 плотности распределения вероятностей. Так как

$$\int_1^\infty f(x)\,dx = \int_1^\infty \frac{A}{x^2}\,dx = -\lim_{a \to \infty} \frac{A}{x}\Bigg|_1^a = A ,$$

то $A = 1$.

2. Находим $F(x)$:

$$F(x) = \begin{cases} \int\limits_{-\infty}^{x} 0\,dx = 0 & \text{при } x < 1, \\ \int\limits_1^x \frac{dx}{x^2} = -\frac{1}{x}\Bigg|_1^x = \frac{x-1}{x} & \text{при } x \geqslant 1. \end{cases}$$

3. Искомая вероятность $P(2 < X < 3) = F(3) - F(2) = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}$.

4. Вероятность того, что при одном испытании СВ X не попадет в интервал $[2; 3]$, равна $1 - 1/6 = 5/6$, а при четырех испытаниях — $(5/6)^4 = 0{,}48$.

ЗАДАЧИ

2.6. Функция распределения непрерывной СВ X задана следующим образом:

$$F(x) = \begin{cases} 0 & \text{при } x \leqslant 0, \\ \frac{1}{2} - \frac{1}{2}\cos x & \text{при } 0 < x \leqslant \pi, \\ 1 & \text{при } x > \pi \end{cases}$$

Определить плотность распределения вероятностей $f(x)$ СВ X. Построить графики функций $F(x)$ и $f(x)$.

2.7. Дана функция $f(x) = Ax^2 e^{-2x}$ $(0 < x < +\infty)$. Определить, при каком значении A функция $f(x)$ будет являться плотностью распределения вероятностей некоторой непрерывной СВ X. Найти функцию распределения $F(x)$ и вероятность попадания СВ X в интервал $[0;\ 1/2]$.

2.8. Дана функция распределения СВ X

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt.$$

Найти плотность распределения вероятностей СВ X.

2.9. Функция распределения случайного времени безотказной работы радиоаппаратуры имеет вид (*экспоненциальный закон распределения*) $F(t) = 1 - e^{-t/T}$ $(t \geqslant 0)$. Найти: 1) вероятность безотказной работы аппаратуры в течение времени T; 2) плотность распределения вероятностей.

2.10. При каком значении a функция $f(x) = a/(1+x^2)$ $(-\infty < x < +\infty)$ является плотностью распределения вероятностей СВ X? Найти: 1) функцию распределения СВ X; 2) вероятность попадания СВ X в интервал $]-1;\ 1[$.

2.11. Шкала секундомера имеет цену деления 0,2 с. Какова вероятность отсчета времени с ошибкой более 0,05 с по этому секундомеру, если отсчет выполняется с точностью до целого деления с округлением в ближайшую сторону?

2.12. Азимутальный лимб имеет цену деления 1°. Какова вероятность сделать ошибку в пределах $\pm 10'$ при измерении азимутального угла, если результат отсчета округляется до ближайшего целого числа градусов?

2.3. ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН

Математическим ожиданием СВ X называется число, обозначаемое $M(X)$ и определяемое по формуле

$$M(X) = \begin{cases} \sum\limits_i x_i p_i & \text{для дискретной СВ } X, \\ \int\limits_{-\infty}^{\infty} x f(x)\, dx & \text{для непрерывной СВ } X, \end{cases} \qquad (2.1)$$

где $p_i = P(X = x_i)$; $f(x)$ – плотность распределения вероятностей. Предполагается, что ряд $\sum\limits_i x_i p_i$ и несобственный интеграл $\int\limits_{-\infty}^{\infty} x f(x) dx$ абсолютно сходятся, в противном случае считают, что математическое ожидание СВ X не существует.

Перечислим *свойства математического ожидания*.

1. *Математическое ожидание постоянной величины C постоянно, т.е. $M(C) = C$, так как единственному значению C соответствует вероятность $p = 1$. По определению $M(C) = C \cdot 1 = C$.*

2. *Если C – постоянная, то $M(CX) = CM(X)$.*

3. *Для любых* СВ $X_1, X_2, \ldots, X_n$

$$M(X_1 + X_2 + \ldots + X_n) = M(X_1) + M(X_2) + \ldots + M(X_n).$$

4. *Математическое ожидание произведения независимых СВ равно произведению математических ожиданий сомножителей:*

$$M(X_1 X_2 \cdots X_n) = M(X_1)\, M(X_2) \cdots M(X_n).$$

Дисперсией $D(X)$ *СВ* X называют математическое ожидание квадрата отклонения СВ X от ее математического ожидания:

$$D = M(X - M(X))^2. \tag{2.2}$$

Дисперсия СВ X вычисляется по формуле

$$D(X) = \begin{cases} \sum\limits_i (x_i - m_x)^2 p_i & \text{для дискретной СВ } X, \\ \int\limits_{-\infty}^{\infty} (x - m_x)^2 f(x)\,dx & \text{для непрерывной СВ } X, \end{cases}$$

где $m_x = M(X)$.

Формулу (2.2) можно преобразовать к виду

$$D(X) = \begin{cases} \sum\limits_i x^2 p_i - M^2(X) & \text{для дискретной СВ } X, \\ \int\limits_{-\infty}^{\infty} x^2 f(x)\,dx - M^2(x) & \text{для непрерывной СВ } X. \end{cases}$$

Приведем *основные свойства дисперсии.*

1. *Для любой СВ* X $D(X) > 0$.
2. $D(C) = 0$.
3. $D(CX) = C^2 D(X)$, *где* C *– постоянная.*
4. $D(X_1 + X_2) = D(X_1) + D(X_2) \pm 2M((X_1 - m_{x_1})(X_2 - m_{x_2}))$.
5. *Если СВ* $X_1, X_2, \ldots, X_n$ *независимы, то*

$$D(X_1 + X_2 + \ldots + X_n) = D(X_1) + D(X_2) + \ldots + D(X_n).$$

В случае двух СВ X_1 и X_2 имеем $D(X_1 + X_2) = D(X_1) + D(X_2)$.

6. *Дисперсия числа появления события в n независимых испытаниях, при которых вероятность наступления события p постоянна, равна произведению числа испытаний на вероятность появления p и непоявления q события, т.е.* $D(X) = npq$.

Положительное значение квадратного корня из дисперсии СВ X называется ее *средним квадратичным отклонением* и обозначается σ_X: $\sigma_X = \sqrt{D(X)}$.

Модой дискретной СВ называется ее наиболее вероятное значение. *Модой непрерывной СВ* называется такое ее значение M_0, при котором плотность распределения вероятностей имеет максимум (рис. 2.2).

Медианой M_d *произвольной СВ* называется такое ее значение, относительно которого равновероятно получение большего или меньшего значения СВ (см. рис. 2.2).

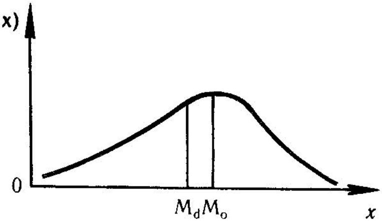


Рис. 2.2

Пример 2.7. Пусть X – случайная величина, ряд распределения которой имеет вид

x_i	0	1
p_i	q	p

где $q = 1 - p$. Найти математическое ожидание и дисперсию СВ X.

Решение. Так как X принимает только два значения: $x_1 = 0$ и $x_2 = 1$ с вероятностями $p_1 = 1 - p$ и $p_2 = p$ соответственно, то по определению математическое ожидание

$$M(X) = x_1 p_1 + x_2 p_2 = 0 \cdot (1-p) + 1 \cdot p = p.$$

Для отыскания дисперсии СВ X воспользуемся формулой (2.2):

$$D(X) = M(X^2) - (M(X))^2 = p - p^2 = p(1-p).$$

Пример 2.8. Из партии численностью 25 изделий, среди которых имеется 6 нестандартных, выбраны случайным образом для проверки их качества 3 изделия. Найти математическое ожидание и среднее квадратичное отклонение СВ X нестандартных изделий, содержащихся в выборке.

Решение. По условию задачи СВ X принимает значения $x_1 = 0, x_2 = 1, x_3 = 2$, $x_4 = 3$. Вероятность того, что в этой выборке окажется ровно k ($k = 0, 1, 2, 3$) нестандартных изделий, вычисляется по формуле

$$P(X - x_k) = C_6^k C_{19}^{3-k} / C_{25}^3,$$

откуда $p_1 = 0{,}41, p_2 = 0{,}43, p_3 = 0{,}11$, $p_4 = 0{,}05$. По формуле (2.1) $M(X) = 0 \cdot 0{,}41 + 1 \cdot 0{,}43 + 2 \cdot 0{,}11 + 3 \cdot 0{,}05 = 0{,}80$. Дисперсию $D(X)$ найдем по формуле (2.2). Для этого запишем ряд распределения X^2:

x_i^2	0	1	4	9
p_i	0,41	0,43	0,11	0,05

Тогда

$$M(X^2) = 0{,}041 + 1 \cdot 0{,}43 + 4 \cdot 0{,}11 + 9 \cdot 0{,}05 = 1{,}32,$$

$$D(X) = M(X^2) - M^2(X) \doteq 1{,}32 - (0{,}8)^2 = 0{,}68.$$

Так как $\sigma(X) = \sqrt{D(X)}$, то $\sigma(X) = \sqrt{0{,}68} \cong 0{,}82$.

Пример 2.9. При сборке прибора для наиболее точной подгонки основной детали может потребоваться (в зависимости от удачи) 1, 2, 3, 4, 5 проб с вероятностями 0,07, 0,21, 0,55, 0,16, 0,01 соответственно. Требуется обеспечить сборщика необходимым количеством деталей для сборки 30 приборов. Сколько деталей надо отпустить сборщику?

Р е ш е н и е. Число проб, необходимых для достижения удовлетворительной сборки прибора, есть СВ X, ряд распределения которой имеет вид

x_i	1	2	3	4	5
p_i	0,07	0,21	0,55	0,16	0,01

Среднее число проб, необходимых для сборки одного прибора, равно $M(X)$. Тогда

$$30\,M(X) = 30\,(1\cdot 0{,}07 + 2\cdot 0{,}21 + 3\cdot 0{,}55 + 4\cdot 0{,}16 + 5\cdot 0{,}11) = 30\cdot 2{,}83 \cong 85.$$

Следовательно, для сборки 30 приборов необходимо 85 деталей.

Пример 2.10. Кривая распределения СВ X представляет собой полуэллипс с полуосями a и b. Полуось a известна. Определить b. Найти $M(X)$, $D(X)$ и функцию распределения $F(x)$.

Р е ш е н и е. Полуось b находится из условия равенства единице площади, ограниченной кривой распределения:

$$\pi ab/2 = 1, \quad b = 2/(\pi a).$$

Плотность распределения вероятностей

$$f(x) = \begin{cases} \dfrac{b}{a}\sqrt{a^2 - x^2} & \text{при} \quad -a \leq x \leq a, \\ 0 & \text{при } x < -a \text{ или } x > a. \end{cases}$$

Математическое ожидание

$$M(X) = \int_{-a}^{a} x \frac{2}{\pi a^2}\sqrt{a^2 - x^2}\,dx = 0.$$

Дисперсия

$$D(X) = \int_{-a}^{a} x^2 \frac{2}{\pi a^2}\sqrt{a^2 - x^2}\,dx = \frac{a^2}{4}.$$

Функция распределения

$$F(x) = \int_{-\infty}^{x} f(u)\,du = \begin{cases} 0 & \text{при } x < -a, \\ \dfrac{1}{\pi a^2}\left(x\sqrt{a^2 - x^2} + a^2 \arcsin\dfrac{x}{a} + \dfrac{a^2\pi}{2}\right) & \text{при } -a \leq x \leq a, \\ 1 & \text{при } x > a. \end{cases}$$

Обобщением основных числовых характеристик случайных величин является понятие моментов СВ.

Начальным моментом k-го порядка СВ X называют математическое ожидание СВ X^k:

$$\nu_k = M(X^k) = \begin{cases} \sum_i x_i^k p_i & \text{для дискретной СВ } X, \\ \int_{-\infty}^{\infty} x^k f(x)\,dx & \text{для непрерывной СВ } X. \end{cases}$$

Начальный момент первого порядка ($k=1$) является математическим ожиданием: $\nu_1 = M(X)$.

Центральным моментом k-го порядка СВ X называют математическое ожидание СВ $(X - M(X))^k$: $\mu_k = M(X - M(X))^k$, вычисляемое по формуле

$$\mu_k = \begin{cases} \sum_i (x_i - m_x)^k p_i & \text{для дискретной СВ}, \\ \int_{-\infty}^{\infty} (x - m_x)^k f(x)\,dx & \text{для непрерывной СВ}. \end{cases}$$

Центральный момент первого порядка всегда равен нулю; центральный момент второго порядка является дисперсией СВ X.

Пример 2.11. Известно, что СВ X принимает значения 3 и 5 с вероятностями 0,2 и 0,8. Найти центральные моменты до третьего порядка включительно.

Р е ш е н и е. Найдем начальные моменты, а затем выразим через них центральные. Имеем:

$$\nu_1 = 3 \cdot 0{,}2 + 5 \cdot 0{,}8 = 4{,}6, \quad \nu_2 = 9 \cdot 0{,}2 + 25 \cdot 0{,}8 = 21{,}8,$$

$$\nu_3 = 27 \cdot 0{,}2 + 125 \cdot 0{,}8 = 105{,}4;$$

$$\mu_1 = 0, \ \mu_2 = \nu_2 - \nu_1^2 = 21{,}8 - 21{,}16 = 0{,}64,$$

$$\mu_3 = 105{,}4 - 3 \cdot 4{,}6 \cdot 21{,}8 + 2 \cdot 4{,}6^3 = -0{,}768.$$

Пример 2.12. Плотность распределения вероятностей СВ X имеет вид

$$f(x) = \begin{cases} 0 & \text{при } x < 0, \\ 2x & \text{при } 0 < x < 1, \\ 0 & \text{при } x > 1, \end{cases}$$

Найти начальные и центральные моменты до третьего порядка включительно.

Р е ш е н и е. Как и в предыдущей задаче, вначале найдем начальные моменты, а затем выразим через них центральные. Имеем:

$$\nu_1 = \int_0^1 x \cdot 2x\,dx = \frac{2x^3}{3}\Big|_0^1 = \frac{2}{3},$$

$$\nu_2 = \int_0^1 x^2 \cdot 2x\,dx = \frac{x^4}{2}\Big|_0^1 = \frac{1}{2},$$

$$\nu_3 = \int\limits_0^1 x^3 \cdot 2x\,dx = \frac{2x^5}{5}\Big|_0^1 = \frac{2}{5};$$

$$\mu_1 = 0, \mu_2 = 1/2 - 4/9 = 1/18,$$

$$\mu_3 = \frac{2}{5} - 3 \cdot \frac{2}{3} \cdot \frac{1}{2} + 2\left(\frac{2}{3}\right)^3 = \frac{2}{5} - 1 + \frac{16}{27} = -\frac{1}{135}.$$

ЗАДАЧИ

2.13. Автомобиль должен проехать по улице, на которой установлены три светофора, дающие независимо друг от друга зеленый сигнал в течение 1,5 мин, желтый – в течение 0,3 мин, красный – в течение 1,2 мин. Найти математическое ожидание, дисперсию и среднее квадратичное отклонение СВ X числа остановок автомобиля.

2.14. На полигоне производится стрельба из орудия по цели. Всякий раз стрельба ведется до первого попадания в цель. Вероятность попадания при одном выстреле $p = 0{,}6$. Каков средний ожидаемый расход снарядов, если производится серия из 100 стрельб, условия которых одинаковы для каждой стрельбы?

2.15. Известно, что размер шарика для подшипников является СВ X. При контроле бракуются все шарики, диаметр которых отличается от номинального больше, чем на 0,1 мм. При этом известно, что средний размер диаметра шарика равен $m_d = (d_1 + d_2)/2$, а брак составляет 10 % всего выпуска. Определить среднее квадратичное отклонение $\sigma(X)$ диаметра шарика.

2.16. Функция распределения СВ X имеет вид

$$F(x) = \begin{cases} 0 & \text{при } x < -1, \\ a + b\arcsin x & \text{при } -1 < x < 1, \\ 1 & \text{при } x > 1. \end{cases}$$

Определить постоянные a и b. Найти $M(X)$ и $D(X)$.

2.17. Плотность распределения вероятностей СВ X имеет вид

$$f(x) = \begin{cases} 0 & \text{при } x < -a, \\ 1/(\pi\sqrt{a^2 - x^2}) & \text{при } -a < x < a, \\ 0 & \text{при } x > a. \end{cases}$$

Определить дисперсию и среднее квадратичное отклонение СВ X.

2.18. Плотность распределения вероятностей СВ X задана в виде

$$f(x) = \begin{cases} 0 & \text{при } x < 0, \\ \dfrac{x^m}{m!} e^{-x} & \text{при } x > 0. \end{cases}$$

Определить $M(X)$ и $D(X)$.

2.19. Функция распределения СВ X имеет вид

$$F(x) = \begin{cases} 1 - x_0^2/3 & \text{при } x > x_0 \quad (x_0 > 0), \\ 0 & \text{при } x < x_0. \end{cases}$$

Найти $M(X)$ и $D(X)$.

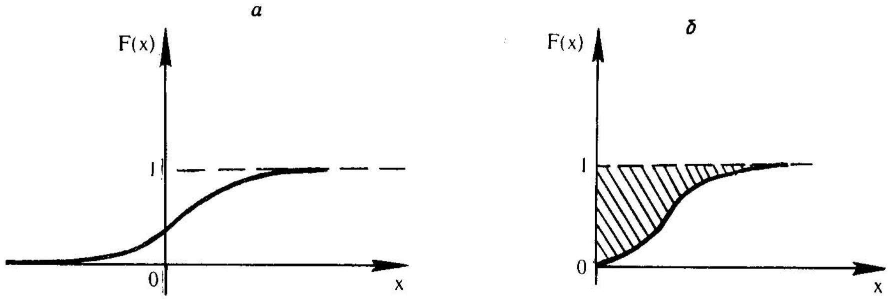


Рис. 2.3

2.20. Функция распределения $F(x)$ неотрицательной СВ X задана графиком (рис. 2.3, *а*). Математическое ожидание СВ X равно $M(X)$. Показать, что $M(X)$ геометрически может быть представлено площадью фигуры, ограниченной кривой $y = F(x)$, прямой $y = 1$ и осью ординат (она заштрихована на рис. 2.3, *б*).

2.21. Дискретная СВ X задана рядом распределения

x_i	2	3	5
p_i	0,1	0,4	0,5

Найти начальные моменты первого и третьего порядков.

2.22. Плотность распределения вероятностей СВ X задана функцией

$$f(x) = \begin{cases} 0 & \text{при } x < 0, \\ \frac{3}{2}x^2 & \text{при } 0 \leqslant x < 1, \\ \frac{3}{2}(2-x)^2 & \text{при } 1 \leqslant x < 2, \\ 0 & \text{при } x \geqslant 2. \end{cases}$$

Определить начальные и центральные моменты до четвертого порядка включительно.

2.23. К СВ X прибавили постоянную (неслучайную) величину a. Как изменятся характеристики X: 1) математическое ожидание; 2) дисперсия; 3) среднее квадратичное отклонение; 4) начальный момент второго порядка?

2.24. Случайную величину X умножили на постоянную величину a. Как изменятся ее характеристики: 1) математическое ожидание; 2) дисперсия; 3) среднее квадратичное отклонение; 4) начальный момент второго порядка?

2.4. ОСНОВНЫЕ ВЕРОЯТНОСТНЫЕ МОДЕЛИ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Законом распределения СВ называется соотношение, учитывающее связь между возможными ее значениями и соответствующими вероятностями.

Приведем часто встречающиеся законы распределения. Сначала перечислим некоторые дискретные распределения.

1. *Вырожденное распределение*

$$P(X=a)=1,$$

где a — постоянная.

2. *Гипергеометрическое распределение*

$$P(X=k)=C_M^k C_{N-M}^{n-k} / C_N^n,$$

где N, M, n — натуральные числа: $M \leqslant N, n \leqslant N$; $k=0,1,..., \min(M, n)$.

3. *Биномиальное распределение*

$$P(X=k)=P_n(k)=C_n^k p^k q^{n-k},$$

где n — натуральное число; $0 \leqslant p \leqslant 1$; $q=1-p$; $k=0, 1,..., n$. Математическое ожидание $M(X)=np$, дисперсия $D(X)=npq$.

4. *Пуассоновское распределение.* Если вероятность появления события мала, а число испытаний велико, то применение формулы Бернулли затруднительно. В этом случае пользуются ее предельным значением

$$P_n(k)=\lambda^k e^{-\lambda}/k!,$$

где $\lambda=np$. Математическое ожидание и дисперсия СВ X, распределенной по закону Пуассона, соответственно равны $M(X)=\lambda$, $D(X)=\lambda$.

Пример 2.13. Из выборки, содержащей M белых и $N-M$ черных шаров, наудачу извлекается n шаров. Какова вероятность того, что среди выбранных n шаров окажется ровно k белых?

Р е ш е н и е. В этом примере СВ, равная числу шаров в выборке, имеет гипергеометрическое распределение

$$P_n(k)=C_M^k C_{N-M}^{n-k} / C_N^n,$$

где $k=k_0, k_0+1,..., k_1$; $k_0=\max(0, M-N+n)$; $k_1=\min(M, n)$.

Распределение СВ, равной числу успехов в k испытаниях, определяется по биномиальному распределению:

$$P(X=k)=P_n(k)=C_n^k p^k q^{n-k}.$$

Для бесконечной последовательности испытаний Бернулли СВ X, равная числу испытаний до первого успеха включительно, имеет геометрическое распределение

$$P(X=k)=(1-p)^{k-1}p.$$

Перечислим теперь некоторые непрерывные распределения, указав их плотность $f(x)$ и числовые характеристики.

1. *Равномерное распределение на отрезке* $[a; b]$, $a<b$.

Плотность распределения вероятностей

$$f(x)=\begin{cases} 1/(b-a) & \text{при } a \leqslant x \leqslant b, \\ 0 & \text{при } x<a, x<b, \end{cases}$$

$$M(X) = \frac{1}{b-a} \int_a^b xdx = \frac{a+b}{2},$$

$$D(X) = \int_{-a}^{a} (x - \frac{a+b}{2})^2 f(x)\, dx =$$

$$= \frac{1}{b-a} \int_a^b (x - \frac{a+b}{2})^2 dx = \frac{(b-a)^3}{12}.$$

2. *Нормальное (или гауссовское) распределение.* Плотность распределения вероятностей имеет вид

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-a)^2/(2\sigma^2)}, \qquad (2.3)$$

где $a = M(X)$; $\sigma = \sqrt{D(X)}$ – среднее квадратичное отклонение. Математическое ожидание

$$M(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} xe^{-(x-a)^2/(2\sigma^2)} dx.$$

Полагая $x = \sigma y + a$ и применяя формулу интегрирования по частям, получаем

$$M(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma y + a) e^{-y^2/2} dy =$$

$$= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} ye^{-y^2/2} dy + \frac{a}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy = 0 + a \cdot 1 = a.$$

Так как функция $ye^{-y^2/2}$ – нечетная, а $\frac{1}{\sqrt{2\pi}} e^{-y^2/2}$ является плотностью нормального распределения с параметрами $(0, 1)$, то первый интеграл в правой части последней формулы равен 0, а второй a. Таким образом, если СВ X распределена нормально с параметрами (a, σ^2), то $M(X) = a$. Дисперсия

$$D(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-a)^2 e^{-(x-a)^2/(2\sigma^2)} dx = \sigma^2.$$

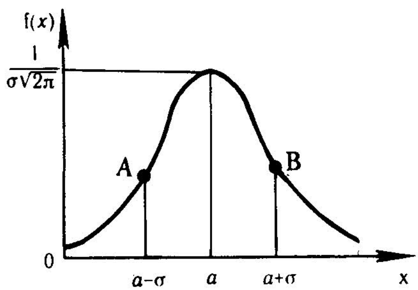


Рис. 2.4

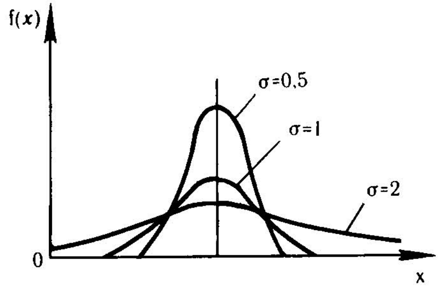


Рис. 2.5

При вычислении $M(X)$ и $D(X)$ использовался интеграл Пуассона: $\int\limits_{-\infty}^{\infty} e^{-z^2/2}dz = \sqrt{2\pi}$.

Если СВ X имеет нормальное распределение с параметрами $M(X) = a$ и $\sigma = \sqrt{D(X)}$, то символически это записывают в виде СВ $X \in N(a\,;\,\sigma)$. График плотности распределения вероятностей, задаваемой формулой (2.3), называют ***нормальной кривой*** или ***кривой Гаусса***. Нормальная кривая представляет собой колоколообразную кривую (рис. 2.4; A, B – точки перегиба), симметричную относительно прямой $x = a$ и асимптотически приближающуюся к оси абсцисс при $x \to \infty$. Приведем без доказательств *основные свойства кривой Гаусса*.

1. *Функция плотности распределения вероятностей определена для любого* $x \in \mathbf{R}$.

2. *Кривая Гаусса расположена над осью Ох для любого* $x \in \mathbf{R}$, $f(x) > 0$.

3. *Ветви кривой Гаусса асимптотически приближаются к оси Ох* $\lim\limits_{x \to \pm\infty} f(x) = 0$.

4. *Кривая Гаусса имеет максимум в точке* $x = a$, *равный* $f(a) = 1/(\sigma\sqrt{2\pi})$.

5. *Кривая Гаусса симметрична относительно прямой* $x = a$. *Следовательно, для СВ* $X \in N(a\,;\,\sigma)$ *математическое ожидание совпадает с модой и медианой распределения.*

На рис. 2.5 изображены кривые Гаусса при $a = 2$ и $\sigma = 2$; 1; 0,5; изменение параметра σ при фиксированном значении a характеризует форму кривой Гаусса.

Нормальному закону распределения подчиняются многие непрерывные СВ, встречающиеся в технике: ошибки измерения, высота микронеровностей на обработанной поверхности, отклонения размеров изготавливаемых деталей от номинального, амплитуды колебания машин, возникающие при движении.

Вероятность попадания СВ X, распределенной нормально, в заданный интервал $]\,\alpha;\,\beta\,[$ вычисляется по формуле

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right), \qquad (2.4$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-t^2/2}\, dt$ – функция Лапласа. Значения функции $\Phi(x)$ приведены в прил. 1 и 2.

Вероятность того, что абсолютная величина отклонения СВ X от своего математического ожидания меньше любого $\epsilon > 0$

$$P(|X - a| < \epsilon) = 2\Phi(\epsilon/\sigma).$$

Пример 2.14. Измерение клиренса автомобиля сопровождается систематическими и случайными ошибками. Систематическая ошибка равна 10 мм (для грузовых автомобилей). Случайные ошибки подчиняются нормальному закону со средним квадратичным отклонением $\sigma(X) = 20$ мм. Найти: 1) вероятность измерения клиренса с ошибкой, не превосходящей по абсолютной величине 30 мм; 2) вероятность того, что измеренный клиренс не превзойдет истинного.

Р е ш е н и е. 1. Обозначим через X суммарную ошибку измерения клиренса. Ее систематическая составляющая $a = 10$ мм. Следовательно, плотность распределения вероятностей суммарной ошибки имеет вид

$$f(x) = \frac{1}{20\sqrt{2\pi}} e^{-(x-10)^2/800}$$

Согласно формуле (2.4), имеем

$$P(|X| < 30) = P(-30 < X < 30) = \left(\Phi\left(\frac{30+10}{20}\right) - \Phi\left(\frac{-30+10}{20}\right)\right) =$$

$$= (\Phi(2) - \Phi(-1)).$$

Так как $\Phi(-1) = -\Phi(1)$, то

$$P(|X| < 30) = (\Phi(2) + \Phi(1)).$$

Из прил. 2 находим: $\Phi(2) = 0{,}4772$, $\Phi(1) = 0{,}3413$. Окончательно имеем

$$P(|X| < 30) = 0{,}8185.$$

2. Вероятность того, что измеренный клиренс не превзойдет истинного,

$$P(-\infty < X < 0) = (\Phi(0{,}5) + \Phi(\infty)).$$

Так как $\Phi(\infty) = \lim\limits_{x \to \infty} \Phi(x) = 0{,}5$, из прил. 2 находим $\Phi(0{,}5) = 0{,}1915$. Тогда

$$P(-\infty < X < 0) = 0{,}6914.$$

Пример 2.15. Для замера напряжений используются специальные тензодатчики. Определить среднюю квадратичную ошибку тензодатчика, если он не имеет систематических ошибок,а случайные распределены по нормальному закону и с вероятностью 0,8 не выходят за пределы $\pm 0{,}2$ мк.

Р е ш е н и е. Из условия задачи следует, что

$$P(|X| < 0{,}2) = 0{,}8.$$

Так как распределение вероятностей случайных ошибок нормальное, а $a = 0$ (систематические ошибки отсутствуют), то

$$P(|X| < 0{,}2) = \Phi(0{,}2/\sigma) - \Phi(-0{,}2/\sigma) = 2\Phi(0{,}2/\sigma),$$

$$2\Phi(0{,}2/\sigma) = 0{,}8.$$

Пользуясь прил. 2, получаем $0{,}2/\sigma = 1{,}28$, откуда $\sigma = 0{,}2/1{,}28 = 1{,}0156$ мк.

Пример 2.16. Браковка шариков для подшипников производится следующим образом: если шарик не проходит через отверстие диаметром d_1, но проходит через отверстие диаметром d_2, то его размер считается приемлемым. Если какое-нибудь из этих условий не выполняется, то шарик бракуется. Известно, что диаметр шарика D есть СВ с характеристиками $m_x = (d_1 + d_2)/2$ и $\sigma_x = (d_2 - d_1)/4$. Определить вероятность P того, что шарик будет забракован.

Р е ш е н и е. Имеем:

$$P = 1 - P(d_1 < d < d_2) = 1 - (\Phi(\frac{d_2 - m_x}{\sigma_x}) -$$

$$- \Phi(\frac{d_1 - m_x}{\sigma_x})) = 1 - (\Phi(\frac{d_2 - d_1}{2\sigma_x}) - \Phi(\frac{d_1 - d_2}{2\sigma_x})) =$$

$$= 1 - 2\Phi(\frac{d_2 - d_1}{2\sigma_x}) = 1 - 2\Phi(2) = 0{,}0456.$$

ЗАДАЧИ

2.25. Срок службы шестерен коробок передач зависит от следующих факторов: усталости материала в основании зуба, контактных напряжений и жесткости конструкции. Вероятность отказа каждого фактора в одном испытании равна 0,1. Составить закон распределения числа отказавших факторов в одном испытании.

2.26. Число неисправностей, обнаруженных во время техосмотра автомобиля, распределено по закону Пуассона с параметром λ. Если неисправностей не обнаружено, техническое обслуживание машины продолжается в среднем 2 ч. Если найдены одна или две неисправности, то на устранение каждой из них затрачивается в среднем 1/2 ч. Если обнаружено больше двух неисправностей, то машина ставится на профилактический ремонт, где она находится в среднем 4 ч. Определить закон распределения среднего времени T обслуживания и ремонта машины и его математическое ожидание.

2.27. Время t между двумя сбоями вычислительной машины распределено по показательному закону с параметром λ:

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{при } t > 0, \\ 0 & \text{при } t < 0. \end{cases}$$

Решение задачи требует безотказной работы машины в течение времени τ. Если за время τ произошел сбой, то задачу приходится решать заново. Сбой обнаруживается только через время τ после начала решения. Найти закон распределения СВ θ (время, за которое задача будет решена).

2.28. В механизмах и узлах ходовой части автомобиля из-за неточности обработки деталей и их неуравновешенности возникают динамические нагрузки, средние значения числа которых достигают 300 за 1 ч езды. Какова вероятность того, что за 1 мин движения автомобиля динамические нагрузки не превысят своего среднего значения ровно в два раза?

2.29. Распределение уклонов на дорогах есть СВ X, которая описывается нормальным законом. Величину математического ожидания $M(X)$ для среднепересеченной местности можно принять равной нулю, так как при движении автомобиля в одном направлении половина уклонов будет представлять собой подъемы, другая половина – спуски. Определить среднее квадратичное отклонение распределения уклонов σ, при котором вероятность $P(0{,}21 < X < 0{,}57)$ была бы наибольшей.

2.30. Завод изготавливает шарики для подшипников. Номинальный диаметр шариков $d_0 = 5$ мм. Вследствие неточности, допущенной при изготовлении шарика, фактический его диаметр – СВ со средним значением d_0 и средним квадратичным отклонением $\sigma_x = 0{,}05$ мм. При контроле бракуются все шарики, диаметр которых отличается от номинального больше, чем на 0,1 мм. Определить, какой процент шариков в среднем будет отбраковываться.

2.31. Какой величины должно быть поле допуска зубчатого колеса, чтобы с вероятностью не более 0,003 изготовленное колесо с контролируемым размером оказалось вне поля допуска? Случайные отклонения размера от середины поля допуска подчинены закону нормального распределения с параметрами $a = 0$, $\sigma = 5$ мк.

2.32. Суммарное максимальное напряжение в станинах состоит из максимального изгибающего напряжения и напряжения от действия продольных сил. Прибором, имеющим среднюю квадратичную 20 кг/см^2 и систематическую 7 кг/см^2 ошибки, производят два измерения указанных напряжений. Какова вероятность того, что обе ошибки измерений, имея разные знаки, по абсолютной величине не превзойдут 7 кг/см^2?

2.33. Высота микронеровностей поверхности обработанных деталей есть СВ X и подчинена закону нормального распределения с математическим ожиданием a и средним квадратичным отклонением σ (оно служит для оценки чистоты поверхности). Годными деталями являются те, для которых $b < X < c$. К деталям, подлежащим дополнительной обработке (шлифовке), относятся те, для которых $X > c$. Найти: 1) функцию распределения микронеровностей для деталей, подлежащих шлифовке; 2) функцию распределения микронеровностей для годных деталей.

2.34. Линейные вертикальные ускорения кузова автомобиля, возникающие при движении, являются СВ X и подчинены закону нормального распределения. При скорости движения автомобиля 90 км/ч среднее квадратичное отклонение ускорений составляет 12 м/с^2. Определить вероятность того, что в результате движения ускорения не превысят 30 м/с^2.

2.35. При проектировании автомобилей определяют конструктивный динамический ход колеса f_d и устанавливают ограничители хода колес по среднему квадратичному отклонению σ_f (конструктивный динамический ход f_d выбирают в зависимости от допустимого значения вероятности пробивания подвески при расчетном режиме, т.е. заданной скорости движения в данных дорожных условиях). Зная, что $f_d/\sigma_f = 2{,}5$, определить вероятность пробивания подвески $P(f_q > f_d)$, где f_q – динамический ход колеса, являющийся СВ, равной разности вертикальных перемещений оси колеса и точки кузова машины, расположенной над осью колеса при движении.

3. МНОГОМЕРНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

3.1. ПОНЯТИЕ МНОГОМЕРНОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ. ЗАКОН РАСПРЕДЕЛЕНИЯ

Если случайное событие описывается упорядоченной совокупностью действительных чисел $X_1, X_2, ..., X_n$, то эта совокупность представляет собой значение n-мерной СВ $(X_1, X_2, ..., X_n)$. *Системой случайных величин (случайным вектором)* называется совокупность СВ, описывающих то или иное случайное событие. Всякое элементарное событие в подобном случае можно рассматривать как результат сложного опыта, состоящего в определении всех величин $X_1, X_2, ..., X_n$, и интерпретировать точкой n-мерного пространства $(X_1, X_2, ..., X_n)$ или вектором $\mathrm{X}\ (X_1, X_2, ..., X_n)$. Каждая из величин $X_1, X_2, ..., X_n$ является одномерной СВ. Если X – случайный вектор (или n-мерная СВ) , то величины $X_1, X_2, ..., X_n$ называются *случайными координатами.*

Геометрически система из двух СВ представляется случайной точкой $M\ (x, y)$ или случайным вектором $\overrightarrow{OM}$ на плоскости Oxy. Аналогично система из трех СВ представляется случайной точкой в трехмерном пространстве $M(x, y, z)$ или случайным вектором $\overrightarrow{OM}$ в этом пространстве $Oxyz$.

Система случайных величин называется *дискретной,* если все входящие в нее СВ дискретны. Система СВ называется *непрерывной,* если все величины, входящие в систему, непрерывны. Если в систему входят как непрерывные, так и дискретные СВ, то ее называют *смешанной.*

Распределением системы СВ называется соотношение, устанавливающее зависимость между значениями системы СВ и вероятностями их появления.

Распределением дискретной двухмерной СВ (X, Y) называется перечень возможных значений этой величины и их вероятностей $p(x_i, y_j)$, $i,j = 1,2,...$

Для системы двух дискретных СВ (X, Y) наиболее простым заданием распределения является таблица (табл. 3.1).

Таблица 3.1

$y_j \backslash x_i$	x_1	x_2	...	x_n
y_1	p_{11}	p_{21}	...	p_{n1}
y_2	p_{12}	p_{22}	...	p_{n2}
y_3	p_{13}	p_{23}	...	p_{n3}
...	...	...	...	...
y_m	p_{1m}	p_{2m}	...	p_{nm}

Всевозможные события $X=x_i$, $Y=y_j$ при $i=1,2,...,n$, $j=1,2,...,m$ составляют *полную группу совместных событий*, т.е.

$$\sum_{ij} p_{ij} = \sum_{ij} P(X=x_i, Y=y_j) = 1.$$

При этом

$$\sum_{j} p_{ij} = \sum_{j} P(X=x_i, Y=y_j) = P(X=x_i),$$

$$\sum_{i} p_{ij} = \sum_{i} P(X=x_i, Y=y_j) = P(Y=y_j).$$

Пример 3.1. Найти распределения составляющих двухмерной СВ (для деталей, работающих на изгиб X и кручение Y), заданной следующей таблицей:

y_j \ x_i	x_1	x_2	x_3
y_1	0,18	0,22	0,16
y_2	0,08	0,16	0,20

Р е ш е н и е. Сложив вероятности по столбцам, получим вероятности возможных значений для деталей, работающих на изгиб X:

$$P(X=x_1) = 0{,}18 + 0{,}08 = 0{,}26,\ P(X=x_2) = 0{,}22 + 0{,}16 =$$
$$= 38,\ P(X=x_3) = 0{,}16 + 0{,}20 = 0{,}36.$$

Ряд распределения составляющей X имеет вид:

x_i	x_1	x_2	x_3
p_i	0,26	0,38	0,36

Контроль: $0{,}26 + 0{,}38 + 0{,}36 = 1$.

Сложив вероятности по строкам, получим вероятности возможных значений нагрузок для деталей, работающих на кручение Y:

$$P(Y=y_1) = 0{,}18 + 0{,}22 + 0{,}16 = 0{,}56,$$
$$P(Y=y_2) = 0{,}08 + 0{,}16 + 0{,}20 = 0{,}44.$$

Ряд распределения составляющей Y имеет вид:

y_j	y_1	y_2
p_j	0,56	0,44

Контроль: $0{,}56 + 0{,}44 = 1$.

ЗАДАЧИ

3.1. Число рабочих циклов двигателя X и пробег автомобиля Y взаимосвязаны. Найти распределения составляющих СВ (X, Y), заданных следующей двухмерной таблицей распределения вероятностей:

y_j \ x_i	x_1	x_2	x_3
y_1	0,106	0,062	0,082
y_2	0,116	0,160	0,070
y_3	0,111	0,111	0,182

3.2. Контроль партии шариков после первой доводки производится по овальности (наибольшее отклонение диаметра от номинала) и гранности (отклонение среднего значения диаметра). При установившемся процессе производства около 6 % шариков после первой доводки не удовлетворяет техническим требованиям, причем 2 % брака вызвано овальностью шариков, 3 % – гранностью и 1 % – обоими признаками. Найти распределения системы двух СВ и ее составляющих.

3.3. Станок-автомат изготавливает валики. Чтобы деталь была годной, она должна удовлетворять допустимым значениям по длине и по диаметру. Вероятность того, что валик будет признан годным по длине, равна 0,8, а по диаметру – 0,7. Найти распределения системы СВ (X, Y) и ее составляющих.

3.4. По цели производится два выстрела. Вероятность попадания при одном выстреле равна 0,7. Найти распределение системы СВ (X, Y), считая, что X – число попаданий, а Y – число промахов.

3.5. Станок-автомат изготавливает кольца. Для того чтобы деталь была годной, она должна удовлетворять допустимым значениям по внутреннему и наружному диаметрам. Вероятность того, что кольцо будет признано годным по внутреннему диаметру, равна 0,75, а по внешнему – 0,80. Найти распределения системы СВ и ее составляющих.

3.2. ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ И ПЛОТНОСТЬ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ МНОГОМЕРНОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Функцией распределения (многомерной функцией распределения) системы СВ $(X_1, X_2,...,X_n)$ *(случайного вектора* $\mathrm{X}(X_1, X_2,..., X_n)$) называется вероятность появления события $\{X_1 < x_1 \cap X_2 < x_2 \cap ... \cap X_n < x_n\}$, т.е.

$$F(x_1, x_2,..., x_n) = P\{(X_1 < x_1) \cap (X_2 < x_2) \cap \ldots \cap (X_n < x_n)\}.$$

Функцией распределения $F(x, y)$ *системы двух СВ* (X, Y) или *функцией распределения двухмерного случайного вектора* называется вероятность совместного выполнения неравенства $X < x$, $Y < y$, т.е.

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

или вероятность появления события $\{(X < x) \cap (Y < y)\}$, т.е.

$$F(x, y) = P\{(X < x) \cap (Y < y)\},$$

где x и y – переменные величины.

Приведем *свойства функции распределения системы* СВ (X, Y).

1. *Функция* $F(x, y)$ *удовлетворяет неравенству*

$$0 \leqslant F(x, y) \leqslant 1.$$

2. *Функция* $F(x, y)$ *является неубывающей, т.е. при* $x_2 > x_1, y_2 > y_1$

$$F(x_2, y_2) \geqslant F(x_1, y_1).$$

3. *Справедливы следующие соотношения:*

$$F(-\infty, y) = 0,\ F(x_1, -\infty) = 0,\ F(-\infty, -\infty) = 0,$$

$$F(\infty, \infty) = 1,\ F(x, \infty) = F_1(x),\ F(\infty, y) = F_2(y).$$

Следовательно, чтобы найти функцию распределения одной составляющей системы двух СВ (одной координаты двухмерного случайного вектора), нужно устремить переменную, соответствующую составляющей (координате) вектора, к ∞.

Плотностью $f(x_1, x_2, \dots, x_n)$ *распределения вероятностей системы СВ* $(X_1, X_2, \dots, X_n)$ *(случайного вектора* $X(x_1, x_2, \dots, x_n)$*)* называется смешанная частная производная порядка n от функции распределения $F(x_1, x_2, \dots, x_n)$, взятая по всем переменным, т.е.

$$f(x_1, x_2, \dots, x_n) = \frac{\partial^n F(x_1, x_2, \dots, x_n)}{\partial x_1 \partial x_2 \dots \partial x_n}.$$

Функция распределения для непрерывных СВ выражается через плотность распределения в виде кратного интеграла

$$F(x_1, x_2, \dots, x_n) = \int_{-\infty}^{x_n} \int_{-\infty}^{x_{n-1}} \dots \int_{-\infty}^{x_1} f(x_1, x_2, \dots, x_n)\, dx_1 dx_2 \dots dx_n .$$

Плотностью $f(x, y)$ *распределения вероятностей системы двух непрерывных СВ* называется вторая смешанная частная производная от функции распределения, т.е.

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}. \tag{3.1}$$

Геометрически плотность $f(x, y)$ можно истолковать как поверхность распределения.

Если $f(x, y)$ – плотность распределения вероятностей в некоторой замкнутой области D, то вероятность попадания в эту область (X, Y) можно определить по формуле

$$P((X, Y) \in D) = \iint_D f(x, y)\, dx dy .$$

Перечислим *свойства плотности распределения вероятностей* $f(x, y)$.

1. $f(x, y) \geqslant 0$.

2. $\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} f(x, y)\, dxdy = 1.$

3. $\int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} f(x, y)\, dxdy = F(x, y).$

4. $\int\limits_{-\infty}^{\infty} f(x, y)dy = f(x),$

$\int\limits_{-\infty}^{\infty} f(x, y)dx = f(y).$

График плотности распределения вероятностей составляющих X, Y изображен на рис. 3.1.

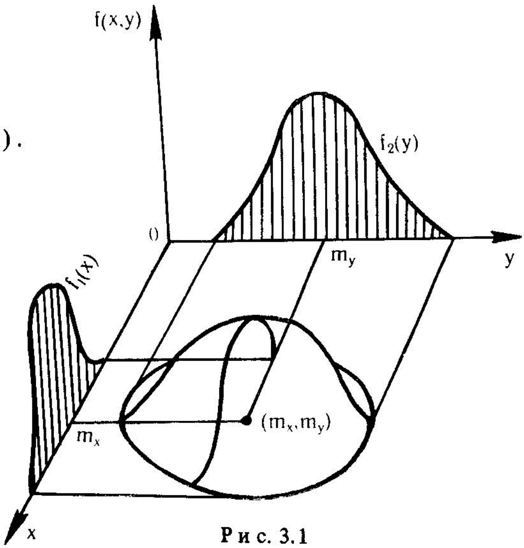


Рис. 3.1

Пример 3.2. Функция распределения системы двух СВ (X, Y) имеет вид

$$F(x, y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-x-y} & \text{при } x > 0,\ y > 0, \\ 0 & \text{при } x \leq 0 \text{ или } y \leq 0. \end{cases}$$

Найти плотность распределения вероятностей $f(x, y)$.

Решение. Согласно формуле (3.1), имеем

$$\frac{\partial F}{\partial x} = \begin{cases} e^{-x} - e^{-x-y} & \text{при } x > 0,\ y > 0, \\ 0 & \text{при } x \leq 0 \text{ или } y \leq 0. \end{cases}$$

Отсюда

$$f(x, y) = \frac{\partial^2 F}{\partial x \partial y} = \begin{cases} -e^{-x-y} & \text{при } x > 0,\ y > 0, \\ 0 & \text{при } x \leq 0 \text{ или } y \leq 0. \end{cases}$$

Пример 3.3. Плотность распределения вероятностей случайного вектора $\mathbf{X}(X, Y)$ имеет вид

$$f(x, y) = \begin{cases} C(x + y) & \text{при } 0 \leq y \leq x,\ 0 \leq x \leq 1, \\ 0 & \text{в остальных случаях} \end{cases}$$

Определить константу C и плотность распределения вероятностей компоненты Y.

Решение. Так как СВ X, Y изменяются в замкнутой области, то

$\iint\limits_{D} f(x, y)\, dxdy = 1.$ Следовательно,

$$1=\int\limits_0^1\int\limits_0^x C\,(x+y)dydx = C\int\limits_0^1 dx\int\limits_0^x (x+y)\,dy =$$

$$= C\int\limits_0^1 dx\,(xy+\frac{y^2}{2})\Big|_0^x = C\int\limits_0^1 (x^2+\frac{x^2}{2})\,dx =$$

$$= C(\frac{x^3}{2})\Big|_0^1 = \frac{C}{2},$$

т.е. $C = 2$;

$$f_2(y) = \int\limits_{-\infty}^{\infty} f(x,y)\,dx = \int\limits_0^1 2(x+y)dx = 2(\frac{1}{2}+y) = 1+2y.$$

ЗАДАЧИ

3.6. Существует несколько способов фиксации величины зерна аустенита в стали. Определение величины зерна производится под микроскопом при стократном увеличении путем сравнения видимых на шлифе зерен с их эталонными изображениями. Размеры X, Y зерен распределены равномерно внутри прямоугольника, ограниченного абсциссами $x = a$, $x = b$ и ординатами $y = c$, $y = d$ ($b > a$, $d > c$). Найти плотность распределения вероятностей и функцию распределения системы СВ (X, Y).

3.7. Плотность распределения вероятностей системы СВ (X, Y) имеет вид

$$f(x,y) = \frac{A}{\pi^2\,(16+x^2)\,(25+y^2)}.$$

Требуется: 1) определить значение A; 2) найти функцию распределения $F(x, y)$.

3.8. Найти вероятность попадания случайной точки (X, Y) в прямоугольник, ограниченный прямыми $x = \pi/4$, $x = \pi/2$, $y = \pi/6$, $y = \pi/3$, если известна функция распределения $F(x,y) = \sin x \sin y$ при $0 \leq x \leq \pi/2$, $0 \leq y \leq \pi/2$.

3.9. Внутри прямоугольника, ограниченного прямыми $x = 0$, $x = \pi/2$, $y = 0$, $y = \pi/2$, плотность распределения вероятностей системы двух СВ имеет вид $f(x,y) = C\sin(x+y)$. Вне прямоугольника $f(x,y) = 0$. Найти значение C и функцию распределения данной системы.

3.10. Найти вероятность того, что составляющая X примет значение, меньшее 1/2, а составляющая Y – значение, меньшее $1/3$, если функция распределения системы СВ

$$F(x,y) = (\frac{1}{\pi}\,\mathrm{arctg}\,2x+\frac{1}{2})\,(\frac{1}{\pi}\,\mathrm{arctg}\,3y+\frac{1}{2}).$$

3.11. Два независимо работающих автомата обрабатывают детали. Случайная величина X выражает число появлений события A, соответствующего качественно обработанной детали на первом станке-автомате, а Y – на втором станке-автомате. Вероятность качественного изготовления детали для первого станка p_1, для второго p_2. Построить функцию распределения $F(x,y)$ системы СВ.

3.12. Для случайного вектора (X, Y) функция распределения $F(x,y)$ имеет вид

$$F(x,y) = \begin{cases} \sin x\sin y & \text{при } x \geq 0,\ y \leq \pi/2, \\ 0 & \text{при } x < 0,\ y > \pi/2. \end{cases}$$

Найти плотности распределения вероятностей СВ X и Y и случайного вектора (X, Y).

3.13. Плотность распределения вероятностей случайного вектора (X, Y) имеет вид

$$f(x,y) = \begin{cases} xe^{-x(1+y)} & \text{при } x \geqslant 0, y \geqslant 0, \\ 0 & \text{при } x < 0, y < 0. \end{cases}$$

Найти функцию распределения вектора (X, Y) и плотности распределения вероятностей СВ X и Y.

3.14. Два автомата независимо друг от друга производят по два изделия. Вероятность изготовления качественного изделия первым автоматом равна p_1, а вторым – p_2; X – количество изделий, изготовленных первым автоматом, Y – количество изделий, изготовленных вторым автоматом. Найти функцию распределения случайного вектора (X, Y).

3.3. УСЛОВНЫЕ РАСПРЕДЕЛЕНИЯ СОСТАВЛЯЮЩИХ МНОГОМЕРНОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Введем понятие условной функции распределения и условной плотности распределения вероятностей для двухмерной СВ (двухмерного вектора).

Условной функцией распределения $F(x/y_1, y_2)$ СВ X относительно определенных значений СВ Y называется условная вероятность выполнения неравенства $X < x$ относительно значений Y, удовлетворяющих неравенству $y_1 \leqslant Y \leqslant y_2$. Из определения следует:

$$F(x/y_1, y_2) = \frac{P(X<x, y_1 \leqslant Y \leqslant y_2)}{P(y_1 \leqslant Y \leqslant y_2)}$$

при условии, что $P(y_1 \leqslant Y \leqslant y_2) > 0$, или

$$F(x/y_1, y_2) = \frac{\int\limits_{-\infty}^{x} \int\limits_{y_1}^{y_2} f(x,y)\,dx\,dy}{\int\limits_{y_1}^{y_2} \int\limits_{-\infty}^{\infty} f(x,y)\,dxdy}.$$

Условной плотностью распределения вероятностей СВ X при условии, что $Y = y$, называется отношение плотности распределения вероятностей системы (X, Y) к плотности распределения вероятностей *СВ* Y, т.е.

$$f_1(x/y) = f(x,y)/f_2(y) \quad \text{при} \quad f_2(y) \neq 0$$

или

$$f_1(x/y) = \frac{f(x,y)}{\int\limits_{-\infty}^{\infty} f(x,y)\,dx}, \quad f_2(y/x) = \frac{f(x,y)}{\int\limits_{-\infty}^{\infty} f(x,y)\,dy}$$

Случайные величины X и Y называются *независимыми*, если события $X<x$ и $Y<y$ являются независимыми при любых значениях x и y. Если события $X<x$ и $Y<y$ являются зависимыми при каких-либо значениях x и y, то СВ называются *зависимыми*. Для независимых *СВ* X и Y справедливы соотношения:

$$F(x,y)=F_1(x)\,F_2(y),\quad f(x,y)=f_1(x)\,f_2(y).$$

Распределение случайного вектора $X(X,Y)$ в прямоугольнике $x_1\leqslant X\leqslant x_2$, $y_1\leqslant Y\leqslant y_2$, площадь которого $S=(x_2-x_1)(y_2-y_1)$, называют *равномерным*, если плотность распределения вероятностей X для всех (x,y), принадлежащих прямоугольнику, постоянна: $f(x,y)=1/S$ и $f(x,y)=0$ вне прямоугольника. Каждая из составляющих X и Y будет равномерно распределенной на соответствующем интервале:

$$f_1(x)=\frac{1}{x_2-x_1},\ x_1\leqslant x\leqslant x_2;\quad f_2(y)=\frac{1}{y_2-y_1},\ y_1\leqslant y\leqslant y_2.$$

Пример 3.4. Случайная точка (X,Y) распределена с постоянной плотностью, равной 1/2 внутри квадрата с вершинами $(1,0)$, $(0,1)$, $(-1,0)$, $(0,-1)$. Найти плотности распределения вероятностей $f_1(x)$, $f_2(x)$ отдельных величин X, Y и условные плотности распределения вероятностей $f_1(x/y)$ и $f_2(y/x)$. Установить, зависимы или независимы случайные величины X и Y.

Р е ш е н и е. Площадь квадрата равна 2. Следовательно,

$$f(x,y)=\begin{cases}1/2 \text{ при } |x|<1,\ |y|<1,\\ 0 \quad \text{при } |x|>1;\end{cases}$$

$$f_1(x)=\begin{cases}\dfrac{1}{2}\displaystyle\int_{-(1-x)}^{1-x}dy=1-x & \text{при } 0<x<1,\\ \displaystyle\int_{-(1+x)}^{1+x}dy=1+x & \text{при } -1<x<0,\\ 0 & \text{при } x<-1 \text{ или } x>1\end{cases}$$

или

$$f_1(x)=\begin{cases}1-|x| & \text{при } |x|<1,\\ 0 & \text{при } |x|>1.\end{cases}$$

Аналогично

$$f_2(x)=\begin{cases}1-|y| & \text{при } |y|<1,\\ 0 & \text{при } |y|>1.\end{cases}$$

При $|y| < 1$ имеем

$$f_1(x/y) = \frac{f(x,y)}{f_2(y)} = \begin{cases} \dfrac{1}{2(1-|y|)} & \text{при } |x| < 1-|y|, \\ 0 & \text{при } |x| > 1-|y|. \end{cases}$$

Аналогично при $|x| < 1$ получаем

$$f_2(x/y) = \begin{cases} \dfrac{1}{2(1-|x|)} & \text{при } |y| < 1-|x|, \\ 0 & \text{при } |y| > 1-|x|. \end{cases}$$

Случайные величины X, Y зависимы.

Пример 3.5. Изготавливаемые в цехе втулки сортируются по отклонениям внутреннего диаметра от номинального на четыре группы по значениям 0,01, 0,02, 0,03, 0,04 мм и по овальности – на четыре группы по значениям 0,02, 0,04, 0,06, 0,08. Распределения отклонений диаметра X и овальности Y приведены в табл. 3.2 и 3.3 соответственно.

Таблица 3.2

y_j \ x_i	0,01	0,02	0,03	0,04
0,02	0,01	0,02	0,04	0,04
0,04	0,03	0,24	0,15	0,06
0,06	0,04	0,10	0,08	0,08
0,08	0,02	0,04	0,03	0,02

Таблица 3.3

x_i	0,01	0,02	0,03	0,04
p_i	0,10	0,40	0,30	0,20
y_j	0,02	0,04	0,06	0,08
p_j	0,11	0,48	0,30	0,11

Найти условные вероятности $P(X = x_i / y = 0{,}06)$ и условную функцию распределения отклонения внутреннего диаметра от номинального втулок, отнесенных в группу по овальности 0,06 : $F(x/y = 0{,}06)$.

Р е ш е н и е. Условные вероятности:

$$P(x=0{,}01/y=0{,}06) = \frac{P(x=0{,}01;\ y=0{,}06)}{P(y=0{,}06)} = \frac{0{,}04}{0{,}30} \cong 0{,}13,$$

$$P(x=0{,}02/y=0{,}06) = \frac{P(x=0{,}02;\ y=0{,}06)}{P(y=0{,}06)} = \frac{0{,}10}{0{,}30} \cong 0{,}33,$$

$$P(x=0{,}03/y=0{,}06) = \frac{P(x=0{,}03;\ y=0{,}06)}{P(y=0{,}06)} = \frac{0{,}08}{0{,}30} \cong 0{,}27,$$

$$P(x=0{,}04/y=0{,}06) = \frac{P(x=0{,}04;\ y=0{,}06)}{P(y=0{,}06)} = \frac{0{,}08}{0{,}30} \cong 0{,}27.$$

Контроль: $0{,}13 + 0{,}33 + 0{,}27 + 0{,}27 = 1$.

Найдем условную функцию распределения ($F(x/y=0{,}06)$). По определению

$$F(x/y=0{,}06) = \begin{cases} 0 & \text{при } x \leq 0{,}01, \\ 0{,}13 & \text{при } 0{,}01 < x \leq 0{,}02, \\ 0{,}46 & \text{при } 0{,}02 < x \leq 0{,}03, \\ 0{,}73 & \text{при } 0{,}03 < x \leq 0{,}04, \\ 1 & \text{при } x > 0{,}04. \end{cases}$$

Пример 3.6. Поверхность распределения системы СВ (X, Y) представляет собой прямой круговой конус, основанием которого служит круг с центром в начале координат и радиусом r_0. Вне этого круга плотность распределения вероятностей равна нулю. Найти плотность распределения вероятностей $f(x, y)$ системы, плотности распределения вероятностей величин, входящих в систему: $f_1(x)$, $f_2(x)$, условные плотности распределения вероятностей $f_1(x/y)$ и $f_2(x/y)$ и определить, являются ли СВ зависимыми.

Р е ш е н и е. Имеем

$$f(x,y) = \begin{cases} \dfrac{3}{r_0^3\pi}\left(r_0 - \sqrt{x^2+y^2}\right) & \text{при } (x^2+y^2) < r_0^2, \\ 0 & \text{при } (x^2+y^2) > r_0^2. \end{cases}$$

Плотности распределения вероятностей отдельных СВ X, Y, входящих в систему (X, Y), выразятся через плотность распределения вероятностей системы:

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y)\,dy, \quad f_2(y) = \int_{-\infty}^{\infty} f(x,y)\,dx,$$

$$f_1(x) = \begin{cases} \dfrac{3}{r_0^3\pi}\left(r_0\sqrt{r_0^2-x^2} - x^2\ln\left(\dfrac{r_0+\sqrt{r_0^2-x^2}}{|x|}\right)\right) & \text{при } |x| < r_0, \\ 0 & \text{при } |x| > r_0, \end{cases}$$

$$f_2(y) = \begin{cases} \dfrac{3}{r_0^3\pi}\left(r_0\sqrt{r_0^2-y^2} - y^2\ln\left(\dfrac{r_0+\sqrt{r_0^2-y^2}}{|y|}\right)\right) & \text{при } |y| < r_0, \\ 0 & \text{при } |y| > r_0. \end{cases}$$

Условные плотности распределения $f_1(x/y)$ и $f_2(y/x)$ выражаются через безусловные, т.е.

$$f_1(x/y) = \frac{f(x,y)}{f_2(y)} = \begin{cases} \dfrac{r_0-\sqrt{x^2+y^2}}{r_0\sqrt{r_0^2-y^2}-y^2\ln\dfrac{r_0+\sqrt{r_0^2-y_2}}{|y|}} & \text{при } |x| < \sqrt{r_0^2-y^2}, \\ 0 & \text{при } |x| > \sqrt{r_0^2-y^2}, \end{cases}$$

$$f_2(y/x) = \frac{f(x,y)}{f_1(x)} = \begin{cases} \dfrac{r_0-\sqrt{x^2+y^2}}{r_0\sqrt{r_0^2-x^2}-x^2\ln\dfrac{r_0+\sqrt{r_0^2-x^2}}{|x|}} & \text{при } |y| < \sqrt{r_0^2-x^2}, \\ 0 & \text{при } |y| > \sqrt{r_0^2-x^2}. \end{cases}$$

Так как $f_1(x/y) \neq f_1(x)$, то СВ X и Y зависимы.

ЗАДАЧИ

3.15. Система двух СВ задана плотностью распределения вероятностей

$$f(x,y) = \frac{3\sqrt{3}}{\pi}e^{-4x^2-6xy-9y^2}.$$

Найти условные плотности распределения вероятностей $f_1(x/y)$ и $f_2(y/x)$.

3.16. Система СВ равномерно распределена внутри круга радиусом r:

$$f(x,y) = \begin{cases} 1/(\pi r^2) & \text{при } x^2+y^2 < r^2, \\ 0 & \text{при } x^2+y^2 > r^2. \end{cases}$$

Найти плотности распределения вероятностей СВ X и Y и их условные плотности распределения вероятностей, а также установить их зависимость.

3.17. Закон распределения системы двух СВ (X, Y) задан следующей таблицей:

y_j \ x_i	10	20	30
20	3λ	2λ	λ
40	λ	4λ	2λ
60	0	2λ	5λ

Вычислить λ. Найти ряд распределения для каждой из СВ X и Y.

3.18. Найти плотность распределения вероятностей суммы $X_1 + X_2$ независимых СВ X_1 и X_2, каждая из которых имеет равномерное распределение на отрезке $[0;\ 1]$.

3.4. ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СИСТЕМ СЛУЧАЙНЫХ ВЕЛИЧИН

К основным числовым характеристикам системы непрерывных СВ (X_1, $X_2,\ldots,X_n$) (случайных векторов) относятся *математические ожидания*

$$M(X_i) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} x_i\, f_i\,(x_1, x_2,\ldots, x_n)\,dx_1 dx_2\ldots dx_n\ ,$$

дисперсии

$$D(X_i) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}(x_i - M(x_i))^2\, f(x_1, x_2,\ldots, x_n)\, dx_1 dx_2\ldots dx_n$$

и *вторые смешанные центральные моменты*

$$K_{ij} = M((X_i - M(X_i))(X_j - M(X_j))) =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}(x_i - M(x_i))(x_j - M(x_j))\, f(x_1, x_2,\ldots,x_n)\,dx_1 dx_2\ldots dx_n\,.$$

Для дискретных СВ числовые характеристики определяют аналогично, заменяя интегрирование суммированием по всем возможным значениям СВ.

Вторые центральные моменты составляют *корреляционную (ковариационную) матрицу*

$$[K_{ij}] = \begin{bmatrix} K_{11} & K_{12} & \ldots & K_{1n} \\ K_{21} & K_{22} & \ldots & K_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ K_{n1} & K_{n2} & \ldots & K_{nn} \end{bmatrix},$$

где $K_{ij} = K_{ji}$.

Коэффициент корреляции для дискретных СВ

$$r_{ij} = K_{ij} / \sqrt{D(X_i)\, D(X_j)}\,.$$

Коэффициенты корреляции составляют ***нормированную корреляционную матрицу***

$$[r_{ij}] = \begin{bmatrix} 1 & r_{12} & \dots & r_{1n} \\ r_{21} & 1 & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & 1 \end{bmatrix}.$$

Если корреляционный момент СВ равен нулю, то эти величины называются ***некоррелируемыми***, в противном случае – ***коррелируемыми***.

Для независимых СВ $X_1, X_2, \dots, X_n$ справедливо равенство

$$f(x_1, x_2, \dots, x_n) = f_1(x_1) f_2(x_2) \cdots f_n(x_n)\,.$$

Если СВ зависимы, то

$$f(x_1, x_2, \dots, x_n) \neq f_1(x_1) f_2(x_2) \cdots f_n(x_n)\,.$$

Для системы двух непрерывных СВ X и Y математическое ожидание

$$M(X) = \int_{-\infty}^{\infty} x f_1(x) dx = m_x\,,$$
$$M(Y) = \int_{-\infty}^{\infty} y f_2(y)\, dy = m_y\,; \tag{3.2}$$

для дискретных

$$m_x = \sum_i \sum_j x_i p_{ij}\,,$$
$$m_y = \sum_i \sum_j y_j p_{ij}\,.$$

Начальным моментом порядка $k+s$ ***системы СВ*** (X, Y) называется математическое ожидание произведения $X^k Y^s$, т.е.

$$\nu_{k,s} = M(X^k Y^s)\,.$$

Для непрерывных СВ

$$\nu_{k,s} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^k y^s f(x, y)\, dx dy\,,$$

для дискретных

$$\nu_{k,s} = \sum_i \sum_j x_i^k y_j^s \; p_{ij}.$$

Центральным моментом порядка $k+s$ *системы СВ* (X, Y) называют математическое ожидание произведения $(X - m_x)^k (Y - m_y)^s$, т.е.

$$\mu_{k,s} = M((X - m_x)^k (Y - m_y)^s),$$

где m_x, m_y – математические ожидания составляющих X и Y.

Для непрерывных СВ

$$\mu_{k,s} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - m_x)^k (y - m_y)^s f(x, y)\, dxdy,$$

для дискретных СВ

$$\mu_{k,s} = \sum_i \sum_j (x_i - m_x)^k (y_j - m_y)^s p_{ij}.$$

Дисперсии составляющих X и Y

$$D_x = M((X - m_x)^2 (Y - m_y)^0) = M((X - m_x)^2),$$

$$D_y = M((Y - m_y)^2).$$

Корреляционным моментом (моментом связи) системы СВ (X, Y) называется математическое ожидание произведения отклонений этих величин от своих математических ожиданий (второй смешанный центральный момент)

$$K_{xy} = M((X - m_x)(Y - m_y)). \tag{3.3}$$

Для непрерывных СВ

$$K_{xy} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - m_x)(y - m_y) f(x, y)\, dxdy, \tag{3.4}$$

для дискретных СВ

$$K_{xy} = \sum_i \sum_j (x_i - m_x)(y_j - m_y) p_{ij}, \quad K_{xy} = K_{yx}.$$

Коэффициент корреляции для непрерывных СВ

$$r_{xy} = K_{xy} / \sqrt{D_x D_y}.$$

Если СВ X и Y независимы, то $K_{xy} = 0$, $r_{xy} = 0$. Если система содержит n СВ, то корреляционные моменты каждой пары этих СВ

$$K_{ij} = K_{x_i x_j} = M((X_i - m_{x_i})(X_j - m_{x_j})). \qquad (3.5)$$

Если в формуле (3.5) $i = j$, то $K_{ij} = D_i = \sigma_i^2$. Учитывая это, все корреляционные моменты и дисперсии можно представить в виде корреляционной матрицы

$$[K_{ij}] = \begin{bmatrix} \sigma_1^2 & K_{12} & \dots & K_{1n} \\ K_{21} & \sigma_2^2 & \dots & K_{2n} \\ \dots & \dots & \dots & \dots \\ K_{n1} & K_{n2} & \dots & \sigma_n^2 \end{bmatrix}.$$

Пример 3.7. Плотность распределения вероятностей СВ X и Y (координат амплитуд колебаний кузова автомобиля при движении)

$$f(x, y) = \begin{cases} 0,5 \sin(x+y) & \text{при } 0 < x < \pi/2,\ 0 < y < \pi/2 \\ 0 & \text{при } y < 0 \text{ или } y > \pi/2. \end{cases}$$

Найти: 1) математические ожидания составляющих системы; 2) дисперсии $D(X)$, $D(Y)$; 3) корреляционный момент K_{xy}.

Р е ш е н и е. 1. Согласно формуле (3.2),

$$M(X) = 0,5 \int_0^{\pi/2}\int_0^{\pi/2} x\sin(x+y)\,dxdy = 0,5 \int_0^{\pi/2} x(\sin x + \cos x)\,dx = \frac{\pi}{4},$$

$$M(Y) = 0,5 \int_0^{\pi/2}\int_0^{\pi/2} y\sin(x+y)\,dxdy = \frac{\pi}{4}.$$

2. Найдем дисперсию:

$$D(X) = 0,5 \int_0^{\pi/2} (x - \frac{\pi}{4})^2 dx \int_0^{\pi/2} \sin(x+y)\,dy =$$

$$= 0,5 \int_0^{\pi/2} (x - \frac{\pi}{4})^2 (\sin x + \cos x)\,dx =$$

$$= (\pi/8 + \pi - 4)\,0,5 = \pi^2/16 + \pi/2 - 2.$$

Так как выражение для $D(Y)$ имеет такой же вид, как и для $D(X)$, то можем записать, что $D(Y) = \pi^2/16 + \pi/2 - 2$.

3. Согласно формуле (3.3),

$$K_{xy} = 0,5 \int_0^{\pi/2}\int_0^{\pi/2} (x - \frac{\pi}{4})(y - \frac{\pi}{4})\sin(x+y)\,dxdy =$$

$$= 0,5 \int_0^{\pi/2}\int_0^{\pi/2} (xy - \frac{\pi}{4}(x+y) + \frac{\pi^2}{16})\sin(x+y)\,dxdy = \pi/2 - 1 - \pi^2/16.$$

ЗАДАЧИ

3.19. Определить математические ожидания системы СВ (X, Y) и ее составляющих, а также корреляционную матрицу, если плотность распределения вероятностей

$$f(x, y) = \frac{2}{\pi(x^2 + y^2 + 1)^2}.$$

3.20. При нескольких заездах автомобиля на мерном участке его левые и правые колеса наезжают на короткие и длинные неровности. Вероятность того, что при одном заезде автомобиль наедет на короткую неровность левыми колесами, равна 0,4, а правыми – 0,05; вероятность того, что автомобиль наедет на длинную неровность левыми колесами, равна 0,1, а правыми – 0,45. Найти математические ожидания и дисперсии числа наездов на короткие и длинные неровности при одном заезде.

3.21. Определить плотность распределения вероятностей, математические ожидания и корреляционную матрицу системы СВ (X, Y), если функция распределения системы

$$F(x, y) = \begin{cases} \sin x \sin y & \text{при } 0 \leqslant x \leqslant \pi/2,\ 0 \leqslant y \leqslant \pi/2, \\ 0 & \text{в других случаях.} \end{cases}$$

3.5. ДВУХМЕРНОЕ НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Распределение системы двухмерной СВ называют *нормальным*, если плотность распределения вероятностей имеет вид

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} \exp\left(-\frac{1}{2(1-r^2)}\left(\frac{(x-m_x)^2}{\sigma_x^2} + \frac{(y-m_y)^2}{\sigma^2 y} - \frac{2r(x-m_x)(y-m_y)}{\sigma_x\sigma_y}\right)\right), \qquad (3.6)$$

где σ_x, σ_y – средние квадратичные отклонения составляющих; m_x, m_y – математические ожидания; $r = r_{xy}$ – коэффициент корреляции.

Если $r_{xy} = 0$, то формула (3.6) примет вид

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{(x-m_x)}{2\sigma_x^2} - \frac{(y-m_y)^2}{2\sigma_y^2}\right).$$

Плотность распределения вероятностей $f_1(x)$ составляющей X

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \frac{1}{\sigma_x\sqrt{2\pi}} \exp\left(-\frac{(x-m_x)^2}{2\sigma_x^2}\right).$$

Аналогично определяется плотность распределения вероятностей для составляющей Y. Условная плотность распределения вероятностей

$$f(x/Y=y) = \frac{1}{\sigma_x\sqrt{2\pi(1-r^2)}}\exp\left(-\frac{\left(x-m_x-r\frac{\sigma_x}{\sigma_y}(y-m_y)\right)^2}{2(1-r^2)\,\sigma_x^2}\right),$$

$$f(y/X=x) = \frac{1}{\sigma_y\sqrt{2\pi(1-r^2)}}\exp\left(-\frac{\left(y-m_y-r\frac{\sigma_y}{\sigma_x}(x-m_x)\right)^2}{2(1-r^2)\,\sigma_y^2}\right).$$

Вероятность попадания случайной точки, распределенной по нормальному закону, в прямоугольник R со сторонами, параллельными координатным осям, вершины которого имеют координаты (α,γ), (α,δ), (β,γ), (β,δ), выражается формулой

$$P((x,y)\in R) = \left[\Phi\left(\frac{\beta-m_x}{\gamma_x}\right)-\Phi\left(\frac{\alpha-m_x}{\gamma_x}\right)\right]\left[\Phi\left(\frac{\delta-m_y}{\gamma_y}\right)-\right.$$

$$\left.-\Phi\left(\frac{\gamma-m_y}{\gamma_y}\right)\right].$$

Эллипсом рассеивания называется эллипс, во всех точках которого плотность распределения $f(x,y)$ СВ, распределенной по нормальному закону, постоянна:

$$f(x,y) = \text{const}.$$

Полуоси эллипса рассеивания пропорциональны σ_x, σ_y: $a = k\sigma_x$, $b = k\sigma_y$. Вероятность попадания случайной точки в область E_k, ограниченную эллипсом рассеивания,

$$P((x,y)\in E_k) = 1 - e^{-k^2/2}.$$

Если в формуле (3.6) положить $m_x = m_y = 0$, $\sigma_x = \sigma_y = \sigma$, получим

$$f(x,y) = \frac{1}{2\pi\sigma^2}\,e^{-\frac{1}{2}\frac{x^2+y^2}{\sigma^2}}$$

или

$$f(z) = \frac{1}{2\pi\sigma^2}\,e^{-z^2/(2\sigma^2)},$$

где $z = x^2 + y^2$.

Распределение, для которого

$$f(z) = \begin{cases} \dfrac{z}{\sigma^2} e^{-z^2/(2\sigma^2)} & \text{при } z > 0, \\ 0 & \text{при } z < 0, \end{cases}$$

называется *распределением Рэлея.*

Функция распределения Рэлея имеет вид

$$F(z) = \frac{1}{\sigma^2} \int_0^r ze^{-z^2/(2\sigma^2)} dr = 1 - e^{-z^2/(2\sigma^2)}.$$

Пример 3.8. Производится штамповка детали, имеющей форму эллипса

$$(x-1)^2 + \frac{(y+1)^2}{4} = 1.$$

Отклонение оси пуансона в результате износов распределено по нормальному закону с параметрами $m_x = 1$, $m_y = 1$, $\sigma_x = 1$, $\sigma_y = 2$, $r_{xy} = 0$. Найти вероятность того, что деталь из-под штампа выйдет годной.

Р е ш е н и е. Область D штампованной детали ограничена эллипсом рассеивания E_1 с полуосями $a = \sigma_x = 1$, $b = \sigma_y = 2$, а вероятность попадания в эту область

$$P((x, y) \in E_1) = 1 - e^{-1/2} \cong 0{,}393.$$

ЗАДАЧИ

3.22. Координаты (X, Y) случайной точки A на плоскости подчинены нормальному закону распределения

$$f(x, y) = \frac{1}{2\pi ab} e^{-\frac{1}{2}\left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right)}.$$

Определить вероятность того, что точка A окажется внутри эллипса с полуосями ka и kb, совпадающими с координатными осями Ox и Oy.

3.23. Система СВ (X, Y) распределена по закону

$$f(x, y) = \frac{a}{1 + x^2 + x^2y^2 + y^2}.$$

Вычислить a, установить, являются ли величины X, Y зависимыми, найти $f_1(x), f_2(x)$, найти вероятность попадания СВ (случайной точки) в квадрат S, центр которого совпадает с началом координат, а стороны параллельны осям координат и равны 2.

3.24. Независимые СВ X, Y распределены по нормальным законам с параметрами $m_x = 2$, $m_y = -3$, $\sigma_x = 1, \sigma_y = 2$. Вычислить вероятности следующих событий: 1) $(X < m_x)$ $(Y < m_y)$; 2) $X < 3$; 3) $Y < (x-5)$; 4) $|X| < 1$; 5) $(|X| < 1)$ $(|Y| < 2)$.

3.25. Плотность распределения вероятностей системы двух СВ (X, Y) задана выражением

$$f(x,y) = a\exp\left(-\frac{(x+3)^2}{8} - \frac{(y-1)^2}{8}\right).$$

Найти коэффициент a. Установить, являются ли СВ X и Y зависимыми. Определить вероятность одновременного выполнения неравенств $X < -3$, $Y < 4$.

3.26. Система двух СВ (X, Y) распределена по нормальному закону с параметрами $m_x = m_y = 0$, $\sigma_x = \sigma_y = \sigma$, $r_{xy} = 0$. Определить вероятность следующих событий:
1) $|Y| < X$; 2) $Y < x$; 3) $Y < |x|$.

3.27. Случайная величина эксцентриситета детали характеризуется функцией распределения Рэлея

$$F(z) = \frac{1}{\sigma^2}\int_0^r z e^{-z^2/(2\sigma^2)}\,dz = 1 - e^{-z^2/(2\sigma^2)}.$$

Найти плотность распределения вероятностей $f(z)$.

3.28. Плотность распределения вероятностей случайных амплитуд A боковой качки корабля определяется *законом Рэлея*

$$f(a) = \frac{a}{\sigma^2}\, e^{-a^2/(2\sigma^2)} \quad (a > 0),$$

где σ^2 – дисперсия угла крена. Как часто встречаются амплитуды, меньшие и большие средней?

4. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

4.1. СХОДИМОСТЬ СЛУЧАЙНЫХ ВЕЛИЧИН

При достаточно большом числе испытаний некоторые характеристики случайных событий и величин обладают свойством устойчивости. Например, это свойство характерно для относительных частот. При большом числе испытаний среднее арифметическое $\bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k$ также обладает свойством устойчивости: $\bar{x} \to M(X)$ при $n \to \infty$.

Группа теорем, устанавливающих соответствие между теоретическими и эмпирическими характеристиками СВ и случайных событий при большом числе испытаний в одинаковых условиях, объединена под общим названием *"предельные теоремы теории вероятностей"*.

Сходимость СВ, используемых при изучении предельных теорем теории вероятностей, рассмотрим на примере последовательности СВ $X_1, X_2, \ldots, X_n, \ldots$, определенной на вероятностном пространстве (Ω, A, P). Так как СВ, образующие эту последовательность, являются функциями элементарных исходов (событий) $\omega \in \Omega$, то, выбрав $\omega_0 \in \Omega$, получим числовую последовательность $X_1(\omega_0), X_2(\omega_0), \ldots, X_n(\omega_0), \ldots$, которая может сходиться или расходиться.

Рассмотрим некоторое множество элементарных событий $A \subset \Omega$, которому соответствуют сходящиеся числовые последовательности. Пусть $X(\omega)$ – предел числовой последовательности: $\lim X_n(\omega) = X(\omega)$ для любого $\omega \in A$. В этом случае говорят, что *последовательность СВ X_n, $n = 1,2,\ldots$, сходится по распределению к X на множестве A.*

Последовательность СВ X_n, $n = 1, 2, \ldots$, *сходится по вероятности к случайной или неслучайной величине X*, если для любого

$$\epsilon > 0 \quad \lim_{n \to \infty} P\{|X_n - X| < \epsilon\} = 1.$$

Примером сходимости по распределению является сходимость биномиального распределения к распределению Пуассона.

Ряд теорем, связывающих различные виды сходимости, составляют содержание *закона больших чисел.*

4.2. НЕРАВЕНСТВО ЧЕБЫШЕВА

Если СВ X имеет конечные дисперсию $D(X)$ и математическое ожидание $M(X)$, то при любом $\epsilon > 0$ справедливо неравенство, называемое *неравенством Чебышева: вероятность того, что отклонение СВ от ее математического ожидания будет по абсолютной величине не меньше любого положительного числа ϵ, ограничена сверху величиной $D(X)/\epsilon^2$:*

$$P(|X - M(X)| \geqslant \epsilon) \leqslant D(X)/\epsilon^2.$$

Так как $\{|X - M(X)| \geqslant \epsilon\} \cup \{|X - M(X)| < \epsilon\} = \Omega$, то неравенство Чебышева можно записать в виде

$$P(|X - M(X)| < \epsilon) \geqslant 1 - D(X)/\epsilon^2. \qquad (4.1)$$

С помощью неравенства (4.1) можно оценить сверху вероятность того, что СВ X отклонится от $M(X)$ не менее, чем на $3\sigma(X)$:

$$P(|X - M(X)| \geqslant 3\sigma(X)) \leqslant \frac{D(X)}{9\sigma^2(X)} = \frac{1}{9}.$$

Неравенство Чебышева имеет теоретическое значение, оно служит основой для ряда теорем, входящих в закон больших чисел.

4.3. ТЕОРЕМА ЧЕБЫШЕВА

Теорема Чебышева является одной из важнейших форм закона больших чисел. Она устанавливается между средним арифметическим СВ и средним арифметическим их математических ожиданий. Сформулируем ее.

Теорема Чебышева. *При неограниченном увеличении числа независимых опытов среднее арифметическое наблюденных значений СВ, имеющей ограниченную дисперсию $D(X)$, сходится по вероятности к ее математическому ожиданию:*

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} x_k - M(X)\right| \leqslant \epsilon\right) = 1.$$

При доказательстве теоремы Чебышева с помощью неравенства Чебышева получаем оценку

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} x_k - M(X)\right| < \epsilon\right) \geqslant 1 - \frac{D(X)}{n\epsilon^2}. \qquad (4.2)$$

Теорема Чебышева распространяется и на более сложную ситуацию, когда распределение СВ от опыта к опыту изменяется. В этом случае имеет место

Обобщенная теорема Чебышева. *Если $X_1, X_2, \ldots, X_n$ – последовательность независимых СВ с математическими ожиданиями $M(X_1), M(X_2), \ldots, M(X_n)$ и дисперсиями $D(X_1), D(X_2), \ldots, D(X_n)$, ограниченными одной и той же постоянной $|D(X)| \leqslant C$, то для любого $\epsilon > 0$*

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} M(X_k)\right| < \epsilon\right) = 1. \qquad (4.3)$$

При доказательстве предельного равенства (4.3) с помощью неравенства Чебышева получаем оценку

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} M(X_k)\right| < \epsilon\right) \geqslant 1 - \frac{C}{n\epsilon^2}. \qquad (4.4)$$

Частным случаем теоремы Чебышева являются *теоремы Бернулли* и *Пуассона*. Связь между частотой события и его вероятностью устанавливает

Теорема Бернулли. Если в каждом из n независимых испытаний вероятность появления события A постоянна и равна p, то относительная частота сходится по вероятности к p при $n \to \infty$, т.е.

$$\lim_{n \to \infty} P\left(\left|\frac{m}{n} - p\right| < \epsilon\right) = 1.$$

При доказательстве этой теоремы получаем оценку

$$P\left(\left|\frac{m}{n} - p\right| < \epsilon\right) \geqslant 1 - \frac{pq}{n\epsilon^2}, \tag{4.5}$$

которая применяется на практике. Теорема Бернулли утверждает устойчивость частоты при постоянных условиях опыта.

Устойчивость частоты при переменных условиях опыта устанавливает

Теорема Пуассона. Если производится n независимых опытов и вероятность появления события A в k-м опыте равна p_k, то при увеличении n частота m/n события A сходится по вероятности к среднему арифметическому вероятностей p_k:

$$\lim_{n \to \infty} P\left(\left|\frac{m}{n} - \frac{1}{n}\sum_{k=1}^{n} p_k\right| < \epsilon\right) = 1,$$

где ϵ — сколь угодно малое положительное число.

4.4. ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

Ранее были рассмотрены вопросы приближения некоторых СВ к определенным предельным значениям независимо от их закона распределения. В теории вероятностей существует другая группа теорем, касающихся предельных распределений СВ $X_n = \sum_{k=1}^{n} X_k$ в случае, когда число слагаемых неограниченно возрастает. Она носит общее название *центральной предельной теоремы* и указывает условия, при которых закон распределения суммы большого числа независимых СВ близок к нормальному.

Теорема Ляпунова. Пусть X_n, $n = 1, 2, \ldots$, — последовательность независимых СВ, для каждой из которых существует математическое ожидание $M(X_k) = m_k$, дисперсия $D(X_k) = \sigma_k^2$, третий центральный момент $M(X_k - m_k)^3$ и выполняется условие Ляпунова

$$\lim_{n \to \infty} \frac{\sum_{k=1}^{n} M|X_k - m_k|^3}{\left(\sum_{k=1}^{n} \sigma_k^2\right)^{3/2}} = 0.$$

Тогда при $n \to \infty$ распределение СВ $X_n = \sum_{k=1}^{n} X_k$ сходится по распределению к нормальному

$$P(X_n < x) = \frac{1}{\sqrt{2\pi}\sigma_{x_n}} \int\limits_{-\infty}^{x} e^{-(X_n - M(X_n))^2/(2\sigma^2_{x_n})} dx,$$

где $M(X_n) = \sum\limits_{k=1}^{n} M(X_k)$; $\sigma^2_{x_n} = \sum\limits_{k=1}^{n} \sigma^2_k$.

Опыт показывает, что распределение суммы независимых СВ, сравниваемых по своему рассеиванию, уже при числе слагаемых порядка десяти приближается к нормальному.

Пример 4.1. При изготовлении конической поверхности детали в среднем 75 % деталей укладывается в поле допуска. Оценить вероятность того, что среди 2000 деталей в поле допуска окажутся от 1450 до 1550 деталей включительно.

Р е ш е н и е. Оценить вероятность можно с помощью неравенства Чебышева (4.1). Искомое число изделий является СВ X, распределенной по биномиальному закону с $M(X) = 2000 \cdot 0{,}75 = 1500$, $D(X) = 2000 \cdot 0{,}75 \cdot 0{,}25 = 375$. Так как 1450 и 1550 – границы допустимых значений СВ – симметричны относительно математического ожидания, равного 1500, неравенство $1450 < X < 1550$ можно заменить эквивалентным ему: $|X - 1500| < 50$. При $\epsilon = 50$, используя неравенство (4.1), получаем

$$P(1450 < X < 1550) = P(|X - 1500| < 50) > 1 - 375/50^2 = 0{,}85.$$

Итак, вероятность искомого события не меньше 0,85.

Пример 4.2. Сколько деталей необходимо проверить, чтобы с вероятностью, не меньшей 0,95, можно было утверждать, что абсолютная величина отклонения частоты годных деталей от вероятности, равной 0,9, не превысит 0,01 ?

Р е ш е н и е. Используя неравенство (4.5), имеем

$$1 - pq/(n\epsilon^2) > 0{,}95.$$

Решая данное неравенство относительно n, получаем

$$n > \frac{pq}{0{,}05\epsilon^2} = \frac{0{,}9 \cdot 0{,}1}{0{,}05 \cdot (0{,}01)^2} = \frac{0{,}09}{0{,}000005} = 18\,000,$$

т.е. наименьшее число деталей, которые следует проверить, равно 18 000.

Пример 4.3. Вероятность наступления события в каждом испытании равна 0,3 . Найти вероятность того, что в 10 000 испытаний отклонение относительной частоты события от его вероятности не превзойдет по абсолютной величине 0,01.

Р е ш е н и е. Используя неравенство Чебышева (4.1), можем записать:

$$P(|m/n - p| < \epsilon) = 1 - pq/(n\epsilon^2).$$

По условию задачи имеем: $n = 10\,000$, $p = 0{,}3$, $q = 1 - 0{,}3 = 0{,}7$, $\epsilon = 0{,}01$. Подставив эти значения в последнее неравенство, получим

$$P\left(\left|\frac{m}{n} - 0{,}3\right| < 0{,}01\right) = 1 - \frac{0{,}3 \cdot 0{,}7}{10\,000 \cdot 0{,}0001} = 1 - 0{,}3 \cdot 0{,}7 = 0{,}79.$$

Пример 4.4. Среднее квадратичное отклонение каждой из 2500 независимых СВ не превосходит 3. Найти вероятность того, что абсолютная величина отклонения среднего арифметического этих СВ от среднего арифметического их математических ожиданий не превзойдет 0,3.

Р е ш е н и е. Согласно теореме Чебышева, можно записать:

$$P(|X - M(X)| < \epsilon) > 1 - \frac{9}{2500 \cdot 0{,}09} = 1 - \frac{1}{25} = 0{,}96.$$

Здесь $n = 2500$, $C = D(X) = \sigma^2 = 9$, $\epsilon = 0{,}3$.

ЗАДАЧИ

4.1. Завод выпускает 90 % изделий первого сорта. Наугад выбирают 1000 изделий. Найти вероятность того, что число первосортных изделий окажется в пределах от 900 до 940.

4.2. Дисперсия каждой из 800 независимых СВ не превышает 9. Какой должна быть верхняя граница абсолютной величины отклонения среднего арифметического СВ от среднего арифметического их математических ожиданий, чтобы вероятность такого отклонения превышала 0,997?

4.3. Определить число испытаний, которые нужно провести, чтобы отклонение частоты появления события A от его средней вероятности в проведенных испытаниях не превышало по абсолютной величине 0,02 с вероятностью 0,99.

4.4. Из 4000 деталей для 500 вероятность появления детали номинального размера составляет 0,4, для 1200–0,5 и для 2300–0,7. Найти границы, в которых должна находиться частость появления ожидаемого результата с гарантированной вероятностью 0,99.

4.5. Вероятность положительного исхода в отдельном испытании равна 0,8. Оценить вероятность того, что при 1000 независимых повторных испытаний отклонение частости положительных исходов от вероятности при отдельном испытании по своей абсолютной величине будет меньше 0,05.

4.6. Вероятность наличия внутренних трещин в отливках равна 0,2. Оценить вероятность того, что из 1000 отливок отклонение числа годных от 800 не превысит 5 %.

4.7. Вероятность появления события в каждом испытании равна 1/4. Найти вероятность того, что число X появлений события заключено в пределах от 150 до 250, если производится 800 испытаний.

4.8. Дискретная СВ X задана рядом распределения

x_i	0,1	0,4	0,6
p_i	0,2	0,3	0,5

Пользуясь неравенством Чебышева, найти вероятность того, что $|X - M(X)| < \sqrt{0{,}4}$.

4.9. Распределение СВ X задается следующим рядом распределения:

x_i	–1	0	2	4	6
p_i	0,2	0,4	0,3	0,05	0,05

Чему равна вероятность того, что $|X - M(X)| < 5$? Оценить эту вероятность, пользуясь неравенством Чебышева.

4.10. Среднее значение длины детали 50 см, дисперсия 0,1. Пользуясь неравенством Чебышева, оценить вероятность того, что длина случайно взятой детали окажется в поле допуска (не меньше 49,5 см и не больше 50,5 см).

5. ЭЛЕМЕНТЫ ТЕОРИИ СЛУЧАЙНЫХ ПРОЦЕССОВ

5.1. ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

Случайной называется функция, которая в результате опыта может принять тот или иной вид (неизвестный заранее), или функция, значение которой при каждом значении аргумента является СВ. Конкретный вид, принимаемый случайной функцией в результате опыта, называется *реализацией случайной функции*. *Случайным процессом* является случайная функция $X(t)$ от независимой переменной t.

Двухмерным законом распределения случайной функции $X(t)$ называется совместный закон распределения ее значений $X(t)$, $X(t')$ при двух произвольно взятых значениях t и t' аргумента t: $f_2(x, x', t, t')$, где $x = X(t)$, $x' = X(t')$; *n-мерным законом распределения случайной функции* $X(t)$ называется закон распределения совокупности значений $X(t_1), X(t_2), \ldots, X(t_n)$ случайной функции $X(t)$ при n произвольно взятых значениях аргумента $t_1, t_2, \ldots, t_n$: $f_n(x_1, x_2, \ldots, x_n, t_1, t_2, \ldots, t_n)$.

Сечением случайной функции называется ее значение при некотором фиксированном значении аргумента.

Математическим ожиданием случайной функции $X(t)$ называется такая неслучайная функция, значение которой при каждом значении аргумента равно математическому ожиданию соответствующего сечения случайной функции:

$$M(X(t)) = \int_{-\infty}^{\infty} x f_1(x, t) dx = m_x(t), \tag{5.1}$$

где $f_1(x, t)$ – одномерный закон распределения.

Дисперсией случайной функции $X(t)$ называется такая неслучайная функция, значение которой при каждом значении аргумента равно дисперсии соответствующего сечения случайной функции: $D(X(t)) = D_x(t)$ или, в интегральном выражении,

$$D(X(t)) = \int_{-\infty}^{\infty} (x(t) - m_x(t))^2 f_1(x, t)\, dx . \tag{5.2}$$

Среднее квадратичное отклонение случайной функции равно квадратному корню из дисперсии случайной функции:

$$\sigma_x(t) = \sqrt{D(x(t))} . \tag{5.3}$$

Корреляционной функцией случайной функции $X(t)$ называется такая неслучайная функция двух аргументов $K_x(t, t')$, значение которой при каждой паре значений t и t' равно корреляционному моменту соответствующих сечений случайной функции, т.е.

$$K_x(t, t') = M((X(t) - m_x(t))(X(t') - m_x(t')))$$

или

$$K_x(t,t') = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} \overset{0}{x}(t)\,\overset{0}{x}(t') f_2(x,x',t,t')\,dx\,dx',$$

где $\overset{0}{x}(t) = X(t) - m_x(t)$; $\overset{0}{x}(t') = X(t') - m_x(t')$; $f_2(x,x',t,t')$ — двухмерный закон распределения.

Приведем *свойства корреляционной функции.*

1. *Функция* $K_x(t, t')$ *симметрична относительно своих аргументов*: $K_x(t,t') = K_x(t',t)$.

2. $|K_x(t,t')| \leqslant \sigma_x(t)\,\sigma_x(t')$.

3. *Если к случайной функции прибавить неслучайную, то ее корреляционная функция не изменится.*

4. *Если случайную функцию умножить на неслучайную* $\varphi(t)$, *то корреляционная функция случайной функции умножится на* $\varphi(t)\,\varphi(t')$.

5. $K_x(t,t) = D_x(t)$.

Взаимной корреляционной функцией или *корреляционной функцией связи* двух случайных функций $X(t)$ и $Y(s)$ называется такая неслучайная функция двух аргументов $K_{xy}(t,s)$, значение которой при каждой паре значений t и s равно корреляционному моменту соответствующих сечений случайных функций, т.е.

$$K_{xy}(t,s) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} \overset{0}{x}\,\overset{0}{y}\, f(x,y,t,s)\,dx\,dy.$$

Функции $X(t)$ и $Y(s)$ называются *коррелируемыми*, если $K_{xy}(t,s) \neq 0$, и *некоррелируемыми*, если $K_{xy}(t,s) = 0$ при всех значениях t и s.

Пример 5.1. Плотность распределения вероятностей случайной функции $X(t)$

$$f(x,t) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-(x-a\sin t)^2/(2\sigma^2)}.$$

Найти математическое ожидание и дисперсию случайной функции, если a и σ – постоянные, причем $\sigma > 0$.

Р е ш е н и е. Согласно формуле (5.1),

$$m_x(t) = \int\limits_{-\infty}^{\infty} x\,\frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-(x-a\sin t)^2/(2\sigma^2)}\,dx =$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma}\int\limits_{-\infty}^{\infty} x e^{-(x-a\sin t)^2/(2\sigma^2)}\,dx.$$

Сделаем замену переменной $x - a\sin t/\sigma = z$, $x = \sigma z + a\sin t$, $dx = \sigma\,dz$. Тогда математическое ожидание

$$m_x(t) = \frac{\sigma}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} z e^{-z^2/2}\,dz + \frac{a\sin t}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{-z^2/2}\,dz.$$

Так как

$$\int\limits_{-\infty}^{\infty} z e^{-z^2/2} dz = 0, \quad \int\limits_{-\infty}^{\infty} e^{-z^2/2} dz = \sqrt{2\pi}\ ,$$

то $m_x(t) = a \sin t$.

Находим дисперсию:

$$D_x(t) = \int\limits_{-\infty}^{\infty} (x - a\sin t)^2 \frac{1}{\sqrt{2\pi}\ \sigma} e^{-(x-a\sin t)^2/(2\sigma^2)} dx\ .$$

Используя замену переменной, приведенную выше, получаем

$$D_x(t) = \frac{1}{\sqrt{2\pi}\,\sigma} \int\limits_{-\infty}^{\infty} \sigma^2 z^2 e^{-z^2/(2\sigma)} dz\ .$$

Интегрируя последнее выражение по частям и учитывая формулу (5.2), имеем $D_x(t) = \sigma^2$.

Пример 5.2. Двухмерный закон распределения случайной функции имеет вид

$$f_2(x_1, x_2, t_1, t_2) = \frac{1}{2\pi} \exp\left(-\frac{(x_1 + \sin t_1)^2 + (x_2 + \sin t_2)^2}{2}\right).$$

Найти математическое ожидание, дисперсию и корреляционную функцию.

Р е ш е н и е. Найдем математическое ожидание случайной функции $f_1(x_1, t_1)$. Для этого сначала запишем одномерный закон распределения:

$$f_1(x_1, t_1) = \int\limits_{-\infty}^{\infty} f_2(x_1, x_2, t_1, t_2)\, dx_2\ .$$

В данном случае

$$f_1(x_1, t_1) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \exp\left(-\frac{(x_1 + \sin t_1)^2 + (x_2 + \sin t_2)^2}{2}\right) dx_2 =$$

$$= \frac{1}{2\pi} e^{-\frac{(x_1 + \sin t_1)^2}{2}} \int\limits_{-\infty}^{\infty} e^{-\frac{(x_2 + \sin t_2)^2}{2}} dx =$$

$$\left| \begin{array}{l} (x^2 + \sin t_2)/\sqrt{2} = u, \\ dx_2 = \sqrt{2}\, du \end{array} \right| = \frac{1}{\pi\sqrt{2}} e^{-\frac{(x_1 + \sin t_1)^2}{2}} \int\limits_{-\infty}^{\infty} e^{-u^2} du = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_1 + \sin t_1)^2}{2}}$$

Математическое ожидание

$$m_x(t) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} x e^{-(x + \sin t)^2/2} dx = \left| \begin{array}{l} (x + \sin t)/\sqrt{2} = u\ ,\quad x = \sqrt{2}\, u - \sin t, \\ dx = \sqrt{2}\, du \end{array} \right| =$$

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (\sqrt{2}\, u - \sin t) e^{-u^2} du = \frac{\sqrt{2}}{\sqrt{\pi}} \int_{-\infty}^{\infty} u e^{-u^2} du -$$

$$- \frac{\sin t}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-u^2} du = -\sin t .$$

Находим корреляционную функцию:

$$K_x(t_1, t_2) = \int_{-\infty}^{\infty} (x_1 + \sin t_1)(x_2 + \sin t_2) \frac{1}{2\pi} e^{-((x_1+\sin t_1)^2 + (x_2 + \sin t_2)^2)/2} dx_1 dx_2 =$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} (x_1 + \sin t_1) e^{-(x_1+\sin t_1)^2/2} dx_1 \int_{-\infty}^{\infty} (x_2 + \sin t_2) e^{-(x_2+\sin t_2)^2/2} dx_2 = 0.$$

Дисперсия $D_x(t) = 1$.

ЗАДАЧИ

5.1. Двухмерная плотность распределения вероятностей $f(x_1, x_2, t_1, t_2)$ случайной функции $X(t)$ равна $\frac{1}{2\pi\sigma^2} e^{(-x_1^2+x_2^2)/(2\sigma^2)}$. Найти математическое ожидание, дисперсию и корреляционную функцию.

5.2. Одномерный закон распределения $f_1(x_1, t)$ случайной функции $X(t)$ имеет вид $\frac{1}{\sqrt{2\pi}\sigma} e^{-(x-a\cos t)^2/(2\sigma^2)}$. Найти математическое ожидание и дисперсию.

5.3. Двухмерная плотность распределения вероятностей $f_2(x_1, x_2, t_1, t_2)$ случайной функции $X(t)$ имеет вид

$$f_2(x_1, x_2, t_1, t_2) = \begin{cases} x_1 x_2 e^{-(x_1^2+x_2^2)/2} & \text{при } x_1 \geqslant 0,\ x_2 \geqslant 0, \\ 0 & \text{при } x_1 < 0,\ x_2 < 0. \end{cases}$$

Найти математическое ожидание, дисперсию и корреляционную функцию случайной функции $X(t)$.

5.4. Дана двухмерная плотность распределения вероятностей случайной функции $X(t)$:

$$f_2(x_1, x_2, t_1, t_2) = \frac{1}{2\pi} \exp\left(-\frac{(x_1 - t_1)^2 - (x_2 - t_2)^2}{2}\right).$$

Найти математическое ожидание, дисперсию и корреляционную функцию случайной функции $X(t)$.

5.2. ОПЕРАЦИИ НАД СЛУЧАЙНЫМИ ФУНКЦИЯМИ

Пусть случайная функция имеет вид $Z(t) = X(t) + Y(t)$. Тогда ее математическое ожидание

$$M(Z(t)) = M(X(t)) + M(Y(t)),$$

а корреляционная функция

$$K_z(t, t') = K_{xy}(t, t') + K_{yx}(t, t') + K_y(t, t') + K_x(t, t').$$

Если $X(t)$ и $Y(t)$ – некоррелируемые, то

$$K_z(t, t') = K_x(t, t') + K_y(t, t').$$

Если $Z(t) = \varphi(t)\, X(t)$, где $\varphi(t)$ – неслучайная функция, то

$$M(Z(t)) = \varphi(t) M(X(t)),\ K_z(t, t') = \varphi(t)\, \varphi(t') K_x(t, t').$$

Если $Y(t) = \dfrac{dX(t)}{dt}$, то

$$M(Y(t)) = \frac{dm_x(t)}{dt} = m'_x(t),$$

$$K_y(t, t') = \frac{d^2 K_x(t, t')}{dt\, dt'}.$$

Если же $Y(t) = \int\limits_0^t X(\tau)\, d\tau$, то

$$M_y(t) = \int\limits_0^t m_x(\tau)\, d\tau, \quad K_y(t, t') = \int\limits_0^t \int\limits_0^{t'} K_x(\tau, \tau')\, d\tau\, d\tau'.$$

Пример 5.3. На вход интегратора поступает случайная функция $X(t)$, математическое ожидание которой $m_x(t) = 4t + 5$, а корреляционная функция $K_x(t_1, t_2) = \cos t_1 \cos t_2$. Найти характеристики на выходе системы.

Решение. Пусть $Y(t) = \int\limits_0^t X(\tau)\, d\tau$. Тогда:

$$m_y(t) = \int\limits_0^t m_x(t')\, dt' = \int\limits_0^t (4t' + 5)\, dt' = 2t^2 + 5t,$$

$$K_y(t_1, t_2) = \int\limits_0^{t_1} \int\limits_0^{t_2} \cos t'_1 \cos t'_2\, dt'_1\, dt'_2 = \sin t_1 \sin t_2.$$

Дисперсия $D_y(t) = \sin^2 t$.

ЗАДАЧИ

5.5. Случайная функция $X(t)$ имеет следующие характеристики:

$$M_x(t) = t^3 + 3t, \; K_x(t_1, t_2) = e^{-t_1^2 - t_2^2}.$$

Найти характеристики случайной функции $Y_1(t) = t^2 \dfrac{dx}{dt} - 3t^2$.

5.6. Характеристики случайной функции $X(t)$ заданы выражениями: $m_x(t) = t+4$, $K_x(t_1, t_2) = t_1 t_2$. Найти характеристики случайной функции $Z(t) = 5tX(t)+1$.

5.7. Работа динамической системы описывается оператором вида $Y(t) = \dfrac{2}{3t}\int\limits_0^t X(\tau)\,d\tau + t^2$. На вход этой системы поступает случайная функция $X(t)$ с характеристиками: $m_x(t) = 2e^t$, $K_x(t_1, t_2) = t_1 t_2 e^{t_1} e^{t_2}$. Найти характеристики на выходе системы.

5.8. На вход динамической системы, описываемой оператором $Y(t) = \dfrac{dX}{dt}$, поступает случайная функция $X(t)$, математическое ожидание которой $m_x(t) = A\sin t$, а корреляционная функция $K_x(t_1, t_2) = De^{-\alpha(t_2 - t_1)^2}$, где $D = \text{const}$. Определить характеристики на выходе системы.

5.3. КАНОНИЧЕСКИЕ РАЗЛОЖЕНИЯ

Элементарной случайной функцией называется случайная функция вида $X(t) = X\varphi(t)$, где X – случайная величина, математическое ожидание которой равно 0, а $\varphi(t)$ – неслучайная функция. Математическое ожидание элементарной случайной функции равно нулю, а корреляционная функция имеет вид $K_x(t, t') = \varphi(t)\varphi(t')D_x$. Производная элементарной случайной функции $X'(t) = X\varphi'(t)$.

Каноническим разложением случайной функции $X(t)$ называется представление этой функции в виде суммы ее математического ожидания и взаимно некоррелируемых элементарных случайных функций:

$$X(t) = m_x(t) + \sum_{i=1}^{k} X_i \varphi_i(t), \tag{5.4}$$

где X_i – случайные величины, называемые *коэффициентами канонического разложения;* $\varphi_i(t)$ – неслучайные функции, которые называются *координатными функциями канонического разложения.*

Если случайная функция $X(t)$ представлена в виде (5.4), то

$$K_x(t, t') = \sum_{i=1}^{k} D_{x_i}(t)\,\varphi_i(t)\,\varphi_i(t'), \tag{5.5}$$

где D_{x_i} – дисперсии случайных величин X_i. Справедливо и обратное утверждение.

Пример 5.4. Корреляционная функция случайной функции $X(t)$ задана каноническим разложением

$$K_x(t_1, t_2) = 3t_1t_2 + t_1^2t_2^2 + 5t_1^3t_2^3.$$

Найти каноническое разложение случайной функции.

Решение. Из равенства (5.4) находим

$$\overset{0}{X}(t) = X(t) - m_x(t) = \sum_{i=1}^{k} X_i\varphi_i(t).$$

Согласно формулам (5.4) и (5.5), в рассматриваемом случае имеем:

$$D(X_1) = 3, \quad D(X_2) = 1, \quad D(X_3) = 5.$$

Следовательно, $\overset{0}{X}(t) = X_1t + X_2t^2 + X_3t^3$.

ЗАДАЧИ

5.9. Случайная функция $X(t)$ задана каноническим разложением $X(t) = \sin t + X_1 + X_2t + X_3t\sin t$; дисперсии случайных величин X_1, X_2, X_3 равны соответственно D_1, D_2, D_3. Найти математическое ожидание, дисперсию и корреляционную функцию случайной функции $X(t)$.

5.10. Корреляционная функция случайной функции $X(t)$ задана каноническим разложением $K_x(t_1, t_2) = 2t_1t_2 + 3t_1^2t_2^2$. Найти каноническое разложение случайной функции.

5.11. На вход динамической системы, работа которой описывается оператором вида $Y_1(t) = 2t\dfrac{dX}{dt} + 3t^2$, поступает случайная функция, заданная каноническим разложением. Найти характеристики случайной функции $X(t)$ на выходе системы.

5.12. Случайная функция $X(t)$ задана каноническим разложением $X(t) = t + X_1\cos 2t + X_2\sin 2t$. Дисперсии $D_{x_1} = D_{x_2} = 2$. Найти каноническое разложение и характеристики случайной функции $Z(t) = 3tX(t) + 2t^2$.

5.4. СТАЦИОНАРНЫЕ СЛУЧАЙНЫЕ ФУНКЦИИ

Случайная функция $X(t)$ называется *стационарной в широком смысле*, если ее математическое ожидание является постоянной величиной, а корреляционная функция зависит только от разности аргументов t_1 и t_2, т.е.

$$m_x(t) = \text{const}, \quad K_x(t_1, t_2) = K_x(\tau), \quad \tau = t_1 - t_2,$$
$$D(X(t)) = K_x(t_1, t_2) = K_x(0).$$

Нормированной корреляционной функцией $\rho_x(\tau)$ называют отношение

$$\frac{K_x(\tau)}{D(X(t))} = \frac{K_x(\tau)}{K_x(0)}.$$

Каноническое разложение стационарной случайной функции имеет вид

$$X(t) = m_x + \sum_{k=1}^{\infty} (v_k \cos\omega_k t + u_k \sin\omega_k t), \tag{5.6}$$

где u_k, v_k — взаимно некоррелируемые случайные величины, математическое ожидание которых равно нулю. Разложение (5.6) называется *спектральным*.

Спектральную плотность S_x рассматривают как "плотность энергии на единицу частоты", которую можно получить путем суммирования квадратов вещественных амплитуд по возможным значениям ω.

Корреляционная функция и спектральная плотность выражаются друг через друга следующим образом:

$$K_x(\tau) = \frac{1}{2}\int_{-\infty}^{\infty} S_x(\omega)\cos\omega\tau d\omega = \int_0^{\infty} S_x(\omega)\cos\omega\tau d\omega, \tag{5.7}$$

$$S_x(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} K_x(\tau)\cos\omega\tau d\tau = \frac{2}{\pi}\int_0^{\infty} K_x(\tau)\cos\omega\tau d\tau. \tag{5.8}$$

Если перейти к *нормированной спектральной плотности*

$$\rho_x(\omega) = S_x(\omega)/D_x$$

и *нормированной корреляционной функции*

$$r_x(\tau) = K_x(\tau)/D_x,$$

то эти функции будут выражаться друг через друга следующим образом:

$$\rho_x(\omega) = \frac{2}{\pi}\int_0^{\infty} r_x(\tau)\cos\omega\tau d\tau,$$

$$r_x(\tau) = \int_0^{\infty} \rho_x(\omega)\cos\omega\tau d\omega.$$

Пример 5.5. Спектральная плотность случайной функции $X(t)$

$$S(\omega) = \begin{cases} a & \text{при} \quad -b \leqslant \omega \leqslant b, \\ 0 & \text{при} \quad |\omega| > b. \end{cases}$$

Найти корреляционную функцию $K_x(\tau)$.

Р е ш е н и е. Согласно формуле (5.7),

$$K_x(\tau) = \frac{1}{2}\int_{-b}^{b} a\cos\omega\tau d\omega = \frac{a\sin b\tau}{\tau}.$$

Пример 5.6. Определить спектральную плотность $S(\omega)$, если корреляционная функция

$$K(\tau) = \begin{cases} \sigma^2(1-|\tau|) & \text{при } |\tau| \leqslant 1, \\ 0 & \text{при } |\tau| > 1. \end{cases}$$

Р е ш е н и е. Согласно формуле (5.8), имеем

$$S(\omega) = \frac{1}{\pi}\int_{-1}^{1} \sigma^2 (1-|\tau|)\cos\omega\tau\, d\tau = \frac{1}{\pi}\sigma^2 \int_{-1}^{1} (1-|\tau|)\cos\omega\tau\, d\tau =$$

$$= \frac{2\sigma^2}{\pi}\int_{0}^{1} (1-\tau)\cos\omega\tau\, d\tau = \frac{2\sigma^2}{\pi}\left(\int_{0}^{1}\cos\omega\tau\, d\tau - \int_{0}^{1}\tau\cos\omega\tau\, d\tau\right) =$$

$$= \frac{2\sigma^2}{\pi}\left(\frac{\sin\omega\tau}{\omega}\Big|_0^1 - \frac{\tau\sin\omega\tau}{\omega}\Big|_0^1 + \frac{1}{\omega}\int_{0}^{1}\sin\omega\tau\, d\tau\right) =$$

$$= \frac{2\sigma^2}{\pi}\left(\frac{\sin\omega}{\omega} - \frac{\sin\omega}{\omega} - \frac{1}{\omega^2}\cos\omega\tau\Big|_0^1\right) = \frac{2\sigma^2}{\pi\omega^2}(1-\cos\omega) =$$

$$= \frac{4\sigma^2}{\pi\omega^2}\sin\frac{\omega}{2} = \frac{\sigma^2}{\pi}\left(\frac{\sin(\omega/2)}{(\omega/2)}\right)^2.$$

ЗАДАЧИ

5.13. Нормированная корреляционная функция микропрофиля разбитой грунтовой дороги имеет вид

$$r_x(l) = 0{,}624 e^{-11|l|} + 0{,}356 e^{-0{,}15|l|}\cos 0{,}36 e,$$

где l – длина дороги. Найти нормированную спектральную плотность $\rho_x(\omega)$.

5.14. Нормированная корреляционная функция микропрофиля булыжной мостовой с удовлетворительным состоянием покрытия имеет вид

$$r_x(l) = e^{-0{,}11|l|}\cos 0{,}2381\, l.$$

Найти нормированную спектральную плотность $\rho_x(\omega)$.

5.15. Колебания автомобиля при движении по булыжному покрытию характеризуются нормированной спектральной плотностью

$$\rho_x(\omega) = \frac{0{,}4 v_a}{\omega^2 + 0{,}0154 v_a^2},$$

где v_a – поступательная скорость автомобиля. Найти нормированную корреляционную функцию.

5.16. Случайная функция имеет вид $X(t) = 1 + X_1\cos\omega_1 t + X_2\sin\omega_1 t + X_3\cos\omega_2 t + X_4\sin\omega_2 t$, дисперсии коэффициентов разложения $D_{x_1} = D_{x_2} = 1, D_{x_3} = D_{x_4} = 2$. Установить, является ли эта функция стационарной.

5.17. Работа динамической системы описывается дифференциальным уравнением $2Y'(t) + Y(t) = X'(t) + 3X(t)$. На вход системы поступает стационарная случайная функция $X(t)$, математическое ожидание которой $m_x = 1$, а корреляционная функция $K_x(\tau) = 2e^{-2(\tau)}$. Определить математическое ожидание и дисперсию случайной функции на выходе системы.

5.5. ПУАССОНОВСКИЙ ПРОЦЕСС. ПРОСТЕЙШИЙ ПОТОК ОДНОРОДНЫХ СОБЫТИЙ

Под *потоком событий* будем понимать последовательность событий, происходящих одно за другим в некоторые моменты времени (например, поток вызовов на телефонной станции, поток включений приборов в электросети и т.д.). Под *потоком однородных событий* будем понимать события, отличающиеся лишь моментами появления. Поток событий называется *регулярным*, если события следуют одно за другим через строго определенные промежутки времени.

Простейшим (пуассоновским) потоком называется поток событий, обладающий следующими свойствами:

1) *стационарностью*, т.е. вероятность попадания того или иного числа событий на участок времени длиной τ зависит только от длины участка и не зависит от того, где расположен этот участок;

2) *независимостью* (отсутствием последействия), т.е. число событий, попадающих на один из неперекрывающихся участков времени, не зависит от числа событий, попадающих на другие;

3) *ординарностью*, т.е. вероятность попадания на элементарный участок Δt двух или более событий пренебрежимо мала по сравнению с вероятностью попадания одного события.

Отметим, что при соблюдении условий 1–3 число событий, попадающих на интервал времени $]0;\ t[$, будет распределено по закону Пуассона

$$P_t(k) = (\lambda t)^k e^{-\lambda t} / k!,$$

где λ – *интенсивность потока*, равная среднему числу событий, появляющихся в единицу времени.

Пример 5.7. Среднее число автомобилей, прибывающих на станцию обслуживания для ремонта в течение 1 ч, равно 2. Найти вероятность того, что за 4 ч прибудет: 1) 3 автомобиля; 2) менее 3 автомобилей; 3) не менее 3 автомобилей.

Р е ш е н и е. Предположим, что процесс поступления автомобилей на станцию обслуживания в течение t ч является пуассоновским. Тогда возможными значениями переменной k будут 0, 1, 2. Вероятности этих значений определяем по приведенным ниже формулам.

$$1.\ P_4(k=3) = \frac{8^3 e^{-8}}{3!} = \frac{512 \cdot 0{,}000335}{6} = 0{,}028.$$

$$2.\ P_4(k<3) = P(k=0) + P(k=1) + P(k=2) = e^{-8} + \frac{8e^{-8}}{1} + \frac{8^2 e^{-8}}{2} \cong 0{,}01.$$

3. В связи с тем, что события "поступило менее 3 автомобилей" и "поступило не менее 3 автомобилей" вместе образуют достоверное событие, то $P_4(k<3) + P_4(k \geqslant 3) = 1$. Отсюда

$$P_4(k \geqslant 3) = 1 - P_4(k<3) = 1 - 0{,}01 = 0{,}99.$$

ЗАДАЧИ

5.18. Число электронов, вылетающих из нагретого катода электронной лампы в течение времени t, подчиняется закону Пуассона со средним числом выпускаемых в единицу времени электронов, равным λ. Найти вероятности того, что: 1) число электронов за время t_1 будет меньше m; 2) за время t_2 вылетит четное число электронов.

5.19. Среднее число вызовов, поступающих на АТС в минуту, равно 120. Найти вероятности того, что: 1) за 2 с на АТС не поступит ни одного вызова; 2) за 3 с на АТС поступит не менее 6 вызовов.

5.20. Среднее число автомобилей, прибывающих на автозаправочную станцию за 1 ч, равно 4. Найти вероятность того, что за 3 ч прибудет: 1) 6 автомобилей; 2) менее 6 автомобилей; 3) не менее 6 автомобилей.

5.21. Найти вероятность того, что за 2 ч на предприятие бытового обслуживания поступит 4 заявки, если число заявок в среднем за 1 ч равно 3.

5.6. МАРКОВСКИЕ ПРОЦЕССЫ С ДИСКРЕТНЫМИ СОСТОЯНИЯМИ

Марковским процессом, протекающим в физической системе, называется такой случайный процесс $X(t)$, при котором для каждого момента времени вероятность любого состояния системы в будущем зависит только от состояния системы в настоящий момент и не зависит от того, каким образом система пришла в это состояние.

Марковский процесс называется *процессом с дискретными состояниями,* если имеется лишь конечное или счетное число различных фазовых состояний системы.

Если переходы системы из состояния в состояние возможны только в строго определенные, заранее фиксированные моменты времени t_1, t_2, то марковский процесс $X(t)$ называется *процессом с дискретным временем,* если же переходы системы из состояния в состояние возможны в любые случайные моменты времени t, то марковский процесс называется *процессом с непрерывным временем.*

Пусть имеется физическая система A, которая может находиться в различных фазовых состояниях: A_1, A_2, ..., A_m, причем переходы системы из состояния в состояние возможны только в моменты времени t_1, t_2, ..., t_k, ... Будем называть эти моменты *шагами процесса* и рассматривать марковский случайный процесс $X(t)$, происходящий в системе A, как функцию целочисленного аргумента 1, 2, ..., k (номер шага). Тогда $X(k)$ обозначает состояние системы A через k шагов. Предположим, что цепочка последовательных переходов

$$X(0) \to X(1) \to X(2) \to \dots$$

зависит от вмешательства случая, причем соблюдается следующая закономерность: если на каком-либо шаге k система находилась в состоянии A_i^k, то, независимо от предшествующих обстоятельств, она на следующем шаге с вероятностью P_{ij} переходит в состояние $A_{ij}^{(k+1)}$:

$$P_{ij} = P(A_j^{(k+1)} | A_i^{(k)}), \quad i, j = 1, 2, \dots, m .$$

Такая случайная последовательность называется *марковской цепью,* а вероятности P_{ij} – *переходными вероятностями.*

Марковская цепь называется *однородной,* если переходные вероятности не зависят от номера шага; в противном случае она называется *неоднородной.*

Если для каждого состояния физической системы A известна вероятность перехода в любое другое состояние за один шаг, то переходные вероятности P_{ij} записывают в виде матрицы:

$$\mathrm{P} \; = \; \begin{pmatrix} P_{11} & P_{12} & \dots & P_{1m} \\ P_{21} & P_{22} & \dots & P_{2m} \\ \dots & \dots & \dots & \dots \\ P_{m1} & P_{m2} & \dots & P_{mm} \end{pmatrix},$$

которая называется *матрицей перехода.* Так как в каждой строчке матрицы записаны вероятности всех возможных переходов из выбранного состояния и эти переходы образуют полную систему событий, то

$$\sum_{j=1}^{m} P_{ij} = 1, \quad i = 1, 2, \dots, m.$$

Обозначим через $P_{ij}(n)$ вероятность того, что в результате n шагов (испытаний) система перейдет из состояния i в состояние j. Например, $P_{25}(10)$ – вероятность перехода за 10 шагов из второго состояния в пятое. Заметим, что при $n = 1$ получим переходные вероятности.

Поставим следующую задачу: зная переходные вероятности P_{ij}, найти вероятности $P_{ij}(n)$ перехода системы из состояния i в состояние j за n шагов. Для этого введем промежуточное (между i и j) состояние r. Другими словами, будем считать, что из первоначального состояния i за m шагов система перейдет в промежуточное состояние r с вероятностью $P_{ir}(m)$, после чего за оставшиеся $n-m$ шагов из промежуточного состояния r она перейдет в конечное состояние j с вероятностью $P_{rj}(n-m)$. По формуле полной вероятности

$$P_{ij}(n) \; = \; \sum_{r=1}^{k} P_{ir}(m) P_{rj}(n-m).$$

Покажем, что, зная все переходные вероятности $P_{ij} = P_{ij}(1)$, т.е. матрицу P_1 перехода из состояния в состояние за один шаг, можно найти вероятности $P_{ij}(2)$ перехода из состояния в состояние за два шага, а следовательно, и саму матрицу перехода P_2, по известной матрице P_2 можно найти матрицу P_3 перехода из состояния в состояние за три шага и т.д. Действительно, положив $n = 2, m = 1$ в равенстве Маркова, получим

$$P_{ij}(2) = \sum_{r=1}^{k} P_{ir}(1) P_{rj}(2-1)$$

или

$$P_{ij}(2) = \sum_{r=1}^{k} P_{ir} P_{rj}.$$

По этой формуле можно найти все вероятности $P_{ij}(2)$, а следовательно, и саму матрицу P_2. Запишем вытекающее из полученной формулы соотношение в матричной форме:

$$\mathrm{P}_2 = \mathrm{P}_1 \mathrm{P}_1 = \mathrm{P}_1^2.$$

Положив $n = 3$, $m = 2$, аналогично получим

$$P_3 = P_1 P_2 = P_1 P_1^2 = P_1^3 .$$

В общем случае

$$P_n = P_1^n .$$

Для описания случайного процесса, протекающего в физической системе A с дискретными состояниями $A_1, A_2, \ldots, A_m$, часто пользуются вероятностями состояний $P_1(k), P_2(k), \ldots, P_m(k)$, где $P_l(k)$ $(k = 1, 2, \ldots, m)$ – вероятность того, что через k шагов система A будет находиться в состоянии A_l. Вероятности $P_l(k)$ удовлетворяют условию $\sum\limits_{l=1}^{m} P_l(k) = 1$.

Если система A в начальный момент (перед первым шагом) находится в каком-то определенном состоянии, например A_r, то для начального момента времени ($t = 0$) будем иметь $P_1(0) = 0$, $P_2(0) = 0, \ldots, P_{r-1}(0) = 0, P_r(0) = 1$, $P_{r+1}(0) = 0, \ldots, P_m(0) = 0$.

Вероятности состояний системы A через k шагов определяются рекуррентной формулой

$$P_l(k) = \sum_{i=1}^{m} P_i(k-1) P_{ij} .$$

Если переход системы A из одного состояния в другое возможен в любой момент времени t, а вероятности $P_{ij}(t, \tau)$ перехода системы из состояния A_i в момент времени t в состояние A_k в момент времени τ не зависят от поведения системы до момента времени t, то $X(t)$ является марковским случайным процессом с дискретным числом состояний и для вероятностей перехода справедливо соотношение

$$P_{ik}(t, \tau) = \sum_{j=0}^{m} P_{ij}(t, s) P_{jk}(s, \tau) .$$

Процесс называется *однородным*, если $P_{ik}(t, \tau) = P_{ik}(\tau - t)$. В этом случае для марковского процесса

$$P_{ik}(\tau - t) = \sum_{j=0}^{m} P_{ij}(s-t) P_{jk}(\tau - s), \quad t \leqslant s \leqslant \tau .$$

Марковский процесс называется *регулярным*, если выполняются следующие условия:

1) для любого состояния системы A_k существует предел

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \left(1 - P_{kk}(t, t + \Delta t)\right) = c_k(t); \tag{5.9}$$

2) для каждой пары состояний системы A_i, A_k существует непрерывная по t плотность распределения вероятностей перехода $P_{ik}(t)$, определяемая равенством

$$P_{ik}(t) = \frac{1}{c_i(t)} \lim_{\Delta t \to 0} \frac{P_{ik}(t, t+\Delta t)}{\Delta t} . \qquad (5.10)$$

Для регулярных марковских процессов вероятности P_{ik} определяются системами дифференциальных уравнений:

$$\frac{\partial P_{ik}(t, \tau)}{\partial t} = -c_k(\tau) P_{ik}(t, \tau) + \sum_{j \neq k} P_{ij}(t, \tau) c_j(\tau) P_{ik}(\tau) ,$$

$$\frac{\partial P_{ik}(t, \tau)}{\partial t} = c_i(t) P_{ik}(t, \tau) - c_i \sum_{j \neq 1} P_{jk}(t, \tau) P_{ij}(t) ,$$

$$i, j, k = 0, 1, 2, \ldots, m ,$$

с начальными условиями $P_{ik}(t, t) = \delta_{ik}$:

$$\delta_{ik} = \begin{cases} 1 \text{ при } i = k , \\ 0 \text{ при } i \neq k . \end{cases}$$

Для однородных марковских процессов вследствие того, что $c_i(t)$ и $P_{ij}(t)$ не зависят от t , а $P_{ik}(t, \tau) = P_{ik}(\tau - t)$, системы дифференциальных уравнений имеют вид:

$$\frac{dP_{ik}(t)}{dt} = -c_k P_{ik}(t) + \sum_{j \neq k} c_j P_{jk} P_{ij}(t) ,$$

$$\frac{dP_{ik}(t)}{dt} = -c_i P_{ik}(t) + c_i \sum_{j=k} P_{ij} P_{jk}(t), \quad i, j, k = 0, 1, 2, \ldots, m$$

при начальных условиях $P_{ik}(0) = \delta_{ik}$.

Вероятности P_k нахождения системы в состоянии A_k в момент времени t определяются системой

$$\frac{dP_k}{dt} = -c_k P_k(t) + \sum_{j \neq k} c_j(t) P_{jk}(t), \; j, k = 0, 1, 2, \ldots, m ,$$

при соответствующих начальных условиях для $P_j(t)$. Если начальное состояние A_i задано, то начальными условиями будут $P_k(t) = \delta_{ik}$ при $t = 0$.

Для однородных процессов последняя система имеет вид

$$\frac{dP_k(t)}{dt} = -c_k P_k(t) + \sum_{j \neq k} c_j P_{jk} P_j(t)$$

при начальных условиях $p_k(t) = \delta_{ik}$, $t = 0$.

Если для однородного процесса $P_{ik}(t) > 0$ при $t > 0$ и любых i, k, то процесс называется *транзитивным* и для него существует не зависящий от номера исходного состояния предел

$$\lim_{t \to \infty} P_{ik}(t) = \lim_{t \to \infty} P_k(t) = P_k , \tag{5.11}$$

причем предельные вероятности P_k в этом случае определяются из системы

$$c_k P_k = \sum_{j \neq k} c_j P_{jk} P_j , \quad j, k = 0, 1, 2, ..., m . \tag{5.12}$$

Процессом Маркова является простейший поток, обладающий следующими свойствами:

1) *стационарностью,* т.е. при любом $\Delta t > 0$ и целом $k \geqslant 0$ вероятность того, что за промежуток времени $]t;\ t + \Delta t[$ произойдет k событий, одна и та же для всех $t \geqslant 0$;

2) *отсутствием последействия,* т.е. вероятность наступления событий за промежуток времени $]t;\ t + \Delta t[$ не зависит от числа наступления событий до момента времени t;

3) *ординарностью*, т.е.

$$\lim_{\Delta t \to 0} \frac{R_2(\Delta t)}{\Delta t} = 0 ,$$

где $R_2(\Delta t)$ – вероятность наступления не менее двух событий за промежуток времени Δt .

Пример 5.8. Система обслуживания состоит из m приборов, каждый из которых может выполнить одновременно только одно требование, затрачивая на обслуживание случайное время, распределенное по показательному закону с параметром μ . В систему поступает простейший поток требований с параметром λ . Обслуживание требования начинается сразу после его поступления, если в этот момент имеется хотя бы один свободный прибор, в противном случае требование "получает отказ" и не возвращается в систему. Определить предельную вероятность отказа.

Р е ш е н и е. Пусть A_i – состояние, при котором i приборов заняты обслуживанием, тогда $P_{ik}(t) > 0$ для конечного t и, следовательно, согласно формуле (5.11),

$$P_n = \lim_{t \to \infty} P_n(t) .$$

Вероятность P_n определяем из системы (5.12). Имеем

$$c_n P_n = c_{n-1} P_{n-1,n} P_{n-1} + c_{n+1} P_{n+1,n} P_{n+1} . \tag{5.13}$$

Так как поток требований – простейший, а время обслуживания подчиняется показательному закону, то для промежутка времени $]t;\ t + \Delta t[$ можем записать:

$$P_{n,n+1}(t, t + \Delta t) = \lambda \Delta t\,(1 - n\mu \Delta t) + o(\Delta t) = \lambda \Delta t + o(\Delta t),$$

$$P_{n,n-1}(t, t + \Delta t) = (1 - \lambda \Delta t);\ \mu \Delta t + o(\Delta t) = n\mu \Delta t + o(\Delta t),$$

$$P_{n,n}(t, t+\Delta t) = (1-\mu\Delta t)(1-n\mu\Delta t) + o(\Delta t) = 1 - (\lambda + n\mu) + o(\Delta t).$$

Отсюда, согласно формулам (5.11) и (5.12), имеем:

$$c_n = \lambda + n\mu, \; P_{n,n+1} = \frac{\lambda}{\lambda + n\mu} \; P_{n,n-1} = \frac{n\mu}{\lambda + n\mu},$$

$$c_m = m\mu, \; P_{m,m+1} = 0,$$

остальные $P_{jk} = 0$. Подставляя полученные значения в равенство (5.13), находим:

$$(\lambda + n\mu) P_n \lambda P_{n-1} + (n+1)\mu P_{n+1} \quad (0 \leqslant n \leqslant m-1),$$

$$m\mu P_m = \lambda P_{m-1}.$$

Отсюда

$$P_k = \frac{\lambda}{n\mu} P_{n-1} \quad \text{или} \quad P_n = \frac{1}{n!} \left(\frac{\lambda}{\mu}\right)^m P_0 .$$

Так как $\sum\limits_{n=0}^{m} P_k = 1$, то

$$P_0 = 1 / \sum_{n=0}^{m} \frac{1}{n!} \left(\frac{\lambda}{\mu}\right)^n .$$

Вероятность отказа в обслуживании выражается *формулой Эрланга* :

$$P_m = \frac{1}{m!} \left(\frac{\lambda}{\mu}\right)^n / \sum_{n=0}^{m} \frac{1}{n!} \left(\frac{\lambda}{\mu}\right)^n .$$

Пример 5.9. Дана матрица перехода $P_1 = \begin{pmatrix} 0{,}4 & 0{,}6 \\ 0{,}3 & 0{,}7 \end{pmatrix}$. Найти матрицу перехода

$$P_2 = \begin{pmatrix} p_{11}^{(2)} & p_{12}^{(2)} \\ p_{21}^{(2)} & p_{22}^{(2)} \end{pmatrix}.$$

Р е ш е н и е. Воспользуемся формулой $P_2 = P_1^2$:

$$P_2 = \begin{pmatrix} 0{,}4 & 0{,}6 \\ 0{,}3 & 0{,}7 \end{pmatrix} \begin{pmatrix} 0{,}4 & 0{,}6 \\ 0{,}3 & 0{,}7 \end{pmatrix}.$$

Перемножив матрицы, окончательно получим

$$P_2 = \begin{pmatrix} 0{,}34 & 0{,}66 \\ 0{,}33 & 0{,}67 \end{pmatrix}.$$

ЗАДАЧИ

5.22. Дана матрица перехода P_1. Найти матрицу перехода P_2, если:

$$1)\ P_1 = \begin{pmatrix} 0{,}3 & 0{,}7 \\ 0{,}4 & 0{,}6 \end{pmatrix}; \qquad 2)\ P_1 = \begin{pmatrix} 0{,}8 & 0{,}2 \\ 0{,}2 & 0{,}8 \end{pmatrix}.$$

5.23. Найти вероятности перехода из состояния в состояние за два шага, если вероятности перехода за один шаг заданы матрицей:

$$1)\ \begin{pmatrix} 1/2 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \end{pmatrix}; \qquad 2)\ \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}.$$

5.24. Один рабочий обслуживает m автоматических станков, которые при нормальной работе не требуют его вмешательства. Остановки каждого станка вследствие неполадок образуют независимый простейший поток с параметром λ. Для устранения неполадки рабочий тратит случайное время, распределенное по показательному закону с параметром μ. Найти предельные вероятности того, что k станков не работают.

5.25. Сколько мест должно быть на станции обслуживания автомобилей, чтобы с вероятностью, не меньшей 0,9, автомашина, нуждающаяся в ремонте, обеспечивалась местом для ремонта, если заявки образуют простейший поток, а время обслуживания подчиняется показательному закону? Среднее время ремонта – сутки. В течение суток на станцию поступает в среднем пять автомашин.

5.26. В ремонтную мастерскую, в которой трое рабочих, поступает в среднем 4 заказа в час. Среднее время выполнения заказа 0,4 ч. Определить среднее число заказов, ожидающих начала исполнения.

6. ОСНОВНЫЕ ПОНЯТИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

6.1. ГЕНЕРАЛЬНАЯ И ВЫБОРОЧНАЯ СОВОКУПНОСТИ

Группа объектов, объединенных по некоторому качественному или количественному признаку, называется *статистической совокупностью.* Различают генеральную и выборочную совокупности.

Выборочной совокупностью или *выборкой* называется совокупность случайно отобранных объектов.

Генеральной совокупностью называется совокупность всевозможных объектов, из которых производится выборка.

Объемом совокупности называется число объектов, входящих в эту совокупность. Например, если из 1000 приборов выбрано для обследования 100, то объем генеральной совокупности равен 1000, а выборочной – 100.

На практике используются различные способы получения выборки или отбора.

1. Отбор, не требующий разбиения генеральной совокупности на части, например простой случайный бесповторный отбор и простой случайный повторный отбор.

2. Отбор, при котором генеральная совокупность разбивается на части, например типический, механический и серийный отборы.

Простым случайным называется отбор, при котором объекты извлекают по одному из всей генеральной совокупности. Если при этом выбранный объект перед отбором следующего объекта возвращается в генеральную совокупность, то отбор называется *повторным.* Если же выбранный объект в генеральную совокупность не возвращается, то отбор называется *бесповторным.*

Типическим называется отбор, при котором объекты извлекают не из всей генеральной совокупности, а только из некоторой ее части. Типические отборы целесообразны в случаях, когда значения изучаемого признака заметно колеблются в различных частях выборки. (Например, продукция изготавливается на нескольких станках, среди которых есть более или менее изношенные.)

Механическим называется отбор, при котором генеральная совокупность произвольным образом делится на столько групп, сколько объектов должно войти в выборку, и из каждой группы извлекается один объект.

Серийным называется отбор, при котором из генеральной совокупности объекты выбирают не по одному, а группами (сериями). Серийный отбор применяется тогда, когда значение исследуемого признака незначительно колеблется в различных сериях.

Возможен и *комбинированный отбор,* при котором используются названные выше способы. Например, генеральную совокупность разбивают на серии одинакового объема, затем простым случайным отбором находят несколько серий и отдельные объекты в каждой серии.

Статистическая совокупность, расположенная в порядке возрастания или убывания значения признака, называется *вариационным рядом,* а ее объекты – *вариантами.*

Вариационный ряд называется *дискретным,* если его члены принимают конкретные изолированные значения. Если члены вариационного ряда заполняют некоторый интервал, то такой ряд называют *непрерывным.*

Статистическим распределением выборки называется соотношение между значениями вариант x_i и соответствующими им частотами m_i или частостями ω_i. Статистическое распределение можно представить в виде таблицы (табл. 6.1), в первую строку которой записывают варианты (члены вариационного ряда), а во вторую – соответствующие им частоты или частости.

Таблица 6.1

x_i	x_1	x_2	x_3	x_4	x_5	...	...	x_n
m_i	m_1	m_2	m_3	m_4	m_5	...	...	m_k
ω_i	ω_1	ω_2	ω_3	ω_4	ω_5	...	...	ω_k

В табл. 6.1 $n = \sum_{i=1}^{k} m_i$, $\sum_{i=1}^{k} \omega_i = 1$, $\omega_i = m_i/n$, $i = 1, 2, ..., k$.

Для непрерывного вариационного ряда составляется таблица (табл. 6.2), в первой строке которой помещены интервалы изменения вариант, а во второй – соответствующие им частоты.

Таблица 6.2

x_i	$x_1 - x_2$	$x_2 - x_3$	$x_3 - x_4$	...	$x_l - x_{l+1}$
m_i	m_1	m_2	m_3	...	m_l

Геометрически дискретное распределение можно интерпретировать следующим образом. На оси абсцисс откладывают варианты x_i, на оси ординат – соответствующие им частоты m_i или частости ω_i. Ломаную линию, соединяющую полученные точки, называют *полигоном.*

Для непрерывного вариационного ряда интервал, в котором заключены все значения ряда, разбивают на несколько частичных интервалов с шагом $x_i - x_{i+1} = h$ и находят для каждого частичного интервала сумму частот m_i или частостей ω_i вариант, попавших в i-й интервал. На оси абсцисс откладывают частичные интервалы, а над ними на расстоянии m_i/h или ω_i/h проводят отрезки, параллельные оси абсцисс. Соединив концы отрезков и интервалов линиями, параллельными оси Oy, получают ступенчатую фигуру, состоящую из прямоугольников, которая называется *гистограммой.* Площадь i-го частичного прямоугольника равна $h \dfrac{m_i}{h} = m_i$, т.е. сумме частот вариант i-го интервала. Значит, площадь гистограммы частот равна сумме всех частот, т.е. объему выборки. Это справедливо и для относительных частот.

Задачами выборочного метода являются: 1) установление закона распределения СВ и его параметров по данным выборки; 2) статистическая проверка гипотез.

Пример 6.1. Составить вариационный ряд для следующих значений длины случайным образом отобранных заготовок: 39, 41, 40, 43, 41, 44, 42, 41, 41, 43, 42, 39, 40, 42, 43, 42, 41, 39, 42, 42, 41, 42, 40, 41, 43, 41, 39, 40, 41, 40 и построить полигон.

Р е ш е н и е. Так как длина имеет конкретное значение, то вариационный ряд будет дискретным. В первой строке таблицы расположим значения признака x (длина заготовки) в порядке возрастания, а во второй – количество заготовок данной длины (частоты):

x_i	39	40	41	42	43	44	
m_i	4	5	9	7	4	1	$\sum\limits_{i=1}^{6} m_i = 30$

Построим полигон. Для этого по оси Ox отложим значения признака x, а по оси Oy – частоты m_i (рис. 6.1), полученные точки соединим отрезками прямых.

Пример 6.2. Из продукции автомата, обрабатывающего ролики диаметром $D =$ $= 20$ мм, взята выборка объемом 100 штук. Диаметр ролика измерен микрометром с ценой деления 0,01 мкм. Составить непрерывный вариационный ряд и построить гистограмму по следующим данным отклонений диаметра от номинального:

−0,07	−0,03	−0,04	−0,08	−0,03	−0,08	−0,09	−0,10	−0,10	−0,10
−0,13	−0,08	−0,06	−0,04	−0,04	−0,03	−0,04	−0,07	−0,11	−0,12
−0,03	−0,07	−0,08	−0,11	−0,05	−0,05	−0,07	−0,03	−0,09	−0,10
−0,11	−0,14	−0,13	0,08	−0,12	−0,07	−0,09	−0,10	−0,11	−0,08
−0,05	−0,12	−0,07	−0,06	−0,08	−0,11	−0,10	−0,12	−0,11	−0,10
−0,08	−0,05	−0,11	−0,07	−0,05	−0,08	−0,09	−0,09	−0,09	−0,02
−0,06	−0,12	−0,05	−0,07	−0,11	−0,05	−0,08	−0,03	−0,09	−0,09
−0,11	−0,06	−0,07	−0,06	−0,06	−0,12	−0,10	−0,08	−0,11	−0,01
−0,05	−0,07	−0,06	−0,05	−0,08	−0,09	−0,04	−0,09	−0,08	−0,09
−0,07	−0,06	−0,06	−0,12	−0,05	−0,03	−0,10	−0,09	−0,09	−0,08

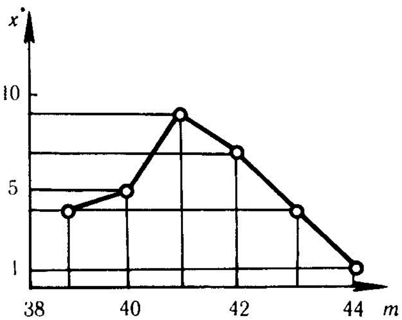
Р е ш е н и е. Так как наименьшее отклонение равно 0,01, а наибольшее – 0,14, то разобьем весь объем на 7 интервалов длиной 0,02. Значения признака и частоты расположим таким же образом, как в примере 6.1 (табл. 6.3).

Таблица 6.3

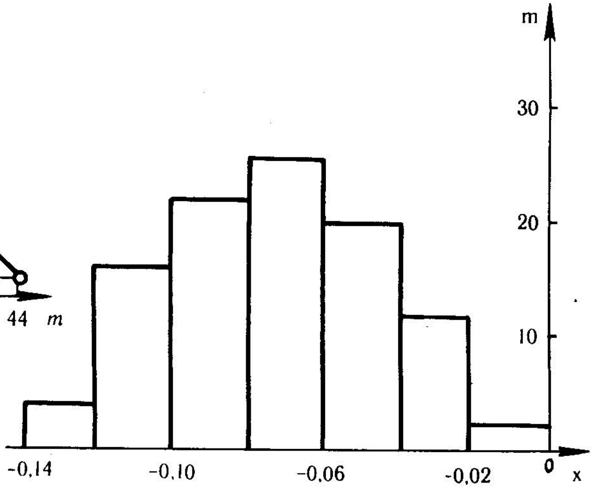
x_i	−0,14 −0,12	−0,12 −0,10	−0,10 −0,08	−0,08 −0,06	−0,06 −0,04	−0,04 −0,02	−0,02 −0,00	
m_i	3	16	22	25	19	13	2	$\sum\limits_{i=1}^{7} m_i = 100$

Значение варианты, принадлежащее границе, отнесем к тому интервалу, у которого граница является левой.

Построим гистограмму: для этого по оси абсцисс отложим интервалы (значения признака), а по оси ординат – частоты (рис. 6.2). Затем через каждую точку проведем прямые, параллельные осям координат. Полученная ступенчатая фигура является гистограммой.



Р и с. 6.1



Р и с. 6.2

ЗАДАЧИ

6.1. Отклонения от номинального размера, полученные при измерении деталей, имеют следующие значения:

0,02	0,07	0,13	0,05	0,11	0,17	0,17	0,05	0,03	0,11
0,04	0,14	0,10	0,11	0,13	0,14	0,16	0,04	0,13	0,04
0,03	0,15	0,11	0,06	0,10	0,15	0,16	0,06	0,14	0,06
0,04	0,08	0,14	0,08	0,08	0,14	0,13	0,07	0,16	0,08
0,02	0,10	0,16	0,04	0,09	0,15	0,12	0,17	0,17	0,16
0,04	0,06	0,17	0,03	0,04	0,04	0,11	0,15	0,15	0,17
0,05	0,07	0,04	0,11	0,05	0,08	0,10	0,14	0,6	0,9
0,03	0,06	0,02	0,17	0,07	0,09	0,09	0,12	0,7	0,10
0,06	0,11	0,09	0,14	0,06	0,10	0,08	0,10	0,11	0,09
0,08	0,12	0,10	0,10	0,10	0,02	0,07	0,06	0,12	0,10

Составить вариационный ряд и построить гистограмму частот.

6.2. Составить вариационный ряд овальности валиков (в мкм) по следующим данным:

1	2	3	4	5	6	7	8	9	10
20	12	14	13	12	13	14	15	16	14
15	14	12	15	14	13	15	12	13	14
15	13	10	10	11	9	10	8	9	11
11	10	9	9	11	8	8	9	10	10
10	9	10	8	7	6	5	7	6	5
4	5	7	7	3	16	17	18	19	16
19	18	17	16	17	18	19	18	17	19
18	20	18	21	22	23	21	10	21	24
25	17	19	28	29	1	2	3	2	17

и построить полигон частот.

6.3. Возраст студентов одного потока представляется следующими данными: 17, 20, 18, 19, 18, 17, 20, 21, 24, 22, 20, 21, 20, 19, 18, 20, 21, 22, 25, 20. Составить вариационный ряд, построить полигон частот.

6.4. В результате проверки партии деталей получены следующие результаты по сортам: 1, 2, 1, 2, 1, 1, 1, 3, 4, 1, 1, 2, 2, 3, 1, 1, 1, 2, 1, 1, 4, 2, 2, 1, 1. Составить вариационный ряд, построить полигон частот.

6.5. Из выпускаемого заводом литья произведена случайная выборка 40 штук литых деталей, взвешивание которых дало следующие результаты (в кг):

99,2	101,5	99,5	103,2	99,7	110,1	100,2	99,2
99,3	100,4	100,3	102,4	98,8	98,8	100,4	99,7
97,6	101,2	99,4	98,2	100,1	98,3	100,7	101,2
97,2	99,7	101,3	100,6	100,7	101,6	102,7	98,6
98,7	99,9	98,2	100,7	101,2	99,6	100,3	99,8

Найти эмпирическое распределение веса литых деталей в данной выборке (вариационный ряд).

6.2. ЭМПИРИЧЕСКАЯ ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ

Эмпирической (статистической) функцией распределения называется функция $F^*(x)$, значения которой равны относительной частоте ω события $X < x$:

$$F^*(x) = \omega(X < x),$$

где $\omega(X < x) = m_x/n$; m_x – число вариант, меньших x; m – объем выборки.

Приведем *свойства эмпирической функции распределения.*

1. $0 \leqslant F^*(x) \leqslant 1$.

2. $F^*(x)$ – *неубывающая функция.*

3. *Если значения вариант расположены в промежутке* $]x_1; x_k[$, *то*

$$F^*(x) = \begin{cases} 0 & \text{при } X < x_1, \\ 1 & \text{при } X > x_k. \end{cases}$$

Функция распределения генеральной совокупности $F(x)$ называется *теоретической функцией распределения.* Различие между эмпирической и теоретической функциями распределения состоит в том, что $F(x)$ определяет вероятность события $X < x$, а $F^*(x)$ – относительную частоту этого события. Из закона больших чисел (теоремы Бернулли) следует, что $F^*(x)$ стремится по вероятности к $F(x)$, т.е. при больших n значения $F^*(x)$ и $F(x)$ мало отличаются друг от друга в том смысле, что

$$\lim_{n \to \infty} P(|F(x) - F^*(x)| < \epsilon) = 1, \quad \epsilon > 0.$$

Отсюда следует, что эмпирическую функцию распределения можно использовать для приближенного представления теоретической функции.

Пример 6.3. Найти эмпирическую функцию распределения СВ X, заданной следующей выборкой:

x_i	39	40	41	42	43	44
m_i	4	5	9	7	4	1

Р е ш е н и е. Находим объем выборки: $m = 4 + 5 + 9 + 7 + 4 + 1 = 30$. Наименьшая варианта равна 39, поэтому $F^*(x) = 0$ при $x < 39$.

Значения $X < 40$, т.е. $x_1 = 39$, наблюдались 4 раза, следовательно, $F^*(x) = 4/30 = 0{,}13$ при $39 \leqslant x < 40$.

Значения $X < 41$, а именно: $x_1 = 39$, $x_2 = 40$, наблюдались $4 + 5 = 9$ раз, поэтому $F^*(x) = 9/30 = 0{,}3$ при $40 \leqslant x < 41$.

Значения $X < 42$, т.е. $x_1 = 39$, $x_2 = 40$, $x_3 = 41$, наблюдались $4 + 5 + 9 = 18$ раз, следовательно, $F^*(x) = 18/30 = 0{,}6$ при $41 \leqslant x < 42$.

Значения $X < 43$, а именно: $x_1 = 39$, $x_2 = 40$, $x_3 = 41$, $x_4 = 42$, наблюдались $4 + 5 + 9 + 7 = 25$ раз, поэтому $F^*(x) = 25/30 = 0{,}71$ при $42 \leqslant x < 43$.

Аналогично находим $F^*(x) = 29/30 = 0{,}97$ при $43 \leqslant x < 44$. Так как $x = 44$ – наибольшая варианта, то $F^*(x) = 1$ при $x > 44$. Запишем эмпирическую функцию:

$$F^*(x) = \begin{cases} 0 & \text{при} \quad x < 39, \\ 0{,}13 & \text{при} \quad 39 \leqslant x < 40, \\ 0{,}3 & \text{при} \quad 40 \leqslant x < 41, \\ 0{,}6 & \text{при} \quad 41 \leqslant x < 42, \\ 0{,}71 & \text{при} \quad 42 \leqslant x < 43, \\ 0{,}97 & \text{при} \quad 43 \leqslant x < 44, \\ 1 & \text{при} \quad 44 \leqslant x. \end{cases}$$

Пример 6.4. Найти эмпирическую функцию по следующему распределению:

x_i	–0,14 –0,12	–0,12 –0,10	–0,10 –0,08	–0,08 –0,06	–0,06 –0,04	–0,04 –0,02	–0,02 –0,00	
m_i	3	16	22	25	19	13	2	$\sum_{i=1}^{7} m_i = 100$

Решение. В данном случае имеем непрерывный вариационный ряд. Эмпирическую функцию найдем таким же способом, как и в примере 6.3, значения X возьмем на концах интервала:

$$F^*(-0{,}14) = 0,\ F^*(-0{,}12) = F^*(-0{,}14) + 3/100 = 0{,}03,$$

$$F^*(-0{,}10) = F^*(-0{,}14) + F^*(-0{,}12) + 16/100 = 0 + 0{,}03 + 0{,}16 = 0{,}19,$$

$$F^*(-0{,}08) = 0 + 0{,}03 + 0{,}16 + 0{,}22 = 0{,}41,$$

$$F^*(-0{,}06) = 0{,}41 + 0{,}25 = 0{,}66,\ F^*(-0{,}04) = 0{,}66 + 0{,}19 = 0{,}85,$$

$$F^*(-0{,}02) = 0{,}85 + 0{,}13 = 0{,}98,\ F^*(0{,}00) = 0{,}98 + 0{,}02 = 1.$$

ЗАДАЧИ

6.6. Найти эмпирическую функцию распределения СВ X и построить ее график для распределения рабочих механического цеха по тарифным разрядам по следующим данным (x_i – тарифный разряд, m_i – количество рабочих):

x_i	1	2	3	4	5	
m_i	4	6	16	26	48	$\sum_{i=1}^{5} m_i = 100$

6.7. Найти эмпирическую функцию и построить ее график по следующим распределениям:

1)

x_i	2	5	7	8
m_i	1	3	2	4

2)

x_i	4	7	8
m_i	5	2	3

3)

x_i	1	4	5	7
m_i	20	10	14	6

6.3. ОСНОВНЫЕ ХАРАКТЕРИСТИКИ ВЫБОРКИ

Пусть случайный эксперимент описывается СВ X. Повторяя его, получаем последовательность наблюденных значений x_1, x_2, ..., x_n СВ X, составляющих выборку.

Определим основные характеристики эмпирического распределения СВ X.

Средним арифметическим наблюденных значений выборки называется величина, определяемая по формуле

$$\overline{x} = x_b = \sum_{i=1}^{k} x_i \frac{m_i}{n},$$

где x_i – наблюденное значение с частотой m_i; n – число наблюдений.

Статистической дисперсией $D(\overline{X}) = S_x^2$ выборочного распределения называется среднее арифметическое квадратов отклонений значений наблюдений от среднего арифметического x:

$$D(\overline{X}) = S_x^2 = \sum_{i=1}^{k} (x_i - x)^2 \frac{m_i}{n},$$

где x_i – наблюденное значение с частотой m_i; $\sum\limits_{i=1}^{k} m_i = n$ – число наблюдений.

Статистическая дисперсия вычисляется по следующей формуле:

$$S_x^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2,$$

которая после преобразований принимает вид

$$S_x^2 = x^2 - \overline{x}^2.$$

Средним квадратичным отклонением или *выборочным средним* называется квадратный корень из статистической дисперсии:

$$S_x = \sqrt{D(\overline{X})} = \sqrt{S_x^2}.$$

Средние величины определяются числовыми характеристиками выборочного распределения.

ЗАДАЧИ

6.8. По данным выборки

x_i	0,1	0,5	0,6	0,8
m_i	5	15	20	10

найти: а) выборочное среднее $\overline{x}$; б) выборочную дисперсию S^2.

6.9. По данным выборки

x_i	8,8	9,3	9,8	10,3	10,8	11,3	11,8
m_i	4	8	11	7	5	3	2

найти выборочную дисперсию.

7. СТАТИСТИЧЕСКОЕ ОЦЕНИВАНИЕ НЕИЗВЕСТНЫХ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ

7.1. ПОСТАНОВКА ЗАДАЧИ

На практике во многих случаях функция распределения рассматриваемой СВ X неизвестна; ее определяют по результатам наблюдений, т.е. по выборке (по виду гистограммы или полигона частостей). Наиболее часто при исследовании непрерывных СВ применяется нормальное распределение, так как оно является предельным для многих распределений.

После того как выбрано распределение, производится оценка параметров. Например, если для описания СВ X выбрано нормальное распределение

$$F(x, a, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-(t-a)^2/(2\sigma^2)}\, dt,$$

то по выборке $x_1, x_2, \dots, x_n$ требуется оценить два параметра – математическое ожидание a и среднее квадратичное отклонение σ. Если же выбрано распределение Пуассона, то требуется оценить только один параметр λ, которым определяется это распределение:

$$P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}.$$

Пусть из генеральной совокупности с функцией распределения $F(x, \theta)$, где θ – неизвестный параметр, извлечена выборка. Задача оценивания неизвестного параметра θ состоит в построении приближенных формул

$$\theta \cong u(x_1, x_2, \dots, x_n), \tag{7.1}$$

где $x_1, x_2, \dots, x_n$ – варианты рассматриваемой выборки. Функцию (7.1) называют *выборочной функцией* или *статистикой*, а ее значение – *оценкой* и обозначают $\hat{\theta} = u(x_1, x_2, \dots, x_n)$.

Так как любая выборка является конечной и случайной, то выборочные функции $\hat{\theta} = u(x_1, x_2, \dots, x_n)$ также являются случайными функциями.

Оценки параметров подразделяются на точечные и интервальные. *Точечная оценка* параметра $\hat{\theta}$ определяется одним числом $\theta^* = u(x_1, x_2, \dots, x_n)$. *Интервальной оценкой* называют оценку, которая определяется двумя числами $\hat{\theta}_1$ и $\hat{\theta}_2$ – концами интервала, накрывающего оцениваемый параметр θ.

7.2. КЛАССИФИКАЦИЯ ТОЧЕЧНЫХ ОЦЕНОК

Для того чтобы точечная оценка неизвестного параметра $\hat{\theta} = u(x_1, x_2, \dots, x_n)$ была наилучшей с точки зрения точности, необходимо, чтобы она была состоятельной, несмещенной, эффективной.

Состоятельной называется оценка $\hat{\theta} = u(x_1, x_2, \dots, x_n)$, сходящаяся по вероятности к оцениваемому параметру, т.е. такая, что для любого $\epsilon > 0$ $\lim P(|\theta - \hat{\theta}| < \epsilon) = 1$ при увеличении числа испытаний.

Несмещенной называется оценка $\hat{\theta} = u(x_1, x_2, ..., x_n)$ (оценка без систематической ошибки), математическое ожидание которой равно оцениваемому параметру, т.е. такая, что $M(\hat{\theta}) = \theta$.

Наряду с несмещенными оценками $\hat{\theta} = u(x_1, x_2, ..., x_n)$ применяются *асимптотически несмещенные оценки*, для которых $M(\hat{\theta}) \to \theta$ при увеличении объема выборки.

Состоятельные, несмещенные или асимптотически несмещенные оценки могут быть получены различными методами. Например, оценка математического ожидания – среднее арифметическое $\overline{x}$ – является несмещенной и состоятельной оценкой математического ожидания.

Известно, что дисперсия любой несмещенной оценки одного параметра θ удовлетворяет *неравенству Рао-Крамера*

$$D(\hat{\theta}) \geqslant -\frac{1}{NM(\partial^2 \ln f(x, \theta)/\partial \theta^2)}, \tag{7.2}$$

где N – число произведенных испытаний; $f(x, \theta)$ – плотность распределения вероятностей СВ. В случае дискретной СВ X плотность $f(x, \theta)$ заменяется функцией распределения вероятностей $P(X = x;\ \theta)$.

Эффективной называется оценка $\hat{\theta} = u(x_1, x_2, ..., x_n)$, при которой неравенство (7.2) обращается в равенство.

Пример 7.1. Определить оценку относительной частоты m/n в испытаниях Бернулли.

Р е ш е н и е. Из теоремы Бернулли следует, что

$$\lim_{n \to \infty} P(|m/n - P| < \epsilon) = 1, \quad \epsilon > 0.$$

Таким образом, m/n является состоятельной оценкой.

Далее определим несмещенность, для чего вычислим математическое ожидание относительной частоты, полагая, что $m = x$:

$$M(\frac{X}{n}) = \frac{1}{n} M(X) = \frac{1}{n} np = p.$$

Так как $M(\frac{X}{n}) = p$, то относительная частота является несмещенной оценкой.

Для того чтобы определить эффективность, воспользуемся неравенством Рао–Крамера (7.2).

Определим дисперсию относительной частоты:

$$D(\frac{m}{n}) = D(\frac{X}{n}) = \frac{1}{n^2} D(X) = \frac{1}{n^2} npq = \frac{pq}{n}.$$

Учитывая, что

$$f(x, \theta) = f(X = x; p) = C_n^x p^x (1 - p)^{n-x},$$

имеем

$$M(\frac{\partial^2 \ln f(x, p)}{\partial p^2}) = M(\frac{\partial^2}{\partial p^2}(\ln C_n^x + x \ln p + (n - x)\ln(1 - p))) =$$

$$= M\left(-\frac{x}{p^2} - \frac{n-x}{(1-p)^2}\right) = -\frac{M(x)}{p^2} - \frac{n-M(x)}{(1-p)^2} = -\frac{np}{p^2} - \frac{n-np}{(1-p)^2} =$$

$$= -\frac{n}{pq}.$$

Так как оценка параметра может быть получена по каждому из N испытаний, то для $N = 1$

$$-\frac{1}{2\,M\left(\dfrac{\partial^2 \ln(x,\theta)}{\partial p^2}\right)} = \frac{pq}{n}.$$

Так как нижняя граница оценки совпадает с дисперсией относительной частоты, то относительная частота является эффективной оценкой.

7.3. МЕТОДЫ ПОЛУЧЕНИЯ ОЦЕНОК

Одним из методов получения оценок неизвестных параметров распределения по выборке является *метод моментов*. Он основывается на том, что эмпирические моменты (или их функции) принимаются за оценки соответствующих теоретических моментов (или их функций) и параметры выражаются через эти моменты.

Эмпирические начальные моменты порядка k определяются формулами:

$$\nu_k^* = \overline{x}^k = \frac{1}{n} \sum_i x_i^k,$$

а соответствующие им теоретические начальные моменты порядка k – формулами:

$$\nu_k = \begin{cases} \sum\limits_i x_i^k p_i\ (x,\ \theta) & \text{для дискретных СВ,} \\ \int\limits_{-\infty}^{\infty} x^k f(x, \theta)\, dx & \text{для непрерывных СВ.} \end{cases}$$

Для нахождения оценок параметров функции распределения $F(x, \theta_1, \theta_2)$, содержащей два неизвестных параметра θ_1 и θ_2, составляется система из двух уравнений

$$\left.\begin{aligned} \nu_1(\hat{\theta}_1, \hat{\theta}_2) &= \nu_1^*(\hat{\theta}_1, \hat{\theta}_2), \\ \mu_2(\hat{\theta}_1, \hat{\theta}_2) &= \mu_2^*(\hat{\theta}_1, \hat{\theta}_2), \end{aligned}\right\}$$

решением которой является оценка $\hat{\theta}_1$, $\hat{\theta}_2$ функции распределения $F(x, \theta_1, \theta_2)$.

Приравнивая теоретический и эмпирический начальные моменты первого порядка, находим, что оценкой математического ожидания СВ X, имеющей произвольное распределение, является среднее арифметическое наблюдений СВ X, т.е. $M(X) = \overline{x} = \frac{1}{n} \sum\limits_i x_i$. Аналогично, приравнивая теоретический и эмпирический центральные моменты второго порядка, находим, что оценка дисперсии СВ X, имеющей произвольное распределение, определяется по формуле

$$D(X) = \sigma_x^2 = S^2 = \Sigma (x_i - x)^2/n .$$

Таким же образом можно найти оценки теоретических моментов любого порядка.

Пример 7.2. Двухмерная СВ (X, Y) имеет нормальное распределение с плотностью распределения вероятностей

$$f(x,y) = (2\pi\sigma_x\sigma_y\sqrt{1-\rho^2})^{-1}\exp((-\frac{1}{2(1-\rho^2)})\times$$
$$\times(\frac{(x-m_x)^2}{\sigma_x^2} - 2\rho\frac{(x-m_x)(y-m_y)}{\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2})) .$$

Оценить по результатам испытаний двухмерной СВ (x_i, y_i), $i = 1, 2, ..., n$, параметры $m_x, m_y, \sigma_x^2, \sigma_y^2, \rho$.

Решение. Используя метод моментов, определяем:

$$\nu_{10} = \nu_{10}^*, \qquad m_x = \overline{x} = \frac{1}{n}\Sigma x_i ,$$
$$\nu_{01} = \nu_{01}^* , \qquad m_y = \overline{y} = \frac{1}{n}\Sigma y_i ,$$
$$\mu_{02} = \mu_{02}^* , \qquad \hat{\sigma}_x^2 = \frac{1}{n}\Sigma(x_i - \overline{x})^2 ,$$
$$\mu_{20} = \mu_{20}^*, \qquad \sigma_y^2 = S_y^2 = \frac{1}{n}(y_i - \overline{y})^2 ,$$
$$\frac{\mu_{11}}{\sigma_x\sigma_y} = \frac{\mu_{11}^*}{S_xS_y} , \qquad \hat{\rho} = r = \frac{1/n\,\Sigma(x_i - \overline{x})(y_i - \overline{y})}{S_xS_y} .$$

Другим методом получения оценок является *метод наибольшего правдоподобия,* суть которого состоит в следующем: для получения оценки неизвестного параметра θ необходимо найти такое значение $\hat{\theta}$, при котором вероятность реализации полученной выборки $X_1, X_2, ..., X_n$ (n-мерной СВ) была бы максимальной. Предположим, что составляющие СВ независимы, тогда вероятность того, что они примут значения, равные наблюдаемым,

$$L = P(X = x_1, X_2 = x_1, ..., X_n = x_n) =$$
$$= P(x_1, \theta)P(x_2, \theta)\cdots P(x_n, \theta). \qquad (7.3)$$

Величина (7.3) называется *функцией правдоподобия,* а $\hat{\theta}$, являющаяся точкой максимума этой функции, – *оценкой, полученной по методу наибольшего правдоподобия.* В случае непрерывной СВ функция правдоподобия имеет вид

$$L = f(x_1, x_2, ..., x_n) = f(x_1, \theta)f(x_2, \theta)\cdots f(x_n, \theta). \qquad (7.4)$$

Выражение (7.4) определяет плотность распределения вероятностей непрерывной СВ $(X_1, X_2, ..., X_n)$ или плотность распределения выборки. Для нахождения оценок в этом случае составляется система m уравнений (m – число оцениваемых параметров):

$$\frac{\partial L}{\partial \theta_i} = 0, \quad i = 1, 2, ..., m,$$

и выбирается то решение, при котором функция правдоподобия достигает максимума.

Поскольку экстремум функции L и $\ln L$ достигается при одних и тех же значениях $\theta = u(x_1, x_2, ..., x_n)$, то для упрощения расчетов иногда пользуются логарифмической функцией правдоподобия.

Метод наибольшего правдоподобия дает состоятельные оценки. Недостаток метода заключается в том, что иногда оценки наибольшего правдоподобия являются смещенными; кроме того, для нахождения оценок часто приходится решать сложные системы уравнений.

Пример 7.3. Случайная величина X имеет показательное распределение с плотностью распределения вероятностей $f(x, \theta) = \theta e^{-\theta x}$. Оценить параметр θ по результатам наблюдений $x_1, x_2, ..., x_n$.

Р е ш е н и е. Запишем функцию правдоподобия:

$$L = \theta e^{-\theta x_1} \theta e^{-\theta x_2} \dots \theta e^{-\theta x_n} = \theta^n e^{-\theta \Sigma x_i}.$$

Логарифмируя ее, получаем

$$\ln L = n \ln \theta - \theta \Sigma x_i .$$

Дифференцируя последнее равенство по параметру θ, находим

$$\frac{\partial \ln L}{\partial \theta} = \frac{n}{\theta} - \Sigma x_i .$$

Приравнивая производную нулю, получаем

$$\frac{n}{\theta} - \Sigma x_i = 0 \Rightarrow \hat{\theta} = \frac{n}{\Sigma x_i} = \frac{1}{\bar{x}}.$$

Метод наименьших квадратов (*МНК*) также применяется для оценивания параметров зависимостей между СВ. Суть его состоит в следующем. Фиксируется некоторая СВ $X(x_1, x_2, ..., x_n)$ и измеряется некоторая СВ $Y(y_1, y_2, ..., y_n)$, которая связана с СВ X следующей зависимостью:

$$y_i = f(x_i, \theta) + \epsilon_i ,$$

где $f(x_i, \theta)$ – известная функция; θ – оцениваемый параметр; ϵ_i – случайная ошибка в i-м измерении. Необходимо по результатам эксперимента $\{x_i, y_i, i = 1, 2, ..., n\}$ найти оценку $\hat{\theta}$. В качестве критерия используется сумма квадратов

$$L(\theta) = \Sigma (y_i - f(x_i, \theta))^2 . \quad (7.5)$$

В качестве оценки для параметра θ принимается такое $\hat{\theta}$, для которого сумма квадратов минимальна, т.е. $L(\hat{\theta}) = \min L(\theta)$.

Рассмотрим линейную модель

$$y_i = \sum_{j=1}^{n} \theta_j x_{ij} + \epsilon_i, \quad i = 1, 2, ..., n, \tag{7.6}$$

где i – номер эксперимента; j – номер переменной; x_{ij} – значение j-й переменной в i-м эксперименте. В матричной форме записи вектор оцениваемых параметров $\theta^T = (\theta_1 \;\; \theta_2 \;\; ... \;\; \theta_m)$, вектор результатов эксперимента $Y^T = (y_1 \;\; y_2 \; ... \; y_n)$, а матрица результатов наблюдений

$$X = \begin{pmatrix} x_{11} & x_{21} & \dots & x_{m1} \\ \dots & \dots & \dots & \dots \\ x_{1n} & x_{2n} & \dots & x_{mn} \end{pmatrix}.$$

Если матрица X – невырожденная, т.е. $|X^T X| \neq 0$, то оценкой по методу наименьших квадратов для параметров модели (7.6) является *статистика*

$$\hat{\theta} = (X^T X)^{-1} X^T Y. \tag{7.7}$$

Оценка (7.7) является несмещенной. Наилучшим в среднем квадратичном смысле прогнозом для θ будет статистика $\hat{Y} = \sum \hat{\theta}_j x_j$.

7.4. ИНТЕРВАЛЬНЫЕ ОЦЕНКИ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ. ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ

При определении точечной оценки мы допускаем ошибку, так как $\hat{\theta} = u(x_1, x_2, ..., x_n)$ является случайной функцией. В связи с этим во многих случаях удобно пользоваться интервальной оценкой, основанной на определении некоторого интервала, внутри которого с данной вероятностью находится неизвестное значение параметра θ. Естественно, чем меньше разность $|\theta - \hat{\theta}|$, тем лучше и точнее оценка. Таким образом, положительное число ϵ характеризует точность оценки

$$|\theta - \hat{\theta}| < \epsilon. \tag{7.8}$$

Задаваясь вероятностью $1 - \alpha$ (*доверительной вероятностью*), с которой неравенство (7.8) должно выполняться, определим доверительный интервал. *Доверительным интервалом* называется интервал $]\hat{\theta} - \epsilon; \; \hat{\theta} + \epsilon[$, накрывающий неизвестный параметр θ с заданной доверительной вероятностью $p = 1 - \alpha$. Наиболее часто полагают $1 - \alpha = 0{,}95$; $0{,}99$; $0{,}9973$.

Для параметров нормального закона распределения вероятностей СВ X пользуются следующей схемой построения доверительных интервалов:

1) по результатам выборки с функцией распределения $F(x, \theta)$ методом наибольшего правдоподобия или методом моментов находят точечную оценку оцениваемого параметра θ;

2) составляют СВ $Y(\theta)$, имеющую плотность распределения вероятностей $f(y, \theta)$;

3) задают доверительную вероятность $1-\alpha$;

4) используя $f(y, \theta)$, находят такие C_1 и C_2, что

$$P(C_1 < X < C_2) = \int_{C_1}^{C_2} f(y, \theta)\, dy = 1 - \alpha . \tag{7.9}$$

Значения C_1 и C_2 получают из условия:

$$P(Y(\theta) < C_1) = \alpha/2, \quad P(Y(\theta) > C_2) = \alpha/2.$$

Используя указанную схему, определим доверительный интервал для математического ожидания a СВ X, имеющий нормальное распределение при известном σ:

$$P(|\bar{x} - a| < \epsilon) = \frac{1}{\sqrt{2\pi}} \int_{-\epsilon\sqrt{n}/\sigma}^{\epsilon\sqrt{n}/\sigma} e^{-x^2/2} dx = 2\Phi(\epsilon\sqrt{n}/\sigma) .$$

Обозначим $u = (\bar{x} - a)(\sigma/\sqrt{n})^{-1}$ (выборочная статистика), тогда доверительный интервал при заданной доверительной вероятности $1-\alpha$ имеет вид

$$\boxed{\left]\, \bar{x} - u_{\alpha/2} \frac{\sigma}{\sqrt{n}} ; \quad \bar{x} + u_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right[.}$$

Он накрывает известное математическое ожидание a с заданной вероятностью $1-\alpha$. Точность оценки математического ожидания (предельная погрешность) $\epsilon = u_{\alpha/2}\, \sigma/\sqrt{n}$.

Приведем квантили стандартизованного нормального распределения для наиболее употребительных значений доверительной вероятности $1-\alpha$:

Доверительная вероятность	Квантили $u_{\alpha/2}$
0,90	1,64
0,95	1,96
0,99	2,58
0,9973	3,00
0,999	3,37

Пример 7.4. Случайная величина X распределена нормально со средним квадратичным отклонением $\sigma = 2$. Найти доверительный интервал для математического ожидания a, если $n = 16$, $x = 20{,}09$. Доверительную вероятность $1-\alpha$ принять равной 0,90.

Р е ш е н и е. По таблице значений функции Лапласа (прил. 2) и заданной вероятности $1-\alpha=0{,}9$ находим $u_{0{,}05}=1{,}64$. Точность оценки $\epsilon = 1{,}64\cdot\frac{2}{\sqrt{16}} = 0{,}82$. Следовательно, искомый доверительный интервал $]20{,}09-0{,}82;\ 20{,}09+0{,}82[$.

Доверительный интервал для математического ожидания СВ X, нормально распределенной при неизвестном σ, с заданной вероятностью $1-\alpha$ имеет вид:

$$\boxed{\left]\overline{x}-t_{\alpha/2;\,n-1}\frac{s}{\sqrt{n-1}},\ \overline{x}+t_{\alpha/2,\,n-1}\frac{s}{\sqrt{n-1}}\right[\ ,}$$

где $s=\sqrt{\Sigma(x_i-\overline{x})^2/n}$; $t=\dfrac{\overline{x}-a}{s}/\sqrt{n-1}$ – вспомогательная функция (выборочная статистика) распределения по закону Стьюдента (прил. 3) с ν степенями свободы: $\nu=n-1$. В прил. 4 помещены квантили t-распределения, соответствующие доверительной вероятности $1-\alpha$ и объему выборки n. Точность (предельная погрешность) оценки математического ожидания $\epsilon = t_{\alpha/2;\,n-1}\dfrac{s}{\sqrt{n-1}}$. Если на основании выборки найдена несмещенная оценка $s=\hat{\sigma}=\sqrt{\Sigma(x_i-x)^2/(n-1)}$, то доверительный интервал имеет вид

$$\boxed{\left]\overline{x}-t_{\alpha/2;\,n-1}\frac{s}{\sqrt{n}};\ \ \overline{x}+t_{\alpha/2;\,n-1}\frac{s}{\sqrt{n}}\right[\ ,}\ .$$

где $t=\dfrac{\overline{x}-a}{s}\sqrt{n}$.

Пример 7.5. Известны результаты измерений некоторого диаметра: $x_1=2$, $x_2=2{,}035$, $x_3=2{,}015$, $x_4=2{,}035$, $x_5=2{,}015$. Оценить с помощью доверительного интервала истинный диаметр, приняв $1-\alpha=0{,}095$.

Р е ш е н и е. Считаем, что СВ X распределена нормально, с неизвестными параметрами a и σ. Находим точечные оценки этих параметров:

$$\hat{a}=\overline{x}=\Sigma x_i/5=10{,}095/5=2{,}019,$$

$$\hat{\sigma}=s=\sqrt{0{,}00007/5}=0{,}0036.$$

По таблицам (прил. 4), доверительной вероятности $1-\alpha=0{,}95$ и числу степеней свободы $\nu=n-1=4$ находим квантиль распределения $t_{0{,}025;\,4}=2{,}776$. Следовательно, предельная погрешность

$$\varepsilon = t_{\alpha/2;\,n-1}\frac{s}{\sqrt{n-1}} = 2{,}776\cdot\frac{0{,}0036}{2} = 0{,}0051.$$

Искомый доверительный интервал $]\overline{x}-\epsilon;\ \overline{x}+\epsilon[\ = \ ]2{,}0139;\ 2{,}0241[$.

Доверительный интервал для среднего квадратичного отклонения σ СВ X, имеющей нормальное распределение, определяется по формуле

$$s\gamma_1<\sigma<s\gamma_2 \quad \text{или} \quad]s\gamma_1;\ s\gamma_2[\ ,$$

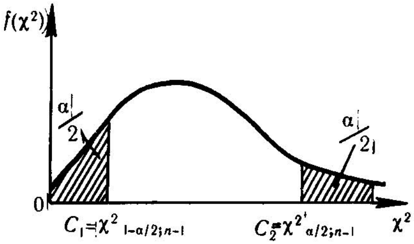


Рис. 7.1

где

$$\gamma_1 = \sqrt{\frac{n-1}{\chi^2_{\alpha/2;\,n-1}}};$$

$$\gamma_2 = \sqrt{\frac{n-1}{\chi^2_{1-\alpha/2;\,n-1}}}.$$

Вспомогательная случайная величина χ^2 имеет распределение с $n-1$ степенями свободы $\chi^2 = (n-1)s^2/\gamma^2$ (прил. 3). Вероятность того, что СВ χ^2 попадает в интервал $]\,c_1;\ c_2\,[$ (рис. 7.1), равна $P(c_1 < \chi^2 < c_2) = \int\limits_{c_1}^{c_2} f(\chi^2)\,d(\chi^2)$.

Коэффициенты γ_1 и γ_2 соответствуют доверительной вероятности $1-\alpha$ и числу степеней свободы $\nu = n-1$ (прил. 7).

Пример 7.6. Построить доверительный интервал, накрывающий среднее квадратичное отклонение σ с надежностью $1-\alpha = 0{,}95$, если $n = 5$, $\bar{x} = 2{,}019$, $\sigma = s = 0{,}0036$.

Решение. По заданной доверительной вероятности 0,95 и числу степеней свободы $\nu = n-1 = 4$ из прил. 7 находим с точностью до 0,01: $\gamma_1 = 0{,}6$, $\gamma_2 = 2{,}87$. Следовательно, искомый доверительный интервал

$$0{,}6\cdot 0{,}0036 < \sigma < 2{,}87\cdot 0{,}0036, \text{ т.е. } 0{,}0022 < \sigma < 0{,}103.$$

ЗАДАЧИ

7.1. Случайная величина имеет нормальное распределение со средним квадратичным отклонением $\sigma = 4$. Требуется найти доверительный интервал для математического ожидания a, если $n = 25$, $\bar{x} = 12{,}8$. Доверительную вероятность $1-\alpha$ принять равной 0,9.

7.2. Найти минимальный объем выборки, на основании которой можно было бы оценить с надежностью $1-\alpha = 0{,}95$ математическое ожидание СВ X – контролируемого размера – с ошибкой, не превышающей 26 мкм. Предположить, что СВ X имеет нормальное распределение с параметрами a, $\sigma = 50$.

7.3. Найти минимальное число измерений, которые надо произвести, чтобы с надежностью 0,95 можно было утверждать, что предельная погрешность точечной оценки измерения некоторого зазора не превышает 0,003, если известны результаты пяти измерений зазора: $x_1 = 2{,}005$, $x_2 = 2{,}030$, $x_3 = 2{,}025$, $x_4 = 2{,}020$, $x_5 = 2{,}015$.

7.4. Даны результаты измерения некоторого угла α в градусах:

Номер измерения	α	Номер измерения	α	Номер измерения	α	Номер измерения	α
1	3,1	7	2,8	13	2,9	19	2,9
2	3,3	8	2,7	14	3,1	20	3,1
3	2,9	9	3,1	15	2,8	21	3,2
4	3,0	10	3,2	16	2,9	22	3,0
5	3,1	11	2,9	17	3,2		
6	3,2	12	3,0	18	3,3		

Найти доверительный интервал, накрывающий среднее квадратичное отклонение σ с надежностью 0,98.

8. СТАТИСТИЧЕСКИЕ РЕШЕНИЯ

8.1. ПОНЯТИЕ О КРИТЕРИЯХ СОГЛАСИЯ. СТАТИСТИЧЕСКИЕ ГИПОТЕЗЫ

Результаты, полученные при обработке выборки из некоторой совокупности, могут быть использованы для нахождения решений относительно всей совокупности, чаще всего в виде функций распределения. Подобные решения называются *статистическими.* Вследствие вероятностного характера принимаемое решение может быть ошибочным, т.е. могут быть расхождения между статистическими и теоретическими функциями распределения. Для оценки характера расхождений используют критерии согласия.

Статистической называется гипотеза о предполагаемом виде неизвестного распределения или о параметрах известных распределений. Например, утверждения "генеральная совокупность распределена по закону Пуассона" или "дисперсии двух совокупностей, подчиняющихся нормальному закону, равны между собой" – статистические гипотезы.

Нулевой называется выдвинутая гипотеза, *конкурирующей* – гипотеза, которая противоречит нулевой.

Простой называется гипотеза, содержащая только одно предположение. *Сложная гипотеза* состоит из конечного или бесконечного числа простых гипотез.

Критерием согласия статистической гипотезы с опытными данными называют правило, позволяющее на основании выборки принять или отвергнуть нулевую гипотезу.

В *статистической оценке гипотезы* могут быть допущены ошибки. Ошибка первого рода состоит в том, что будет отвергнута правильная гипотеза, ошибка второго рода – в том, что будет принята неверная гипотеза. Вероятность совершить ошибку первого рода обозначается буквой α и называется *уровнем значимости критерия.* Наиболее часто уровень значимости принимают равным 0,05 или 0,01 (5 % или 1 %).

Существует несколько критериев. За меру расхождения (статистику) между теоретической функцией распределения и статистическим распределением выбирается некоторая СВ.

Критической областью называется совокупность значений статистики критерия, при которых нулевую гипотезу отвергают.

Областью принятия гипотезы (областью допустимых значений) называется совокупность значений статистики, при которых гипотезу принимают.

Критическими называются точки, отделяющие критические области от областей принятия гипотезы.

Мощностью критерия называется вероятность попадания статистики в критическую область при условии, что справедлива конкурирующая гипотеза, т.е. мощность критерия – это вероятность отвергнуть нулевую гипотезу, если верна конкурирующая.

8.2. КРИТЕРИЙ СОГЛАСИЯ ПИРСОНА

За меру расхождения между теоретическим и статистическим распределениями – статистику – выбирают СВ

$$\chi_q^2 = \sum_{i=1}^{l} \frac{(m_i - m_i')^2}{m_i'},$$

которая, как доказал Пирсон, при достаточно большом объеме выборки имеет закон распределения, приближающийся к χ^2-распределению (см. прил. 3), независимо от вида функции $F(x)$.

Приведем пример применения критерия согласия Пирсона к проверке гипотезы о нормальном распределении генеральной совокупности. Пусть имеется эмпирическое распределение из равноотстоящих вариант и соответствующих им частот:

x_i	x_1	x_2	...	x_l
m_i	m_1	m_2	...	m_l

Проверка гипотезы о нормальном распределении сводится к следующему алгоритму:

1) вычислить теоретические частоты по формуле

$$m_i' = \frac{nh}{\sigma_в} \varphi(u_i),$$

где n – объем выборки; h – шаг, равный разности между двумя соседними вариантами; $\sigma_в$ – выборочное среднее квадратичное отклонение; $u_i = (x_i - \overline{x}_в)/\sigma_в$; $\overline{x}_в$ – выборочное среднее; $\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$;

2) вычислить величину χ_q^2:

$$\chi_q^2 = \sum_{i=1}^{l} \frac{(m_i - m_i')^2}{m_i'};$$

3) найти число степеней свободы $k = l - 3$ (l – число различных значений x_i);

4) выбрать уровень значимости;

5) определить из прил. 10 по данным k и χ_q^2 вероятность $P(\chi^2 \geqslant \chi_q^2)$. Если эта вероятность меньше принятого уровня значимости, то гипотеза отвергается, если больше – принимается.

Если вариационный ряд непрерывен, то проверка гипотезы о нормальном распределении сводится к следующему алгоритму:

1) вычислить выборочное среднее $\overline{x}_в$ и выборочное среднее квадратичное отклонение $\sigma_в^*$, причем вместо вариант $x_в^*$ взять среднее арифметическое концов интервала $x_i^* = (x_i + x_{i+1})/2$;

2) нормировать СВ X, т.е. перейти к новой СВ $Y = (X - X^*)/\sigma_в^*$, причем наименьшее значение приравнять к $-\infty$, а наибольшее – к $+\infty$;

3) вычислить теоретические вероятности попадания Y в интервалы $]y_i;\ y_{i+1}[$

$$P(y_i < Y < y_{i+1}) = \Phi(y_{i+1}) - \Phi(y_i),$$

где $\Phi(y)$ – функция Лапласа;

4) вычислить

$$\chi_q^2 = \sum_{i=1}^{l} \frac{(m_i - np_i)^2}{np_i},$$

где n – объем выборки;

5) найти число степеней свободы $k = l - 3$, где l – число интервалов выборки;

6) выбрать уровень значимости;

7) по значениям k, χ_q^2 в прил. 10 найти $P(\chi^2 \geqslant \chi_q^2)$. Если $P(\chi^2 \geqslant \chi_q^2) > \alpha$, то гипотеза принимается, если $P(\chi^2 \geqslant \chi_q^2) < \alpha$ – отвергается.

Таким образом, обработка статистических данных выборочным методом предполагает:

1) составление вариационного ряда наблюденных значений изучаемой величины;

2) построение эмпирической кривой (полигона) распределения;

3) вычисление характеристик эмпирического распределения;

4) определение по виду графика эмпирической кривой распределения теоретического распределения, к которому приближается эмпирическая кривая;

5) оценку близости эмпирического распределения к предполагаемому теоретическому.

В случае непрерывных СВ обработку статистических данных можно провести также по следующей схеме:

1) составление вариационного ряда;

2) определение эмпирической функции;

3) построение графика эмпирической функции;

4) построение графика теоретической функции распределения путем замены ломаной линии на кривую;

5) определение уравнения кривой распределения;

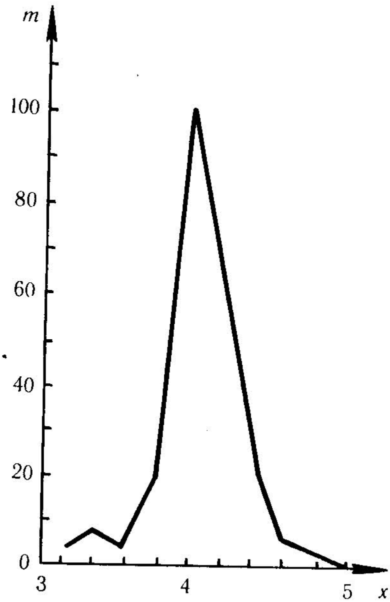
6) определение дифференциального закона распределения.

Пример 8.1. Пользуясь критерием Пирсона, при уровне значимости 0,05 установить, согласуется ли гипотеза о нормальном распределении генеральной совокупности со следующими данными выборки объемом $n = 200$:

x_i	3,2	3,4	3,6	3,8	4,0	4,2	4,4	4,6	4,8	5,0
m_i	1	5	4	18	86	62	14	6	3	1

Р е ш е н и е. Так как выборка представлена в виде вариационного ряда, строим эмпирическую кривую (полигон частот). Вычислим характеристики эмпирического распределения – выборочное среднее и среднее квадратичное отклонение:

$$\overline{x}_в = \sum_{i=1}^{10} x_i \frac{m_i}{n} = 4{,}08, \quad D_в = \overline{x}_в^2 - (\overline{x}_в)^2 = 0{,}0648; \quad \sigma_в = \sqrt{D_в} = 0{,}2545.$$



Р и с. 8.1.

По виду кривой (полигона) предполагаем, что теоретическое распределение подчиняется нормальному закону (рис. 8.1).

Оценим близость эмпирического распределения к теоретическому с помощью критерия Пирсона. Для этого составим таблицу (табл. 8.1).

Таблица 8.1

i	x_i	m_i	$u_i=(x_i-\bar{x})/\sigma$	$\varphi(u_i)$	$m_i'=\dfrac{nh}{\sigma}\varphi(u_i)$	m_i-m_i'	$(m_i-m_i')^2$	$(m_i-m_i')^2/m_i$
1	3,2	1	−3,464	0,0010	0,1547	0,843	0,711	4,517
2	3,4	5	−2,677	0,0113	1,778	3,222	23,261	10,3
3	3,6	4	−1,889	0,0681	10,718	−6,718	45,131	4,210
4	3,8	18	−1,102	0,2179	34,297	−16,297	265,59	7,714
5	4,0	86	−0,315	0,3790	59,654	27,654	764,74	11,144
6	4,2	62	0,472	0,3555	55,955	7,955	63,282	1,130
7	4,4	14	1,259	0,1804	28,394	−14,394	207,187	7,296
8	4,6	6	2,047	0,0488	7,681	−1,681	2,825	0,358
9	4,8	3	2,834	0,0073	1,149	1,851	3,426	2,284
10	4	1	3,622	0,0006	0,0944	0,9056	0,820	8,72
							$\Sigma=$	57,673

Так как $k = 10 - 3 = 7$, $\chi_q^2 = 57{,}673$, из прил. 10 видно, что $P(\chi^2 \geqslant \chi_q^2) = 0$, т.е. меньше, чем 0,05. Значит, распределение отлично от нормального.

Пример 8.2. Произведен выбор 200 метчиков из сменного задания рабочего. Проверяемый размер измерен с точностью до 1 мкм. В табл. 8.2 приведены отклонения x_i от номинального размера, разбитые на варианты, частоты, соответствующие им, и их вероятности. Оценить с помощью критерия Пирсона гипотезу о согласии выборочного распределения с законом нормального распределения при уровне значимости 0,05.

Решение. Определяем значения середин интервалов и находим оценки математического ожидания и дисперсии:

$$M(x) = \sum_{i=1}^{10} x_i p_i \ , M(\overline{x}^2) = \sum_{i=1}^{10} x_i^2 p_i .$$

Таблица 8.2

i	$x_i - x_{i+1}$	m_i	p_i
1	−20 − −15	7	0,035
2	−15 − −10	11	0,055
3	−10 − −5	15	0,075
4	−5 − 0	24	0,120
5	0 − 5	49	0,245
6	5 − 10	41	0,205
7	10 − 15	26	0,130
8	15 − 20	17	0,085
9	20 − 25	7	0,035
10	25 − 30	3	0,015

Вычисления сводим в таблицу (табл. 8.3).

Теоретические вероятности p_i попадания отклонений в интервалы $]x_i;\ x_{i+1}[$ вычисляем по формуле

$$p_i = \Phi(y_{i+1}) - \Phi(y_i),$$

где y_i – левая граница i-го интервала относительно $\overline{x}$ в единицах $\overline{\sigma}$: $y_i = (x_i - \overline{x}) / \overline{\sigma}$. При этом наименьшее $y_i = y_1 = -2{,}5$ заменяем на $-\infty$, а наибольшее $y_i = 2{,}65$ – на $+\infty$.

Таблица 8.3

i	x_i	y_i	$\Phi(y)$	p_i	np_i	$(m_i - np_i)^2/(np_i)$
1	−17,5	$-\infty$	−0,5000	0,0239	4,78	1,04
2	−12,5	−1,99	−0,4761	0,0469	9,38	0,28
3	−7,5	−1,47	−0,4292	0,0977	19,54	1,05
4	−2,5	−0,96	−0,3315	0,1615	32,30	2,13
5	2,5	−0,44	−0,1700	0,1979	39,58	2,24
6	7,5	0,07	0,0279	0,1945	38,90	0,11
7	12,5	0,59	0,2224	0,1419	28,38	0,20
8	17,5	1,10	0,3643	0,0831	16,62	0,01
9	22,5	1,62	0,4474	0,0526	10,52	0,03
10	27,5	2,13	0,4834	–		
11	–	$+\infty$	0,5000	–	200	$\chi_q^2 = 7{,}09$

Значения функции Лапласа $\Phi(y)$ находим из прил. 2. Последний интервал объединяем с предпоследним, следовательно, $s = 9$.

Вычисляем

$$\chi_q^2 = \sum_{i=1}^{9} \frac{(m_i - np_i)^2}{np_i} = 7{,}09.$$

Число степеней свободы $k = 9 - 2 - 1 = 6$. Из прил. 10 по входным величинам χ_q^2 и k находим $P(\chi^2 \geq \chi_q^2) = 0{,}313$. Гипотеза о нормальном отклонении от номинального размера не противоречит наблюдениям.

ЗАДАЧИ

8.1. С помощью критерия Пирсона при уровне значимости 0,05 установить, согласуется ли гипотеза о нормальном распределении генеральной совокупности с данными выборки:

x_i	5	7	9	11	13	15	17	19	21
m_i	15	26	25	30	26	21	24	20	13

8.2. Произвести статистическую обработку данных задачи 8.1 и построить теоретическую кривую.

8.3. Отсчет по шкале измерительного прибора оценивается приблизительно в долях деления шкалы. Приведено 200 результатов отсчета последней цифры между соседними делениями шкалы:

x_i	0	1	2	3	4	5	6	7	8	9
m_i	35	16	15	17	17	19	11	6	30	24

Установить, используя критерий Пирсона, согласуются ли наблюдения с законом равномерного распределения, если при уровне значимости 0,05 принимается вероятность появления любой цифры $p_i = 0{,}10$.

8.4. Приведены отклонения внутренних диаметров шестерен, обработанных на станке, от заданного размера:

Границы интервала, мк (x_i)	0–5	5–10	10–15	15–20	20–25
Численность вариант (m_i)	15	75	100	50	10
Частота (p_i)	0,06	0,30	0,40	0,20	0,04

Проверить, используя критерий Пирсона, гипотезу о согласии наблюдений с законом нормального распределения, приняв уровень значимости 0,05.

8.3. КРИТЕРИЙ СОГЛАСИЯ КОЛМОГОРОВА

Критерий согласия Колмогорова применяется для проверки гипотез о законах распределения только непрерывных СВ. Отличие от критерия согласия χ^2 состоит в том, что сравниваются эмпирическая функция $F^*(x)$ и теоретическая функция $F(x)$ при известных параметрах распределения.

В качестве меры расхождения между теоретическим и статистическим распределениями рассматривается максимальное значение модуля разности $D = \max |F^*(x) - F(x)|$.

А.Н. Колмогоров доказал, что какова бы ни была функция распределения $F(x)$ непрерывной СВ X, при неограниченном увеличении числа независимых наблюдений n вероятность неравенства $P(D\sqrt{n} \geqslant \lambda)$ стремится к пределу

$$P(\lambda) = 1 - \sum_{n \to -\infty}^{\infty} (-1)^n e^{-2n^2\lambda^2} .$$

Уровень значимости α принимается чаще всего в пределах 0,10–0,20.

Проверку гипотезы с помощью критерия Колмогорова проводят в следующем порядке:

1) располагают результаты наблюдений по возрастанию их значений в виде интервального вариационного ряда;

2) находят эмпирическую функцию распределения $F^*(x) = n_x/n$;

3) вычисляют, пользуясь предполагаемой функцией $F(x)$, значения теоретической функции распределения, соответствующие наблюдаемым значениям СВ X;

4) находят для каждого x_i модуль разности между эмпирической и теоретической функциями распределения;

5) определяют величину $\lambda = D\sqrt{n} = \max|F^*(x) - F(x)|\sqrt{n}$;

6) находят критические значения λ_α в зависимости от уровня значимости из табл. 8.4.

Таблица 8.4

α	0,50	0,40	0,30	0,20	0,10	0,05	0,02	0,01	0,001
λ_α	0,828	0,895	0,974	1,073	1,224	1,358	1,510	1,627	1,950

Если опытное значение $\lambda \geqslant \lambda_\alpha$, то гипотеза о согласии теоретического закона распределения с данными выборки опровергается. Если $\lambda < \lambda_\alpha$, то гипотеза принимается.

Пример 8.3. Результаты измерения 1000 деталей помещены в табл. 8.5.

Таблица 8.5

i	x_i	m_i	i	x_i	m_i	i	x_i	m_i
1	98,0	21	5	100,0	181	8	101,5	97
2	98,5	47	6	100,5	201	9	102,0	41
3	99,0	87	7	101,0	142	10	102,5	25
4	99,5	158						

Пользуясь критерием согласия Колмогорова, проверить следующую гипотезу; СВ X подчиняется нормальному закону с математическим ожиданием $M(x) = a = 100{,}25$ мм и средним квадратичным отклонением $\sigma = 1$ мм при уровне значимости $\alpha = 0{,}05$.

Р е ш е н и е. Результаты наблюдений в порядке возрастания их значений помещены в табл. 8.6.

Таблица 8.6

i	$x_i - a$	$\frac{1}{2}\Phi(x_i-a)$	$F(x_i)$	$F^*(x_i)$	$\|F^*(x_i) - F(x_i)\|$
1	−2,25	−0,4877	0,0123	0,0105	0,0018
2	−1,75	−0,4599	0,0401	0,0445	0,0044
3	−1,25	−0,3944	0,1056	0,1115	0,0059
4	−0,75	−0,2734	0,2266	0,2340	0,0074
5	−0,25	−0,0987	0,4013	0,4035	0,0022
6	0,25	0,0987	0,5987	0,5945	0,0042
7	0,75	0,2734	0,7734	0,7660	0,0074
8	1,25	0,3944	0,8944	0,8855	0,0089
9	1,75	0,4599	0,9599	0,9545	0,0054
10	2,25	0,4877	0,9877	0,9875	0,0002

Эмпирическую функцию распределения $F^*(x)$ вычисляем по формуле

$$F^*(x_k) = \sum_{i=1}^{k} m_i/1000 .$$

Теоретическую функцию распределения $F(x)$ определяем по формуле $F(x) = \frac{1}{2} + \frac{1}{2}\Phi(x-a)$, где Φ – функция Лапласа. Составляем для каждого значения x_i разности $F^*(x_i) - F(x_i)$. Результаты вычислений помещены в табл. 8.6.

Из табл. 8.6 выбираем наибольшую из разностей $|F^*(x_i) - F(x_i)| = D = 0{,}0089$. Определяем $\lambda = D\sqrt{n} = 0{,}0089\sqrt{1000} = 0{,}281$. По значению $\alpha = 0{,}05$ находим из таблицы $\lambda_\alpha = 1{,}073$. Видим, что $\lambda < \lambda_\alpha$, следовательно, гипотеза о согласии наблюдений с законом нормального распределения с параметрами $a = 100{,}25$ и $\sigma = 1$ не опровергается.

ЗАДАЧИ

8.5. В табл. 8.7 приведены статистические данные по отклонению диаметров (мкм) валиков, обработанных на станке, от заданного размера.

Таблица 8.7

Границы интервала (x_i)	0–5	5–10	10–15	15–20	20–25
Численность (m_i)	15	75	100	50	10
Частость (ω_i)	0,06	0,30	0,40	0,20	0,04

Проверить гипотезу о согласии наблюдений с законом нормального распределения, используя критерий Колмогорова и приняв уровень значимости $\alpha = 0{,}05$.

8.6. Из таблицы случайных чисел выбрано 150 натуральных чисел, меньших 100. Разбив все элементы выборки на 10 интервалов и при этом приняв значения чисел, попавших в один интервал, равными его середине, проверить гипотезу о согласии наблюдений с законом равномерного распределения при уровне значимости $\alpha = 0{,}05$. Результаты выборки помещены в табл. 8.8.

Таблица 8.8

Границы интервала	Частота m_i	Частость ω_i	Границы интервала	Частота m_i	Частость ω_i
0–9	16	0,107	50–59	19	0,127
10–19	15	0,100	60–69	14	0,093
20–29	19	0,127	70–79	11	0,073
30–39	13	0,087	80–89	13	0,087
40–49	14	0,093	90–99	16	0,107

8.7. Данные по выборке 150 отклонений диаметров от номинального размера (мкм) имеют вид:

Середина интервала x_i	26	29	32	35	38	41	47	50	53
Частота m_i	1	4	13	23	22	29	16	11	2

При уровне значимости $\alpha = 0{,}10$ проверить гипотезу H_0 о нормальном распределении генеральной совокупности.

8.4. ОДНОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ

Задачей дисперсионного анализа является изучение влияния одного или нескольких факторов на рассматриваемый признак (наблюдаемую случайную величину).

Пусть на количественный нормально распределенный признак $\mathbf{X}(X_1, X_2, \ldots, X_k)$ с одинаковыми дисперсиями σ^2 воздействует фактор F. Используя статистические данные, требуется проверить гипотезу H_0 о том, что математические ожидания всех случайных величин x_i, $i = 1, 2, \ldots, k$, одинаковы. Если гипотеза верна, то статистические средние каждой серии не должны значительно отличаться друг от друга. Если же статистические средние серий значительно отличаются, то гипотеза H_0 отвергается.

Обозначим через $\overline{X_i}$ среднее арифметическое i-й серии наблюдений:

$$\overline{X_i} = \frac{1}{n} \sum_{j=1}^{n_i} x_{ij}, \quad i = 1, 2, \ldots, k,$$

а через $\overline{X}$ общее среднее арифметическое всех kn_k наблюдений:

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{k} \overline{X_i} n_i = \frac{1}{n} \sum_{i=1}^{k} \left(\frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} \right) n_i = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij}.$$

Тогда дисперсия D отдельных наблюдений x_{ij} от общего среднего $\overline{X}$ вычисляется по формуле

$$D = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \overline{X})^2 .$$

Общая сумма квадратов отклонений состоит из суммы квадратов отклонений между сериями (рассеивание по факторам)

$$D_2 = \sum_{i=1}^{k} (\overline{X}_i - \overline{X})^2 n_i$$

и суммы квадратов отклонений внутри серий (остаточное рассеивание)

$$D_1 = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \overline{X}_i)^2 .$$

В качестве статистической характеристики для гипотезы H_0 возьмем величину $F_{\text{набл}}$, которая имеет F-распределение Фишера с $\nu_1 = k - 1$ и $\nu_2 = n - 1$ степенями свободы. Критические границы F_{α,ν_1,ν_2} находим из прил. 6, а

$$F_{\text{набл}} = \frac{\frac{1}{k-1} D_2}{\frac{1}{n-k} D_1} = \frac{\frac{1}{k-1} \sum_{i=1}^{k} (\overline{X}_i - \overline{X})^2 n_i}{\frac{1}{n-k} \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \overline{X}_i)^2} .$$

Если $F_{\text{набл}} < F_{\alpha,\nu_1,\nu_2}$, то рассеивание по факторам незначительное и гипотеза H_0 принимается, если же $F_{\text{набл}} > F_{\alpha,\nu_1,\nu_2}$, рассеивание значимое и гипотеза H_0 отвергается.

Схему применения дисперсионного анализа можно представить в виде табл. 8.9.

Таблица 8.9

Дисперсия по факторам	Сумма квадратов	Число степеней свободы	Средняя сумма квадратов отклонений
Между сериями	$\sum_{i=1}^{k} (\overline{X}_i - \overline{X})^2 n_i$	$k-1$	$\frac{1}{k-1} \sum_{i=1}^{k} (\overline{X}_i - \overline{X})^2 n_i$
Внутри серии	$\sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \overline{X}_i)^2$	$n-k$	$\frac{1}{n-k} \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \overline{X}_i)^2$
Общая	$\sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \overline{X})^2$	$n-1$	$\frac{1}{n-1} \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \overline{X})^2$

Таким образом, для выяснения влияния факторов сравнивают дисперсию по факторам с остаточной дисперсией по величине отклонения. В этом сравнении и заключается основная идея дисперсионного анализа.

Пример 8.4. При анализе эксплуатационных качеств автомобилей "Жигули" выделяют один фактор – расход топлива. Исследуются три модели автомобилей "Жигули" 2102, 2106, 2108, по четыре в каждой, с расходом одного и того же вида топлива. Расход топлива приведен в табл. 8.10.

Таблица 8.10

Номер модели	Расход топлива, л,на 100 км			
	1	2	3	4
2102	5,1	5	5,9	5,3
2106	7	7,1	8,6	7
2108	9,1	9,8	8,9	10

Можно ли считать расход топлива в среднем одинаковым для всех автомобилей?

Р е ш е н и е. Значения средних по выделенным трем моделям будут:

$$\overline{X_1} = 5{,}325; \quad \overline{X_2} = 7{,}425; \quad \overline{X_3} = 9{,}450.$$

Промежуточные расчеты сведем в табл. 8.11.

Таблица 8.11

Номер модели	Расход топлива, л на 100 км										
	x_{i1}	x_{i1}^2	x_{i2}	x_{i2}^2	x_{i3}	x_{i3}^2	x_{i4}	x_{i4}^2	$\sum\limits_j x_{ij}$	$(\sum\limits_j x_{ij})^2$	$\sum\limits_j x_{ij}^2$
2102	5,1	26,01	5	25	5,9	34,81	5,3	28,09	21,3	453,69	113,91
2106	7	49	7,1	50,41	8,6	73,96	7	49	29,7	882,09	222,37
2108	9,1	82,81	9,8	96,04	8,9	79,21	10	100	37,8	1428,84	358,06
Σ	21,2	157,82	21,9	171,45	23,4	187,98	22,3	177,09	88,8	2764,62	694,34

Тогда общая дисперсия

$$D(X) = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \overline{X})^2 = \sum_{ij} x_{ij}^2 - \frac{1}{kn_i} (\sum_{ij} x_{ij})^2 = 694{,}34 - \frac{1}{3\cdot 4}(88{,}8)^2 =$$

$$= 37{,}22.$$

Дисперсия внутри групп

$$D_1(X) = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \overline{X_i})^2 = \sum_{ij} x_{ij}^2 - \frac{1}{n_i} \sum_{i=1}^{k} (\sum_{j=1}^{n_i} x_{ij})^2 = 694{,}34 -$$

$$- \frac{1}{4}\, 2764{,}62 = 3{,}185.$$

Дисперсия между группами

$$D_2(X) = \sum_{i=1}^{k} (\overline{X}_i - \overline{X})^2 n_i = \frac{1}{n_i} \sum_{i=1}^{k} (\sum_{j=1}^{n_i} x_{ij})^2 - \frac{1}{kn_i} (\sum_{ij} x_{ij})^2 =$$

$$= \frac{1}{4}\, 2764{,}62 - \frac{1}{3 \cdot 4} (88{,}8)^2 = 34{,}43\,.$$

Таблица 8.9 однофакторного дисперсионного анализа имеет следующий вид (табл.8.12)

Таблица 8.12

Дисперсия	Сумма квадратов	Число степеней свободы	Средняя сумма квадратов отклонений
Между группами	34,43	2	17,22
Внутри групп	3,185	9	0,35
Общая	37,22	11	3,38

Статистическая характеристика

$$F_{\text{набл}} = \frac{3{,}38}{0{,}35} = 9{,}66\,.$$

По таблице F-распределения с учетом двух степеней свободы большей дисперсии и при девяти степенях свободы меньшей дисперсии (прил. 6) находим критические границы F_α, равные 4,256 для $\alpha = 0{,}05$. Расчетное значение по наблюдениям превышает F_α, что дает повод отвергнуть гипотезу H_0 о равенстве в среднем расхода топлива в трех рассматриваемых моделях автомобилей.

ЗАДАЧИ

8.8. Предполагая, что долговечность электрической лампочки зависит от материала и технологии изготовления, исследуют 3 партии лампочек по 5 штук в каждой. Исходные данные приведены в табл. 8.13.

Таблица 8.13

Номер партии	Продолжительность горения, ч				
	1	2	3	4	5
1	1250	1150	1200	1210	1230
2	1460	1400	1350	1380	1470
3	1550	1500	1510	1490	1580

Проверить гипотезу H_0 о равенстве средних долговечностей лампочек в рассматриваемых 3 партиях.

8.9. В табл. 8.14 приводятся результаты пятикратного измерения чистоты поверхности в одном и том же месте образца с помощью трех микроскопов.

Таблица 8.14

Номер микроскопа	Отклонение чистоты поверхности, мм				
	1	2	3	4	5
1	−5	−2	−20	−2	−5
2	8	11	31	29	27
3	19	2	−12	9	3

Проверить нулевую гипотезу H_0 о равенстве систематических ошибок указанных приборов.

9. ОСНОВЫ ТЕОРИИ КОРРЕЛЯЦИИ

9.1. ПОНЯТИЕ О КОРРЕЛЯЦИИ И РЕГРЕССИИ. КОРРЕЛЯЦИОННАЯ ТАБЛИЦА. КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ

Две СВ могут быть связаны функциональной или статистической зависимостью либо быть независимыми.

Статистической называется зависимость, при которой изменение одной СВ ведет к изменению условного распределения другой СВ.

Условным средним $\overline{Y}_x$ называется среднее арифметическое значение СВ Y при $X = x$. Если каждому значению x соответствует одно значение условного среднего, то зависимость условного среднего от x является функциональной; в этом случае говорят о *корреляционной зависимости* СВ Y от СВ X. Уравнение $\overline{Y}_x = f(x)$ называется *уравнением регрессии Y на X*. Функция $f(x)$ называется *регрессией Y на X*, а ее график – *линией регрессии Y на X*.

При большом числе наблюдений одни и те же значения случайных величин X и Y могут встречаться много раз. Для удобства их группируют и записывают в виде *корреляционной таблицы* (табл. 9.1).

В качестве количественной оценки тесноты корреляционной связи между двумя СВ используют *коэффициент корреляции*. При положительном (отрицательном) значении коэффициента корреляции с возрастанием одной СВ в среднем возрастает (убывает) и другая. Близость коэффициента корреляции к единице свидетельствует о зависимости между СВ, близкой к функциональной линейной.

Выборочный коэффициент корреляции определяется по формуле

$$r_в = \frac{\sum\limits_{i\,j} n_{xy}\, x_i y_j - n\bar{x}\bar{y}}{n\,\bar{\sigma}_x\,\bar{\sigma}_y}, \qquad -1 \leqslant r_в \leqslant 1;$$

где n – объем выборки; $\bar{x}$, $\bar{y}$ – выборочные средние; $\bar{\sigma}_x$, $\bar{\sigma}_y$ – выборочные средние квадратичные отклонения.

Таблица 9.1

$y_j \backslash x_i$	x_1	x_2	...	x_k
y_1	n_{11}	n_{12}	...	n_{1k}
y_2	n_{21}	n_{22}	...	n_{2k}
...	...	...	...	...
y_m	n_{m1}	n_{m2}	...	n_{mk}

Пример 9.1. В результате измерений отклонений от номиналов высот моделей (x_i) и отливок к ним (y_j) получены следующие результаты:

x_i	0,90	1,22	1,32	0,77	1,30	1,20	1,32	0,95	0,45	1,30	1,20
y_j	−0,30	0,10	0,70	−0,28	0,25	0,02	0,37	−0,70	0,55	0,35	0,32

Составить корреляционную таблицу и вычислить коэффициент корреляции.

Р е ш е н и е. Разобьем весь интервал, в котором заключены значения признаков, на пять частей. Возьмем для x_i наименьшее значение 0,40 и наибольшее – 1,40, тогда ширина одного интервала будет равна 0,20. Наименьшее $y_j = -0{,}7$, а наибольшее – 0,7. Ширина интервала 0,28. Откладываем интервалы изменений x_i по горизонтали, а y_j – по вертикали; данные заносим в табл. 9.2.

Определим коэффициент корреляции. Для этого найдем средние значения $\overline{x}$ и $\overline{y}$, предполагая, что x_i и y_j – середины соответствующих интервалов:

Таблица 9.2

y_j \ x_i	0,40–0,60	0,60–0,80	0,80–1,00	1,00–1,20	1,20–1,40	n_y
−0,7 – −0,42	–	–	1	–	–	1
−0,42 – −0,14	–	1	1	–	1	3
−0,14 – 0,14	–	–	–	–	1	1
0,14 – 0,42	–	–	–	2	2	4
0,42 – 0,7	1	–	–	–	1	2
n_x	1	1	2	2	5	$n=11$

$$\overline{x} = \frac{0{,}50\cdot 1 + 0{,}70\cdot 1 + 0{,}90\cdot 2 + 1{,}10\cdot 2 + 1{,}30\cdot 2}{11} = \frac{11{,}70}{11} = 1{,}06,$$

$$\overline{y} = \frac{-0{,}56\cdot 1 - 0{,}28\cdot 3 + 0\cdot 1 + 0{,}28\cdot 4 + 0{,}56\cdot 2}{11} = \frac{0{,}84}{11} = 0{,}08,$$

$$D_x = \frac{1}{11}\left((1{,}30 - 1{,}06)^2\cdot 5 + (0{,}50 - 1{,}06)^2\cdot 1 + (0{,}70 - 1{,}06)^2\cdot 1 +\right.$$

$$\left.+ (0{,}90 - 1{,}06)^2\cdot 2 + (1{,}10 - 1{,}06)^2\cdot 2\right) = 0{,}7816/11 = 0{,}078,$$

$$D_y = \frac{1}{11}\left((-0{,}64)^2\cdot 1 + (-0{,}36)^2\cdot 3 + (-0{,}08)^2\cdot 1 + (-0{,}08)^2\cdot 1 +\right.$$

$$\left.+ (0{,}20)^2\cdot 4 + (0{,}48)^2\cdot 2\right) = 2{,}1168/11 = 0{,}211,$$

$$r_в = \frac{1}{11\cdot 0{,}26\cdot 0{,}44}(0{,}58\cdot 0{,}50\cdot 1 - 0{,}28\cdot 0{,}70\cdot 1 - 0{,}56\cdot 0{,}90\cdot 1 -$$

$$- 0{,}28\cdot 0{,}90\cdot 1 + 0{,}28\cdot 1{,}1\cdot 2 + 0{,}28\cdot 1{,}3\cdot 1 + 0{,}28\cdot 1{,}3\cdot 2 + 0{,}56\cdot 1{,}3\cdot 5) =$$

$$= 1{,}046/1{,}258 = 0{,}82.$$

Коэффициент корреляции близок к единице, следовательно, между случайными величинами X и Y достаточно тесная корреляционная связь.

ЗАДАЧИ

9.1. Результаты лабораторных испытаний прочности канатов диаметром 13 мм представлены в табл. 9.3, где x_i – сумма разрывных усилий (Н) отдельных проволок, из которых свит канат; y_j – разрывное усилие всего каната (Н).

Таблица 9.3

x_i	y_j	x_i	y_j	x_i	y_j
14360	11350	14140	11540	14350	11450
13860	11400	14240	11450	14210	11400
13650	11400	14250	11465	14350	11500
14400	11560	14630	11640	14020	11400
13900	11500	14080	11450		

Составить корреляционную таблицу и вычислить коэффициент корреляции. (*Указание.* Весь интервал, в котором заключены значения признаков, разбить на 6 равных частей.)

9.2. На металлургическом заводе исследовалась зависимость предела прочности (Н/мм^2) от предела текучести (Н/мм^2). Результаты замеров прочности (x_i) и текучести (y_j) стали 50 марок приведены в табл. 9.4.

Таблица 9.4

x_i	y_j	x_i	y_j	x_i	y_j	x_i	y_j	x_i	y_j
77	81	81	54	129	100	104	94	96	84
96	77	57	40	145	95	108	84	112	94
86	76	86	61	142	206	93	73	136	162
92	86	80	68	120	118	124	107	104	98
98	53	87	88	95	109	112	94	103	77
53	47	163	145	107	107	113	107	115	88
63	36	153	136	133	120	95	99	123	94
80	40	133	129	140	114	112	100	111	76
64	49	159	126	149	113	116	104	127	84
66	60	134	96	147	123	93	88	129	73

Составить корреляционную таблицу, разбив интервалы изменения X и Y на 6 частей, и вычислить коэффициент корреляции.

9.2. ЛИНЕЙНАЯ КОРРЕЛЯЦИЯ

Уравнение прямой регрессии Y на X имеет вид

$$\overline{y}_x - \overline{y} = r_{\text{в}} \frac{\overline{\sigma}_y}{\overline{\sigma}_x} (x - \overline{x}),$$

где $\overline{y}_x$ – условное среднее; $\overline{x}, \overline{y}$ – выборочные средние признаков X и Y; $\overline{\sigma}_x$, $\overline{\sigma}_y$ – выборочные средние квадратичных отклонений признаков X и Y; $r_{\text{в}}$ – выборочный коэффициент корреляции.

Уравнение прямой регрессии X на Y имеет вид

$$\overline{x}_y - \overline{x} = r_в \frac{\overline{\sigma}_x}{\overline{\sigma}_y} (y - \overline{y}).$$

Величины $r_в \overline{\sigma}_y / \overline{\sigma}_x$ и $r_в \overline{\sigma}_x / \overline{\sigma}_y$ называются *линейными коэффициентами регрессии* и обозначаются соответственно $\rho_{y/x}$ и $\rho_{x/y}$:

$$r_в = \rho_{y/x} \overline{\sigma}_x / \overline{\sigma}_y \quad \text{или} \quad r_в = \rho_{x/y} \overline{\sigma}_y / \overline{\sigma}_x .$$

Если значения признаков X и Y заданы в виде корреляционной таблицы с равноотстоящими вариантами, то с целью упрощения расчетов полезно перейти к условным вариантам:

$$u_i = (x_i - c_1)/h_1 , \qquad v_i = (y_j - c_2)/h_2 ,$$

где c_1, c_2 – варианты признака (обычно за условные нули c_1 и c_2 принимают варианты с наибольшими частотами); h_1, h_2 – разности между соседними вариантами признаков X и Y.

В условных вариантах выборочный коэффициент корреляции

$$r_в = \frac{\Sigma n_{uv} uv - n\overline{u}\overline{v}}{n \sigma_u \sigma_v} ,$$

где $\overline{u}$, $\overline{v}$, σ_u и σ_v можно вычислить по формулам:

$$\sigma_u = \sqrt{\overline{u^2} - (\overline{u})^2} , \qquad \sigma_v = \sqrt{\overline{v^2} - (\overline{v})^2} ;$$

$$\overline{x} = \overline{u} h_1 + c_1 , \quad \overline{y} = \overline{v} h_2 + c_2 ;$$

$$\sigma_x = \sigma_u h_1 , \quad \sigma_y = \sigma_v h_2 .$$

Пример 9.2. Распределение 40 заводов области по количеству Y ремонтных слесарей и числу X станко-смен представлено следующей корреляционной таблицей (табл.9.5).

Таблица 9.5

X \ Y	10–15	15–20	20–25	25–30	30–35	35–40	n_x
0–0,2	4	–	–	–	–	–	4
0,2–0,4	2	2	–	–	–	–	4
0,4–0,6	–	–	2	–	–	–	2
0,6–0,8	–	6	–	4	4	–	14
0,8–1,0	–	–	–	–	6	6	12
1,0–1,2	–	–	–	–	–	4	4
n_y	6	8	2	4	10	10	$n = 40$

Составить уравнение прямой регрессии, установить тесноту связи между признаками. Для каждого интервала значений Y вычислить фактические значения частных средних y_x и теоретические значения, найденные из уравнения регрессии.

Р е ш е н и е. За значения признаков примем середины интервалов и составим корреляционную таблицу (табл. 9.6) в условных вариантах, приняв в качестве условных нулей $c_1 = 0{,}7$ и $c_2 = 27{,}5$. (Эти варианты имеют частоту, равную 4, и находятся в середине корреляционной таблицы.)

Таблица 9.6

u \ v	−3	−2	−1	0	1	2	n_u
−3	4	–	–	–	–	–	4
−2	2	2	–	–	–	–	4
−1	–	–	2	–	–	–	2
0	–	6	–	4	4	–	14
1	–	–	–	–	6	6	12
2	–	–	–	–	–	4	4
n_v	6	8	2	4	10	10	$n = 40$

Находим:

$$\overline{v} = \frac{-3\cdot 6 + (-2)\cdot 8 + (-1)\cdot 2 + 0\cdot 4 + 1\cdot 10 + 2\cdot 10}{40} = -\frac{6}{40} = -0{,}15,$$

$$\overline{u} = \frac{-3\cdot 4 - 2\cdot 4 - 1\cdot 2 + 0\cdot 14 + 1\cdot 12 + 2\cdot 4}{40} = -\frac{2}{40} = -0{,}05,$$

$$\overline{v^2} = \frac{9\cdot 6 + 4\cdot 8 + 1\cdot 2 + 0\cdot 4 + 1\cdot 0 + 4\cdot 10}{40} = \frac{138}{40} = 3{,}45,$$

$$\overline{u^2} = \frac{9\cdot 4 + 4\cdot 4 + 1\cdot 2 + 0\cdot 14 + 1\cdot 12 + 4\cdot 4}{40} = \frac{82}{40} = 2{,}05,$$

$$\sigma_в = \sqrt{3{,}45 - 0{,}0225} = 1{,}86, \quad \sigma_u = \sqrt{2{,}05 - 0{,}0025} = 1{,}43,$$

$$\Sigma\, n_{uv} uv = 4\cdot 9 + 2\cdot 6 + 2\cdot 4 + 2\cdot 1 + 6\cdot 1 + 6\cdot 2 + 4\cdot 4 = 92.$$

Найдем искомый коэффициент корреляции:

$$r_в = \frac{92 - 40\cdot 0{,}15\cdot 0{,}05}{40\cdot 1{,}86\cdot 1{,}43} = \frac{91{,}7}{106} = 0{,}85.$$

Вычислим $\overline{x}$, $\overline{y}$, $\overline{\sigma}_x$, $\overline{\sigma}_y$:

$$\overline{x} = \overline{u}h_1 + c_1 = -0{,}05\cdot 0{,}2 + 0{,}7 = 0{,}69,$$

$$\overline{y} = \overline{v}h_2 + c_2 = -0{,}15\cdot 5 + 27{,}5 = 26{,}75,$$

$$\overline{\sigma}_x = \sigma_u h_1 = 1{,}43\cdot 0{,}2 = 0{,}286, \quad \overline{\sigma}_y = \sigma_v h_2 = 1{,}86\cdot 5 = 9{,}3.$$

Подставим полученные значения в уравнение регрессии:

$$y_x - 26{,}75 = 0{,}85\cdot\frac{9{,}3}{0{,}286}(x - 0{,}69) \text{ или } y_x = 27{,}64x + 7{,}03 .$$

Вычислим для каждого интервала изменения x фактические значения частных средних:

$$y_{x=0,1} = 12{,}5 , \quad y_{x=0,3} = \frac{2\cdot 12{,}5 + 17{,}5}{4} = \frac{60}{4} = 15 ,$$

$$y_{x=0,5} = \frac{2\cdot 22{,}5}{2} = 22{,}5 ,$$

$$y_{x=0,7} = \frac{17{,}5\cdot 6 + 27{,}6\cdot 4 + 32{,}5\cdot 4}{14} = \frac{345}{14} = 24{,}64 ,$$

$$y_{x=0,9} = \frac{32{,}5\cdot 6 + 37{,}5\cdot 6}{12} = \frac{420}{12} = 35, \quad y_{x=1,1} = 37{,}5 .$$

Вычислим для каждого интервала изменения x теоретические значения из полученного уравнения:

$$y_{x=0,1} = 27{,}64\cdot 0{,}1 + 7{,}03 = 9{,}79 ,$$

$$y_{x=0,3} = 27{,}64\cdot 0{,}3 + 7{,}03 = 15{,}32 ,$$

$$y_{x=0,5} = 27{,}64\cdot 0{,}5 + 7{,}03 = 20{,}85 ,$$

$$y_{x=0,7} = 27{,}64\cdot 0{,}7 + 7{,}03 = 26{,}37,$$

$$y_{x=0,9} = 27{,}64\cdot 0{,}9 + 7{,}03 = 31{,}9 ,$$

$$y_{x=1,1} = 27{,}64\cdot 1{,}1 + 7{,}03 = 37{,}43.$$

Сравнивая полученные значения, видим, что они близки к фактическим.

ЗАДАЧИ

9.3. Найти уравнение прямой регрессии Y на X и X на Y по данным, приведенным в таблицах: а) 9.7; б) 9.8.

Таблица 9.7

y_j \ x_i	1	2	3	4	n_y
1	1	–	–	–	1
2	2	1	–	–	3
3	1	2	1	–	4
4	–	1	2	1	4
5	–	–	1	2	3
n_x	4	4	4	3	$n = 15$

Таблица 9.8

y_i \ x_i	65	95	125	155	185	215	n_y
30	5	–	–	–	–	–	5
40	4	12	–	–	–	–	16
50	–	8	5	4	–	–	17
50	–	1	5	7	2	–	15
70	–	–	–	–	1	1	2
n_x	9	21	10	11	3	1	$n = 55$

9.4. Результаты измерений колебаний крутящего момента на полуоси автомобиля (X) и угловых колебаний ведущего моста (Y) приведены в табл. 9.9. Найти уравнение прямой регрессии Y на X и оценить тесноту связи между признаками.

9.5. Результаты измерений угловых колебаний ведущего моста автомобиля (X) и угловых колебаний подрессоренной массы (галопирование) (Y) приведены в табл. 9.10. Найти уравнение прямой регрессии Y на X, установить тесноту связи между признаками. Для каждого интервала значений X_i вычислить фактические значения частных средних y_x и из уравнения регрессии – теоретические значения.

9.3. КРИВОЛИНЕЙНАЯ РЕГРЕССИЯ

Если в уравнении регрессии $f(x)$ является нелинейной функцией, то корреляцию называют *криволинейной.* Теснота криволинейной связи определяется *корреляционным отношением*

$$\eta_{yx} = \sigma_{\bar{y}x}/\sigma_y \quad \text{или} \quad \eta_{xy} = \sigma_{\bar{x}y}/\sigma_x ,$$

где σ_y – среднее квадратичное отклонение значений y_j от $\bar{y}$; σ_x – среднее квадратичное отклонение x_i от $\bar{x}$; $\sigma_{\bar{x}y}$ и $\sigma_{\bar{y}x}$ – средние квадратичные отклонения значений условных средних $\bar{y}_x$ и $\bar{x}_y$ от средних значений $\bar{y}$ и $\bar{x}$ соответственно:

$$\sigma^2_{\bar{y}x} = \frac{\Sigma n_x(\bar{y}_x - \bar{y})^2}{n}, \qquad \sigma^2_{\bar{x}y} = \frac{\Sigma n_y(\bar{x}_y - \bar{x})^2}{n};$$

$\bar{y}_x = \Sigma n_{xy} y / \Sigma n_x$; n – объем выборки.

Приведем *свойства корреляционного отношения.*

1. *Корреляционное отношение удовлетворяет неравенству* $0 \leqslant \eta \leqslant 1$.

2. *Если* $\eta = 0$, *то признак* Y *не связан с признаком* X *корреляционной зависимостью.*

3. *Если* $\eta = 1$, *то признак* Y *связан с признаком* X *функциональной зависимостью.*

4. *Корреляционное отношение не меньше абсолютной величины выборочного коэффициента корреляции, т.е.* $\eta \geqslant |r_в|$. *Если* $\eta = |r_в|$, *то имеет место точная линейная корреляционная зависимость.*

Простейшими случаями криволинейной корреляции являются:

1) параболическая корреляция второго порядка с уравнением регрессии $\bar{y}_x = ax^2 + bx + c$;

Таблица 9.9

y_j \ $x_i \cdot 10^{-3}$	−18,4–(−12,42)	−12,42–(−6,44)	−6,44–(−0,46)	−0,46–5,52	5,52–11,5	11,5–17,8	17,8–23,46	23,46–29,44	29,44–35,42	35,42–41,2	n_y
(−232)–(−171,6)	–	4	–	–	–	–	–	–	–	–	4
(−171,6)–(−111,2)	1	3	1	2	–	–	–	–	–	–	7
(−111,2)–(−50,8)	2	1	–	2	–	–	–	–	–	–	5
(−50,8)–9,6	–	4	1	5	2	–	–	–	–	–	12
9,6–70	–	2	1	2	1	–	1	–	–	–	7
70–130,4	–	–	–	2	1	1	–	–	1	–	5
130,4–190,8	–	1	–	1	3	5	8	1	–	–	19
190,8–251,2	–	–	–	–	3	5	8	3	1	1	21
251,2–311,6	–	1	–	1	3	5	4	1	2	1	18
311,6–372	–	–	–	–	–	–	2	–	–	–	2
n_x	3	16	3	15	13	16	23	5	4	2	$n =$ 100

Таблица 9.10

$y_j \cdot 10^{-3}$ \ $x_i \cdot 10^{-3}$	−18,4–(−12,42)	−12,42–(−6,44)	−6,44–(−0,46)	−0,46–(−0,46)	5,52–11,5	11,5–17,48	17,48–23,46	23,46–29,46	29,44–35,42	35,42–41,44	n_y
−17,5–(−11,9)	–	–	–	–	–	–	2	–	–	–	2
−11,9–(−6,3)	–	–	–	–	–	–	1	1	–	–	2
−6,3–(−0,7)	–	–	–	–	1	3	4	–	–	–	8
−0,7–4,9	–	–	–	1	1	2	5	–	3	–	12
4,9–10,5	–	–	–	4	2	3	7	–	–	1	17
10,5–16,1	1	3	1	3	2	3	4	2	–	–	19
16,1–21,7	1	6	1	2	5	3	2	–	–	–	20
21,7–27,3	–	2	–	2	3	–	–	–	1	–	8
27,3–32,9	–	2	1	3	1	1	–	–	–	1	9
32,9–38,5	1	1	–	–	–	1	–	–	–	–	3
n_x	3	14	3	15	15	16	25	3	4	2	$n =$ 100

2) параболическая корреляция третьего порядка с уравнением регрессии

$$\overline{y}_x = ax^3 + bx^2 + cx + d;$$

3) гиперболическая корреляция с уравнением регрессии

$$\overline{y}_x = a/x + b .$$

Пример 9.3. Найти уравнение параболической регрессии Y на X для экспериментальных данных, помещенных в табл. 9.11.

Таблица 9.11

y_j \ x_i	1	2	3	4	5	6	n_y
1	2	1	–	–	–	–	3
2	1	2	–	–	–	–	3
3	–	3	1	–	–	–	4
4	–	1	3	1	–	–	5
5	–	–	2	2	2	1	7
6	–	–	–	1	1	1	3
n_x	3	7	6	4	3	2	$n = 25$
$\overline{y}_x$	1,33	2,57	4,17	5,0	5,33	5,50	

Решение. Ищем уравнение регрессии в виде

$$\overline{y}_x = ax^2 + bx + c .$$

Для определения неизвестных коэффициентов a, b по МНК записываем систему нормальных уравнений:

$$\left.\begin{aligned} &a\,\Sigma n_x x^2 + b\,\Sigma n_x x + nc = \Sigma\, n_x \overline{y}_x\,,\\ &a\,\Sigma n_x x^3 + b\,\Sigma\, n_x x^2 + c\,\Sigma\, n_x x = \Sigma n_x x \overline{y}_x\,,\\ &a\,\Sigma\, n_x x^4 + b\,\Sigma n_x x^3 + c\,\Sigma n_x x^2 = \Sigma\, n_x x^2 \overline{y}_x\,,\end{aligned}\right\} \qquad (9.1)$$

и составляем вспомогательную таблицу (табл. 9.12).

Таблица 9.12

n_x	x	$n_x x$	$n_x x^2$	$n_x x^3$	$n_x x^4$	$\overline{y}_x$	$n_x \overline{y}_x$	$n_x x \overline{y}_x$	$n_x x^2 \overline{y}_x$
3	1	3	3	3	3	1,33	3,99	3,99	3,99
7	2	14	28	56	112	2,57	17,99	35,98	71,96
6	3	18	54	162	486	4,17	25,02	75,06	225,18
4	4	16	64	256	1024	5,0	20,00	80,00	320,0
3	5	15	75	375	1875	5,33	15,99	79,95	399,75
2	6	12	72	432	2592	5,50	11,00	66,00	396,00
$\Sigma n_x = 25$	21	78	296	1284	6092	23,90	93,99	340,98	1416,88

Теперь уравнения (9.1) примут вид:

$$\left.\begin{aligned} 296a + 78b + 25c &= 93{,}99\,, \\ 1284a + 296b + 78c &= 340{,}98\,, \\ 6092a + 1284b + 296c &= 1416{,}88\,. \end{aligned}\right\}$$

Для упрощения расчетов разделим каждое уравнение на коэффициент при c:

$$\left.\begin{aligned} 11{,}84a + 3{,}12b + c &= 3{,}76\,, \\ 16{,}46a + 3{,}80b + c &= 4{,}38\,, \\ 20{,}58a + 4{,}34b + c &= 4{,}79. \end{aligned}\right\}$$

Решив полученную систему, найдем: $a = -0{,}19$, $b = 2{,}21$, $c = 0{,}89$.
Уравнение регрессии имеет вид

$$y_x = -0{,}19x^2 + 2{,}21x - 0{,}89.$$

Подставив в это уравнение вместо x его значения, получим теоретические значения средних $\overline{y}_x$:

x	1	2	3	4	5	6
$\overline{y}_x$	1,14	2,78	4,07	4,91	5,41	5,52

Сравнивая теоретические значения частных средних $\overline{y}_x$ с экспериментальными, видим, что они достаточно близки.

ЗАДАЧИ

9.6. В результате исследований установлено, что между овальностью колец после их обточки (X) и термической обработки (Y) существует корреляционная связь, представленная табл. 9.13.

Таблица 9.13

y_j \ x_i	5	10	15	20	n_y
10	2	–	–	–	2
20	5	4	1	–	10
30	3	8	6	3	20
40	–	3	6	6	15
50	–	–	2	1	3
n_x	10	15	15	10	$n = 50$

Найти уравнение кривой регрессии Y на X.

9.7. Приведены данные об износе резца, определяемом его толщиной y (мм) в зависимости от времени работы t (ч):

y	t	y	t	y	t	y	t	y	t	y	t
0	30,0	3	28,1	6	27,5	9	26,8	12	26,1	15	24,8
1	29,1	4	28,0	7	27,2	10	26,5	13	25,7	16	24,0
2	28,4	5	27,7	8	27,0	11	26,3	14	25,3		

Найти уравнение параболической регрессии y на t.

9.8. Даны результаты измерений признаков Y и X:

x	1,2	1,8	3,1	4,9	5,7	7,1	8,6	9,8
y	4,5	5,9	7,0	7,8	7,2	6,8	4,5	2,7

Построить эмпирическую кривую. По виду этой кривой определить вид уравнения регрессии, записать это уравнение и установить по нему тесноту связи. Найти теоретические значения $\overline{y}_x$.

9.9. В результате лабораторных испытаний на прочность стальных проволок различных диаметров (X), предназначенных для свивания канатов, были получены следующие данные:

x	0,6	2	2,2	2,45	2,6
y	50	560	690	760	900

Считая, что между диаметром проволоки X (мм) и разрывным усилием Y (Н) существует корреляционная зависимость параболического типа, найти уравнение регрессии.

9.10. Данные о количестве выпускаемых деталей X (тысячи штук) и полных затратах на их изготовление Y (сотни рублей), полученные на 15 машиностроительных заводах, записаны в табл. 9.14.

Таблица 9.14

y_j \ x_i	3	4	5	7	8	10	12	13	14	19	20	24	26	n_x
2	–	–	–	–	–	–	–	–	–	1	1	1	1	4
4	1	1	–	–	1	–	–	1	1	–	–	–	–	5
9	1	–	–	1	–	1	1	–	–	–	–	–	–	4
18	–	–	1	–	1	–	–	–	–	–	–	–	–	2
n_y	2	1	1	1	2	1	1	1	1	1	1	1	1	n=15

Для каждого значения x найти частное среднее $\overline{y}_x$. Построить в системе координат $(x, \overline{y}_x)$ эмпирическую кривую и по ней установить возможность представления рассматриваемой зависимости в виде уравнения регрессии $y_x = a/x + b$; записать это уравнение. Найти корреляционное отношение $n_{y/x}$.

9.4. ОЦЕНКА КОРРЕЛЯЦИОННЫХ ХАРАКТЕРИСТИК. ПРОВЕРКА ГИПОТЕЗЫ ОБ ОТСУТСТВИИ КОРРЕЛЯЦИОННОЙ СВЯЗИ

Оценками корреляционных характеристик генеральной совокупности являются выборочные характеристики. Доверительные интервалы для оцениваемых параметров определяются по методике, изложенной в гл. 7.

Рассмотрим проверку гипотезы о значимости выборочных характеристик. Предположим, что генеральная совокупность (X, Y) распределена нормально. Из нее составлена выборка объемом n и найден выборочный коэффициент корреляции $r_в \neq 0$. Но это еще не означает, что коэффициент корреляции генеральной совокупности будет отличаться от нуля. Появляется необходимость проверить нулевую гипотезу H_0 о равенстве нулю коэффициента корреляции генеральной совокупности $(H_0 : r_г = 0)$ при конкурирующей гипотезе $(H_1 : r_г \neq 0)$. Если нулевая гипотеза отвергается, то это означает, что выборочный коэффициент значимо отличается от нуля, а X и Y коррелированы, т.е. связаны линейной зависимостью. Если нулевая гипотеза будет принята, то выборочный коэффициент корреляции незначим, а X и Y некоррелированны, т.е. не связаны линейной зависимостью.

В качестве критерия при проверке нулевой гипотезы используется СВ

$$T = r_в \sqrt{n-2}/\sqrt{1-r_в^2},$$

которая подчиняется распределению Стьюдента с $k = n - 2$ степенями свободы (см. прил. 3).

Для проверки нулевой гипотезы H_0 о равенстве нулю коэффициента корреляции генеральной совокупности нужно вычислить наблюдаемое значение критерия:

$$T_{набл} = r_в \sqrt{n-2}/\sqrt{1-r_в^2},$$

а затем по заданному уровню значимости α, числу степеней свободы и таблице критических точек распределения Стьюдента найти критическую точку $t_{кр}$ (α, k). Если $|T_{набл}| < t_{кр}$, оснований отвергнуть нулевую точку нет. Если $|T_{набл}| > t_{кр}$, нулевую гипотезу отвергают.

Пример 9.4. Из двухмерной нормальной генеральной совокупности извлечена выборка объемом $n = 122$. Найден выборочный коэффициент корреляции $r_в = 0{,}4$. Проверить нулевую гипотезу H_0 о равенстве нулю генерального коэффициента корреляции при уровне значимости $\alpha = 0{,}05$ и конкурирующей гипотезе H_1.

Р е ш е н и е. Находим

$$T_{набл} = r_в \sqrt{n-2}/\sqrt{1-r_в^2} = 0{,}4\sqrt{120}/\sqrt{1-0{,}16} = 4{,}79.$$

По условию конкурирующая гипотеза $H_1 : r_1 \neq 0$, поэтому критическая область – двусторонняя. По уровню значимости $\alpha = 0{,}05$ и числу степеней свободы $k = 122 - 2 = 120$ находим из таблицы значений распределения Стьюдента для двусторонней критической области $t_{кр} = (0{,}05, 120) = 1{,}98$.

Так как $T_{набл} > t_{кр}$, т.е. $4{,}79 > 1{,}98$, нулевую гипотезу отвергаем, т.е. выборочный коэффициент значимо отличается от нуля, следовательно, X и Y коррелируемы.

ЗАДАЧИ

9.11. Используя условия примера 9.1, проверить нулевую гипотезу о равенстве нулю коэффициента корреляции генеральной совокупности при уровне значимости $\alpha = 0{,}10$ и конкурирующей гипотезе $H_1 : r_1 \neq 0$.

9.12. Используя условия задач 9.1 и 9.2 и считая, что СВ Y подчиняется нормальному закону распределения, проверить нулевую гипотезу о равенстве нулю коэффициента корреляции генеральной совокупности при уровне значимости $\alpha = 0{,}02$ и конкурирующей гипотезе . $H_1 : r_1 \neq 0$.

10. ЭЛЕМЕНТЫ ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА

10.1. ЗАДАЧА ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА. МОДЕЛЬ ОБЪЕКТА. ВЫБОР ФАКТОРОВ

Целью любого эксперимента является получение информации об исследуемом объекте. Опытные данные могут накапливаться либо путем пассивного наблюдения, либо с помощью активного эксперимента. В последнем случае осуществляется искусственное воздействие на объект по заранее спланированной программе, что позволяет вскрыть закономерности, найти оптимальные режимы функционирования этого объекта.

В общем виде объект исследования характеризуется следующими величинами: x_i , $i = 1, 2, ..., n$,– управляющие параметры; y_j , $j = 1, 2, ..., k$,– параметры состояния (выходные); z_m , $m = 1, 2, ..., p$,– возмущающие воздействия.

Управляющие параметры x_i представляют собой независимые переменные, определяющие выходные параметры. К *параметрам состояния* y_j относится совокупность контролируемых или вычисляемых параметров, характеризующих состояние объекта. *Возмущающие воздействия* z_m оказывают влияние на объект, но не могут быть измерены и поэтому проявляют себя как случайные величины или случайные функции времени.

Следовательно, одной из основных задач эксперимента является выявление взаимосвязей между входными и выходными параметрами объекта и представление их в виде математической модели

$$y_j = f_j(x_1, x_2, ..., x_n).$$

Входные параметры x_i называются *факторами*. Каждый фактор имеет область определения, которая устанавливается до эксперимента и может быть как непрерывной, так и дискретной. Комбинация факторов рассматривается как *факторное пространство*. Его называют также *функцией цели, критерием эффективности, критерием оптимальности, параметром оптимизации* и др. Все приведенные термины являются синонимами. Область возможных комбинаций факторов – *план эксперимента*.

При планировании активного эксперимента необходимо включить в рассмотрение все существенные факторы, которые могут влиять на процесс.

Факторы должны удовлетворять требованиям: управляемости, однозначности (трудно управлять фактором, который в свою очередь является функцией других факторов) и независимости (возможности установления фактора на любом уровне вне зависимости от уровней других факторов).

Для того чтобы спланировать эксперимент, необходимо располагать некоторым математическим описанием (математической моделью) исследуемой системы. Построение математической модели исследуемого объекта возможно либо на основании пассивного эксперимента, либо на выдвижении логически обоснованных гипотез. Функция цели представляется в виде полинома (линии регрессии), коэффициенты которого (коэффициент регрессии) определяются по данным эксперимента:

$$y = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{i<j} a_{ij} x_i x_j + \sum_{i=1}^{n} a_{ii} x_i^2 + \dots,$$

где a_0, a_{ij}, a_{ii} – выборочные коэффициенты регрессии.

Для того чтобы правильно выбрать степень полинома, необходимо учитывать два основных требования, предъявляемые к математической модели, – адекватности и простоты. Модели низких степеней являются более простыми, а условия адекватности проверяются после реализации опытов. Проверка адекватности – строго формализованная статистическая процедура, выполняемая с помощью F-критерия Фишера

$$F = S^2_{\text{адк}} / S^2_y \,, \tag{10.1}$$

где $S^2_{\text{адк}}$ – дисперсия адекватности; S^2_y – дисперсия, характеризующая ошибку опыта (см. прил. 4).

Модель считается адекватной в том случае, когда значение F, рассчитанное по формуле (10.1), не превышает соответствующего табличного значения при заданном уровне значимости. Обычно при решении инженерных задач уровень значимости принимается равным 0,05 (5 %).

10.2. ПОЛНЫЙ ФАКТОРНЫЙ ЭКСПЕРИМЕНТ

Полным факторным экспериментом (ПФЭ) называется такой эксперимент, при реализации которого определяется значение функции цели для всех возможных сочетаний уровней варьирования факторов. Если имеется k факторов, каждый из которых может устанавливаться на q уровнях, то для осуществления полного факторного эксперимента необходимо поставить $n = q^k$ опытов.

Наибольшее распространение получили эксперименты, в которых факторы варьируются на двух уровнях, т.е. эксперименты типа 2^k.

Планирование, проведение и обработка результатов ПФЭ состоят из следующих обязательных этапов: кодирование факторов; составление план-матрицы эксперимента; реализация плана эксперимента; проверка воспроизводимости опытов; проверка адекватности линейной модели; оценка значимости коэффициентов регрессии.

Кодирование факторов необходимо для перевода натуральных факторов в безразмерные величины, чтобы иметь возможность построить стандартную ортогональную план-матрицу эксперимента.

После выбора факторов для каждого из них устанавливается основной уровень (т.е. какое-то исходное значение) и интервал варьирования. Прибавление интервала варьирования к основному уровню дает верхний, а вычитание – нижний уровень фактора. Удобно, чтобы верхний уровень соответствовал +1, нижний –1, а основной 0. Для этого факторы кодируют так, чтобы их подобные значения x_i были связаны с натуральными $x_{i\text{н}}$ соотношением

$$x_i = (x_{i\text{н}} - x_{i0}) / \Delta x_i \,,$$

где x_{i0} – нулевой уровень фактора; Δx_i – интервал варьирования i-го фактора. На интервал варьирования накладываются определенные ограничения. Рекомендуется выбирать интервал, не превышающий удвоенной средней квадратичной ошибки в определении данного фактора.

Для перевода натуральных переменных в кодовые x_i заполняют таблицу кодирования переменных на двух уровнях. Затем составляют план-матрицу эксперимента и приведенный план эксперимента. Проиллюстрируем это на простом примере.

Пример 10.1. Исследовать процесс нагрева агрегата в зависимости от времени непрерывной работы. Цель исследования – определить зависимость износа деталей от скорости и конечной температуры нагрева. Температура изменяется от 250 до 450 °С, скорость нагрева – от 2 до 10 К/мин.

Р е ш е н и е. Кодируем факторы, сводя результаты в табл. 10.1 ($i = 1, 2$).

Таблица 10.1

Интервал варьирования и уровень факторов	Температура x_1, °С	Скорость нагрева x_2, К/мин
Нулевой уровень $x_{i0} = 0$	350	6
Интервал варьирования Δx_i	50	2
Нижний уровень $x_i = -1$	300	4
Верхний уровень $x_i = +1$	400	8

Составление плана-матрицы эксперимента (табл. 10.2) осуществляется следующим образом. В примере варьируются только два фактора x_1 и x_2, причем каждый – на двух уровнях +1, –1 (или +, –).

Таблица 10.2

Номер опыта	x_1	x_2
1	–1	–1
2	+1	–1
3	–1	+1
4	+1	+1

Строки в таблице соответствуют различным опытам, а столбцы – значениям факторов. В первом опыте оба фактора находятся на нижнем уровне, во втором опыте фактор x_1 – на верхнем, а фактор x_2 – на нижнем уровнях и т.д. Такие таблицы называют *матрицами планирования эксперимента.*

Реализация плана эксперимента представлена в табл. 10.3. *Приведенный план экспе-*

Таблица 10.3

Номер опыта	x_1	x_2	x_1x_2	y_{n_1}	y_{n_2}	$y_n = (y_{n_1} + y_{n_2})/2$
1	–1	–1	+1	27,0	28,0	27,5
2	+1	–1	–1	15,9	17,1	16,5
3	–1	+1	–1	22,1	22,9	22,5
4	+1	+1	+1	13,4	13,6	13,5

римента представляет собой расширенную матрицу, так как введен столбец x_1x_2, позволяющий оценить коэффициент регрессии при взаимодействии факторов. Здесь y_{n_1}, y_{n_2} – износ деталей.

Проверка воспроизводимости параллельных опытов при одинаковом их числе на каждом сочетании уровней факторов осуществляется по *критерию Кохрена* σ (см. прил. 11):

$$\sigma = \frac{S^2_{n\,\text{наиб}}}{\sum\limits_{u=1}^{n} S_u^2} \leqslant \sigma\,(0{,}05;\ f_n;\ f_u), \tag{10.2}$$

где $S_u^2 = \dfrac{\sum\limits_{p=1}^{m} (y_{up} - \bar{y}_u)^2}{m-1}$ – дисперсия, характеризующая рассеяние результатов опытов на u-м сочетании уровней факторов; $p = 1, 2, ..., m$ – число параллельных опытов; $S^2_{n\,\text{наиб}}$ – наибольшая из дисперсий в строчках плана; $\sigma\,(0{,}05;\ f_n;\ f_u)$ – табличное значение критерия Кохрена при 5 %-м уровне значимости; f_n – число независимых оценок дисперсии; $f_u = m - 1$ – число степеней свободы каждой оценки.

Процесс считается *воспроизводимым,* если выполняется неравенство (10.2). При этом дисперсия воспроизводимости (ошибка опыта) определяется по формуле

$$S_y^2 = \sum_{u=1}^{n} S_u^2/n\ . \tag{10.3}$$

Если неравенство (10.2) не выполняется, то необходимо уточнить измерения в опыте с максимальной дисперсией.

В случае воспроизводимости процесса рассчитывают коэффициенты регрессии:

$$a_0 = \sum_{u=1}^{n} \bar{y}_u/n\,,\ a_i = \sum_{u=1}^{n} x_{iu}\bar{y}_u/n\ , \tag{10.4}$$

$$a_{ij} = \sum_{u=1}^{n} x_{iu}x_{ju}\bar{y}_u/n$$

и строят линейную модель эксперимента. *Проверка адекватности* этой модели выполняется с помощью *критерия Фишера* (см. прил. 6). Адекватность имеет место, если выполняется неравенство

$$F = S^2_{\text{адк}}/S_y^2 \leqslant F(0{,}05;\ f_{\text{адк}};\ f_u)\ ,$$

где

$$S^2_{\text{адк}} = \frac{\sum\limits_{u=1}^{n} (\tilde{y} - y_u)^2}{n-k-1}\ ;$$

y_u – расчетное значение отклика в u-м опыте; $F(0{,}05;\ f_{\text{адк}};\ f_u)$ – критерий Фишера при 5 %-м уровне значимости; $f_{\text{адк}} = n - k - 1$ – число степеней свободы дисперсии адекватности; f_u – число степеней свободы дисперсии воспроизводимости.

Оценка значимости коэффициента регрессии производится с помощью *критерия Стьюдента* (см. прил. 3). Коэффициент считается значимым, если выполняется неравенство

$$|a_i| \geqslant \Delta a_i = t\,(0{,}05;\ f_y)\frac{S_y}{\sqrt{n}},$$

где $t\,(0{,}05;\ f_y)$ – критерий Стьюдента.

Пример 10.2. Для процесса, исследованного в примере 10.1, указать линейную модель, проверить ее адекватность и оценить значимость коэффициентов регрессии.

Решение. Значение оценок дисперсии в каждой точке плана рассчитываем по формуле $S_u^2 = \Delta^2/2$, где Δ – разность между значениями функции цели в параллельных опытах: $\Delta = y_{n_1} - y_{n_2}$ (см. табл. 10.3). Имеем: $S_1^2 = (27{,}0 - 28{,}0)^2/2 = 0{,}50$, $S_2^2 = (15{,}9 - 17{,}1)^2/2 = 0{,}72$, $S_3^2 = 0{,}32$, $S_4^2 = 0{,}02$.

Выясним, воспроизводим ли данный процесс. Для этого найдем значение критерия Кохрена:

$$\sigma = \frac{0{,}72}{0{,}50 + 0{,}72 + 0{,}32 + 0{,}02} = 0{,}4615.$$

Сравнивая найденное значение с табличным $\sigma(0{,}05;\ 4;\ 1) = 0{,}9065$, заключаем, что рассматриваемый процесс воспроизводим. Дисперсия воспроизводимости (ошибка опыта)

$$S_y^2 = \frac{0{,}50 + 0{,}72 + 0{,}32 + 0{,}02}{4} = 0{,}39.$$

Построим линейную модель эксперимента, вычислив коэффициенты регрессии:

$$a_0 = \frac{27{,}5 + 16{,}5 + 22{,}5 + 13{,}5}{4} = 20, \qquad a_1 = \frac{-27{,}5 + 16{,}5 - 22{,}5 + 13{,}5}{4} = -5,$$

$$a_2 = \frac{-27{,}5 - 16{,}5 + 22{,}5 + 13{,}5}{4} = -2.$$

Модель имеет вид $y = 20 - 5x_1 - 2x_2$.

Чтобы проверить адекватность построенной модели условиям эксперимента, воспользуемся критерием Фишера. Найдем:

$$S^2_{\text{адк}} = \frac{1}{4 - 2 - 1} = 1, \quad F = 1/0{,}39 = 2{,}564.$$

Учитывая, что критерий Фишера $F(0{,}05;\ 1;\ 4) = 7{,}7086$ (см. прил. 6) и $2{,}564 < 7{,}7086$, делаем заключение об адекватности рассматриваемой модели.

Так как

$$\Delta a_i = 2{,}7764 \cdot \frac{0{,}39}{4} = 0{,}8669$$

(здесь $t\ (0{,}05;\ f_y) = 2{,}7764$), то $|a_i| > \Delta a_i$, $i = 0, 1, 2$, и коэффициенты регрессии a_0, a_1, a_2 значимы.

Приведенное описание порядка построения планов эксперимента типа 2^3, кодирование факторов и статистических оценок не изменяются с ростом числа факторов, но несколько усложняются.

В табл. 10.4 приведены матрицы ПФЭ 2^k, $k = 2, 3, 4, 5$. Первые четыре выделенные (отчеркнутые) строки представляют собой матрицу планирования 2^2. Для построения матрицы 2^3 первые четыре строки значения факторов x_1, x_2 и x_3 повторяются, а в столбце x_4 записывают четыре значения на верхнем уровне (+1) и четыре на нижнем (−1). Таким образом, выделенные восемь опытов представляют собой матрицу планирования 2^3 и т.д.

Таблица 10.4

ПФЭ 2^k	Номер опыта	Факторы						Параметр
		x_1	x_2	x_3	x_4	x_5	x_6	y
1	2	3	4	5	6	7	8	9
	1	+1	+1	+1	+1	+1	+1	y_1
	2	+1	−1	+1	+1	+1	+1	y_2
2^2	3	+1	+1	−1	+1	+1	+1	y_3
	4	+1	−1	−1	+1	+1	+1	y_4
2^3	5	+1	+1	+1	−1	+1	+1	y_5
	6	+1	−1	+1	−1	+1	+1	y_6
	7	+1	+1	−1	−1	+1	+1	y_7
	8	+1	−1	−1	−1	+1	+1	y_8
	9	+1	+1	+1	+1	−1	+1	y_9
	10	+1	−1	+1	+1	−1	+1	y_{10}
	11	+1	+1	−1	+1	−1	+1	y_{11}
	12	+1	−1	−1	+1	−1	+1	y_{12}
2^4	13	+1	+1	+1	−1	−1	+1	y_{13}
	14	+1	−1	+1	−1	−1	+1	y_{14}
	15	+1	+1	−1	−1	−1	+1	y_{15}
	16	+1	−1	−1	−1	−1	+1	y_{16}
	17	+1	+1	+1	+1	+1	−1	y_{17}
	18	+1	−1	+1	+1	+1	−1	y_{18}
	19	+1	+1	−1	+1	+1	−1	y_{19}
	20	+1	−1	−1	+1	+1	−1	y_{20}
	21	+1	+1	+1	−1	+1	−1	y_{21}
	22	+1	−1	+1	−1	+1	−1	y_{22}
	23	+1	+1	−1	−1	+1	−1	y_{23}
	24	+1	−1	−1	−1	+1	−1	y_{24}

Окончание табл. 10.4

	25	+1	+1	+1	+1	−1	−1	y_{25}
	26	+1	−1	+1	+1	−1	−1	y_{26}
	27	+1	+1	−1	+1	−1	−1	y_{27}
	28	+1	−1	−1	+1	−1	−1	y_{28}
2^5	29	+1	+1	+1	−1	−1	−1	y_{29}
	30	+1	−1	+1	−1	−1	−1	y_{30}
	31	+1	+1	−1	−1	−1	−1	y_{31}
	32	+1	−1	−1	−1	−1	−1	y_{32}

10.3. ДРОБНЫЙ ФАКТОРНЫЙ ЭКСПЕРИМЕНТ

С увеличением числа факторов количество опытов в полном факторном эксперименте резко возрастает. Так, при трех факторах следует поставить $2^3 = 8$ опытов, при пяти – $2^5 = 32$ опыта, а уже при восьми – 256 опытов. Целесообразно сократить число опытов за счет той информации, которую несут эффекты взаимодействия факторов и которая для построения постулируемой линейной модели несущественна.

Рассмотрим снова ПФЭ 2^2 (план этого эксперимента задан табл. 10.5). По нему можно построить модель

$$y = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 .$$

Однако, если есть основание предполагать, что в выбранных интервалах варьирования процесс может быть описан линейной моделью, то достаточно определить всего три коэффициента a_0, a_1, a_2, а квадратичный коэффициент a_{12} должен быть достаточно малым. Поскольку надо определить только три коэффициента, а опытов сделано четыре, остается одна степень свободы, которую можно использовать для включения в схему эксперимента еще одного фактора $x_3 = x_1x_2$. Теперь получилась матрица планирования уже для трех факторов, линейная модель которой

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 .$$

Коэффициенты этой модели рассчитаны по формуле (10.4), а дисперсии – по формуле (10.3).

Наиболее важное отличие приведенного планирования от полного факторного эксперимента заключается в следующем. Уже из табл. 10.5 легко видеть, что значение a_3 в точности совпадает со значением a_{12}. Если в дополнение к столбцам табл. 10.5 построить столбцы x_1x_3 и x_2x_3, они совпадут со столбцами x_2 и x_1 соответственно и, следовательно, коэффициенты a_{13} и a_{23} совпадут с коэффициентами a_2 и a_1, т. е. линейные эффекты смешаны с эффектами взаимодействия. Символически это записывается следующим образом:

$$a_1 \to \alpha_1 + \alpha_{23}, \quad a_2 \to \alpha_2 + \alpha_{13}, \quad a_3 \to \alpha_3 + \alpha_{12},$$

Таблица 10.5

Номер опыта	x_0	x_1	x_2	$x_1x_2 = x_3$	y
1	+	+	+	+	y_1
2	+	–	+	–	y_2
3	+	+	–	–	y_3
4	+	–	–	+	y_4

где a_i – вычисленные выборочные оценки коэффициентов; α_i ,α_{ij} – неизвестные истинные значения коэффициентов, т.е. a_1 является совместной оценкой коэффициентов α_1 и α_{23} и т.д. (или a_1 определяет совместное влияние факторов x_2 и x_3). Вышесказанное свидетельствует о значительной потере информации по сравнению с ПФЭ, но это результат сокращения числа опытов. Действительно, ПФЭ для трех факторов должен содержать $2^3 = 8$ опытов, а в нашем случае (см. табл. 10.5) проводится только 4 опыта. Поскольку эти опыты нужны для построения линейной модели, парными взаимодействиями факторов можно пренебречь, предполагая, что основные эффекты более значимы по сравнению с парными взаимодействиями.

Указанные в табл. 10.5 четыре опыта, поставленные для оценки влияния трех факторов, представляют собой половину ПФЭ или *дробную реплику*. Наиболее распространены так называемые *регулярные дробные реплики*, которые получают делением числа опытов соответствующего ПФЭ на число, кратное 2. Составляют дробные реплики заменой некоторых эффектов взаимодействия новыми независимыми переменными. Эти реплики условно обозначают 2^{k-p}, где p – число линейных эффектов, приравниваемых эффектам взаимодействия. Если, например, ПФЭ 2^6 включает 64 опыта, то его полуреплика содержит $2^{6-1} = 32$ опыта, четвертьреплика – $2^{6-2} = 16$ опытов и т.д. Минимальная дробная реплика для построения линейной модели должна включать $k + 1$ опытов, где k – число факторов.

Дробный факторный эксперимент, указанный в табл. 10.5, представляет собой полуреплику 2^{3-1} , которая построена в результате отождествления

$$x_3 = x_1x_2 . \tag{10.5}$$

Равенство (10.5) называют *генерирующим отношением*, его можно записать в виде

$$x_3^2 = x_1x_2x_3 .$$

Столбец x_3^2 состоит только из +1. Поэтому можно записать:

$$1 = x_1x_2x_3 . \tag{10.6}$$

Если равенство (10.6) умножить на x_1 , получим:

$$x_1 = x_1^2x_2x_3 , \quad x_1^2 = 1, \; x_1 = x_2x_3 .$$

Таблица 10.6

Номер модели	x_0	x_1	x_2	x_3	x_4
1	+	+	+	+	+
2	+	–	+	–	–
3	+	+	–	+	–
4	+	–	–	+	+
5	+	+	+	–	–
6	+	–	+	–	+
7	+	–	–	–	+
8	+	–	–	–	–

Аналогично

$$x_2 = x_1 x_3\,, \quad x_3 = x_1 x_2\,.$$

Следовательно, коэффициенты линейного уравнения будут оценками:

$$a_1 \to \alpha_1 + \alpha_{23}\,, \quad a_2 \to \alpha_2 + \alpha_{13}\,, \quad a_3 \to \alpha_3 + \alpha_{12}\,.$$

Построим матрицу планирования полуреплики 2^{4-1} (табл. 10.6).

Реализация опытов первой полуреплики (см. табл. 10.6) позволяет оценить коэффициенты линейной модели (основной эффект), смешанные с эффектами тройных взаимодействий. Таким образом, с помощью табл. 10.6 можно построить только линейную модель вида

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4\,,$$

а затем определить ее коэффициенты.

10.4. СВОЙСТВА ПОЛНОГО И ДРОБНОГО ФАКТОРНЫХ ЭКСПЕРИМЕНТОВ

Полному и дробному факторным экспериментам присущи следующие свойства.

1. *Симметричность относительно центра эксперимента*:

$$\sum_{u=1}^{N} x_{iu} = 0\,,$$

т.е. сумма элементов любого столбца матрицы планирования равна нулю.

2. *Нормировка*:

$$\sum_{u=1}^{N} x_{iu}^2 = N\,,$$

т.е. сумма квадратов элементов любого столбца равна числу опытов.

3. *Ортогональность*:

$$\sum_{u=1}^{N} x_{iu} x_{ju} = 0, \quad i \neq j \;,$$

т.е. сумма почленных произведений любых двух столбцов равна нулю.

4. *Ротатабельность*, т.е. способность математической модели, полученной в результате полного и дробного экспериментов, предсказывать значение параметра Y с одинаковой точностью на равных расстояниях от центра эксперимента, независимо от направления.

Свойства 1, 2 следуют непосредственно из способа построения матриц планирования, в частности из того, что значения факторов в матрице равны только +1 и −1.

10.5. МАТРИЧНОЕ ОПРЕДЕЛЕНИЕ КОЭФФИЦИЕНТОВ УРАВНЕНИЯ МОДЕЛИ

Рассмотрим уравнение

$$y = a_0 + a_1 x_1 + \ldots + a_k x_k \; . \tag{10.7}$$

В матричной форме уравнение (10.7) можно записать в виде

$$XA = Y \; , \tag{10.8}$$

где X – матрица условий эксперимента:

$$X = \begin{bmatrix} x_{01} & x_{11} & \ldots & x_{k1} \\ x_{02} & x_{12} & \ldots & x_{k2} \\ \ldots & \ldots & \ldots & \ldots \\ x_{0N} & x_{1N} & \ldots & x_{kN} \end{bmatrix} \; ;$$

k – число факторов; N – число опытов; A – матрица неизвестных коэффициентов регрессии: $A^T = [a_0 \; a_1 \ldots a_k]$; Y – матрица результатов наблюдений: $Y^T = [y_1 \; y_2 \; \ldots \; y_N]$.

Умножим равенство (10.8) на транспонированную матрицу X^T:

$$X^T X A = X^T Y \; ,$$

затем на обратную $(X^T X)^{-1}$. Получили:

$$(X^T X)^{-1} (X^T X) A = (X^T X)^{-1} (X^T Y), \quad (X^T X)^{-1} (X^T X) = E,$$

где E – единичная матрица. Следовательно,

$$A = (X^T X)^{-1} (X^T Y) \tag{10.9}$$

и матрица коэффициентов регрессии найдена.

Пример 10.3. Найти коэффициенты регрессии для ПФЭ с двумя факторами.

Р е ш е н и е. В этом случае линейная модель имеет вид

$$y = a_0 + a_1x_1 + a_2x_2 ,$$

т.е. в матричной записи $XA = Y$, где

$$X = \begin{bmatrix} +1 & +1 & +1 \\ +1 & -1 & +1 \\ +1 & +1 & -1 \\ +1 & -1 & -1 \end{bmatrix} ; \qquad A = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} ; \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} .$$

Столбцы матрицы X попарно ортогональны, т.е.

$$\sum_{u=1}^{4} x_{iu}x_{ju} = 0, \qquad i, j = 1, 2, 3 .$$

Например, $\sum\limits_{u=1}^{4} x_{1u}x_{2u} = (+1)(+1) + (-1)(+1) + (+1)(+1) + (-1)(+1) = 0.$

Найдем матрицы X^T и X^TX:

$$X^T = \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \end{bmatrix} ,$$

$$X^TX = \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \end{bmatrix} \begin{bmatrix} +1 & +1 & +1 \\ +1 & -1 & +1 \\ +1 & +1 & -1 \\ +1 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} .$$

Элементы, стоящие на главной диагонали последней матрицы, представляют собой суммы $\sum\limits_{u=1}^{4} x_{iu}^2$. Из условия нормировки каждая такая сумма всегда равна числу опытов.

Матрицы $(X^TX)^{-1}$ и X^TY имеют вид:

$$(X^TX)^{-1} = \begin{bmatrix} 1/4 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1/4 \end{bmatrix} ,$$

$$X^TY = \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} y_1 + y_2 + y_3 + y_4 \\ y_1 - y_2 + y_3 - y_4 \\ y_1 + y_2 - y_3 - y_4 \end{bmatrix} .$$

По формуле (10.9) получим:

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1/4 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1/4 \end{bmatrix} \begin{bmatrix} y_1 + y_2 + y_3 + y_4 \\ y_1 - y_2 + y_3 - y_4 \\ y_1 + y_2 - y_3 - y_4 \end{bmatrix},$$

откуда

$$a_0 = \frac{y_1 + y_2 + y_3 + y_4}{4}, \quad a_1 = \frac{y_1 - y_2 + y_3 - y_4}{4},$$

$$a_2 = \frac{y_1 - y_2 - y_3 - y_4}{4}.$$

В общем случае ортогонального планирования

$$\boxed{a_i = \frac{\sum\limits_{u=1}^{N} x_{iu} y_u}{\sum\limits_{u=1}^{N} x_{iu}^2} = \frac{\sum\limits_{u=1}^{N} x_{iu} y_u}{N}, \quad i = 0, 1, ..., k.}$$

Ортогональность столбцов исходной матрицы планирования X приводит к получению диагональной обратной матрицы $(X^T X)^{-1}$, диагональные элементы которой c_{ii} определяют дисперсии коэффициентов регрессии:

$$S_i^2 = c_{ii} S_y^2,$$

причем последние оказываются одинаковыми:

$$S_i^2 = S_y^2 / \sum_{u=1}^{N} x_{iu}^2.$$

Из равенства дисперсий следует свойство ротатабельности матрицы полного и дробного факторных экспериментов.

ЗАДАЧИ

10.1. Построить математическую модель зависимости частоты вращения, времени непрерывной работы и теплонагруженности агрегата. Исследуемый агрегат представляет собой двухступенчатый редуктор, смазка картерная. Анализ работы конструкции агрегата показал, что основной причиной теплонагруженности агрегата являются условия внутрикартерной гидродинамики, т.е. нагрев зависит от частоты вращения n мин$^{-1}$ (x_2), времени непрерывной работы τ мин (x_3), объема залитого масла V л (x_1). Данные по ПФЭ и его расширенная матрица планирования типа 2^3 приведены соответственно в табл. 10.7 и 10.8.

Таблица 10.7

Пределы варьирования	Факторы		
	Объем залитого масла V, л (x_1)	Частота вращения n, мин$^{-1}$ (x_2)	Время непрерывной работы T, мин (x_3)
Основной уровень (x_{i0})	10	1750	40
Интервал варьирования (Δx_i)	5	250	20
Верхний уровень (+1)	15	2000	60
Нижний уровень (−1)	5	1500	20

Таблица 10.8

Номер опыта	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	y $(t - t_{о.с})$ °C
1	+1	+1	+1	+1	+1	+1	+1	25
2	+1	−1	+1	+1	−1	−1	+1	20
3	+1	+1	−1	+1	−1	+1	−1	38
4	+1	−1	−1	+1	+1	−1	−1	41
5	+1	+1	−1	+1	−1	−1	−1	45
6	+1	−1	+1	−1	−1	+1	−1	26
7	+1	+1	−1	−1	−1	−1	+1	25
8	+1	−1	−1	−1	+1	+1	+1	28

В табл. 10.8 $t_{о.с}$ – температура окружающей среды.

10.2. По результатам опыта (табл. 10.9 и 10.10) построить математическую модель получения продукта (масел) при дистилляции сырья. На выход продукта влияют три фактора: температура в конце дистилляции x_1, скорость нагрева x_2 и продолжительность изотермической выдержки.

Таблица 10.9

Пределы варьирования	Факторы		
	Температура дистилляции x_1, °C	Скорость нагрева x_2, К/мин	Изотермическая выдержка x_3, ч
Основной уровень (x_{i0})	400	6	3
Интервал варьирования (Δx_i)	15	1	1
Верхний уровень (+1)	415	7	4
Нижний уровень (−1)	385	5	2

Таблица 10.10

Номер опыта	x_0	x_1	x_2	x_3	y
1	+				10,5
2	+	–			8,3
3	+		–		6,8
4	+	–	–		4,6
5	+			–	16,4
6	+	–		–	14
7	+		–	–	12,7
8	+	–	–	–	10,5

Опыты в матрице планирования не дублированы.

ОТВЕТЫ

1

1.2. 1) $\overline{A}$ – левое колесо не наехало на препятствие; 2) $\overline{B}$ – правое колесо не наехало на препятствие; 3) $A+B$ – либо левое, либо правое, либо оба колеса одновременно наехали на препятствие; 4) $\overline{A+B}$ – ни одно колесо не наехало на препятствие; 5) $\overline{A}\overline{B}$ – оба колеса не наехали на препятствие. **1.3.** $A = BC$. **1.4.** 1) $A=\Omega$, $B=\varnothing$; 2) $A=\Omega$, $B=\varnothing$; 3) $A=B$. **1.5.** $X=\overline{B}$. **1.6.** Нет, так как $\overline{A+B}=\overline{A}\overline{B}$. **1.8.** $C=(A_1+A_2)(B_1B_2\overline{B}_3+B_1\overline{B}_2B_3+\overline{B}_1B_2B_3+B_1B_2B_3)$. **1.9.** $P=17/22$. **1.10.** 1) $P=560/1111$; 2) $P=480/1111$; 3) $P=40/1111$; 4) $P=30/1111$; 5) $P=1/1111$. **1.11.** $P=1/6$. **1.12.** 1) $P=5/9$; 2) $P=2/9$; 3) $P=7/9$. **1.13.** $P=0{,}08$. **1.14.** $P=1-l/L$. **1.15.** $P=1/4$. **1.16.** $P=0{,}504$. **1.17.** $P=0{,}96$. **1.18.** $P(A)=0{,}197$, $P(B)=0{,}345$. **1.19.** $P(A)=0{,}078$, $P(B)=0{,}683$, $P(C)=0{,}605$. **1.20.** $P=0{,}712$. **1.21.** 0,0022, 0,11. **1.22.** $P=2/9$. **1.23.** $P=13/132$. **1.24.** $P=0{,}9428$. **1.25.** $p_1=0{,}002$, $p_2=0{,}01$, $p_3=0{,}0988$. **1.26.** $p=0{,}78$. **1.27.** $p=0{,}214$. **1.28.** $P=6/13$. **1.29.** $P_1=0{,}103$, $P_2=0{,}277$, $P_3=0{,}620$. **1.30.** $P_{100}(70;\ 86)=0{,}927$. **1.31.** 1) $P_{100}(20;\ 100)\Rightarrow 0{,}55$; 2) $P_{100}(0;\ 28)=0{,}98$; 3) $P_{100}(14;\ 26)=0{,}91$. **1.32.** $P_{100}(50)=0{,}0782$. **1.33.** $P_{21}(11;\ 21)=0{,}96$. **1.34.** $n=100$. **1.35.** $n=177$.

2

2.1. 1)

x_i	0	1	2	3
p_i	0,125	0,375	0,375	0,125

2.2.

x_i	1	2	3	4	5
p_i	0,1	0,09	0,081	0,0729	0,6561

2.3. 1) $P(X=m)=1/2^m$; 2) один опыт.

2.4.

x_i	0	1	2	...	m	...	n
p_i	q^n	$C_n^1pq^{n-1}$	$C_n^2p^2q^{n-2}$	...	$C_n^mp^mq^{n-m}$	...	p^n

2.5.

x_i	0	1	2	3	4
p_i	0,0009	0,0208	0,1525	0,4286	0,3968

2.6. 1) $f(x)=\begin{cases}0 & \text{при } x<0,\\ \dfrac{1}{2}\sin x & \text{при } 0\leqslant x\leqslant\pi,\\ 0 & \text{при } x>\pi.\end{cases}$ **2.7.** 1) $A=4$; 2) $F(x)=1-e^{-2x}(2x^2+2x+1)$; 3) $P(0\leqslant x\leqslant 1/2)=0{,}086$. **2.8.** $f(x)=\dfrac{1}{\sqrt{2\pi}}e^{-x^2/2}$. **2.9.** 1) $p=1/e$; 2) $f(x)=\dfrac{1}{T}e^{-t/T}$.

2.10. 1) $F(x)=\dfrac{1}{2}+\dfrac{1}{\pi}\operatorname{arctg}x$; 2) $P(|x|<1)=1/2$, $a=1/\pi$. **2.11.** $p=1/2$. **2.12.** $p=1/3$.

2.13. $M(X)=1{,}5$, $D(X)=0{,}75$, $\sigma(X)=0{,}86$. **2.14.** $M(X)=60$. **2.15.** $\sigma(X)=(d_2-$

$-d_1)/25$. **2.16.** $M(X) = 0$, $D(X) = 0{,}5$, $a = 1/2$, $b = 1/\pi$. **2.17.** $D(X) = a^2/2$, $\sigma(X) = a/\sqrt{2}$. **2.18.** $M(X) = D(X) = m + 1$. **2.19.** $M(X) = \frac{3}{2}x_0$, $D(X) = \frac{3}{4}x_0^2$. **2.21.** $\mu_1 = 3{,}9$, $\mu_3 = 74{,}1$, **2.22.** $\mu_1 = 1$, $\mu_2 = 1{,}1$, $\mu_3 = 1{,}3$, $\mu_4 = 1{,}6$, $\nu_1 = 0$, $\nu_2 = 0{,}1$, $\nu_3 = 0$, $\nu_4 = 1/35$. **2.23.** 1) Прибавится слагаемое a; 2) не изменится; 3) не изменится; 4) прибавится выражение $a^2 + 2aM(X)$. **2.24.** 1) Умножится на a; 2) умножится на a^2; 3) умножится на $|a|$; 4) умножится на a^2

2.25.

x_i	0	1	2	3
p_i	0,729	0,243	0,027	0,001

2.26.

t_i	2	2,5	3	6
p_i	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2}{2}e^{-\lambda}$	$1 - e^{-\lambda}(1 + \lambda + \frac{\lambda^2}{2})$

$M(T) = 6 - e^{-\lambda}(4 + 3{,}5\lambda + 1{,}5\lambda^2)$.

2.27.

θ_i	τ	2τ	...	$n\tau$
p_i	p	pq	...	pq^{n-1}

$i = \overline{1, n}$.

2.28. $p = 0{,}9$. **2.29.** $\sigma = \sqrt{\dfrac{0{,}57^2 - 0{,}21^2}{2\ln(0{,}57/0{,}21)}}$. **2.30.** 4,6 %. **2.31.** 15 мкм. **2.32.** $P(|X| < 7) = 0{,}56$. **2.33.** 1) $F(x) = (\Phi(\frac{x-a}{\sigma}) - \Phi(\frac{b-a}{\sigma}))/(1 - \Phi(\frac{c-a}{\sigma}))$; 2) $F(x) = (\Phi(\frac{x-a}{\sigma}) - \Phi(\frac{b-a}{\sigma}))/(\Phi(\frac{c-a}{\sigma}) - \Phi(\frac{b-a}{\sigma}))$. **2.34.** $P(|X| < 30) = 0{,}007$. **2.35.** $P(f_q \geqslant f_d) = 0{,}07$.

3

3.1.

x_i	x_1	x_2	x_3
p_i	0,333	0,333	0,334

y_j	y_1	y_2	y_3
p_j	0,250	0,346	0,404

3.2.

y_j \ x_i	0	1
0	0,94	0,02
1	0,03	0,01

x_i	0	1
p_i	0,97	0,03

y_i	0	1
p_i	0,96	0,04

3.3.

y_j \ x_i	0	1
0	0,56	0,14
1	0,24	0,06

x_i	0	1
p_i	0,8	0,2

y_i	0	1
p_i	0,7	0,3

3.4.

y_j \ x_i	0	1	2
0	0	0	0,49
1	0	0,42	0
2	0,09	0	0

3.5.

y_j \ x_i	0	1
0	0,6	0,2
1	0,15	0,05

x_i	0	1
p_i	0,75	0,25

y_i	0	1
p_i	0,8	0,2

3.6.

$$f(x,y) = \begin{cases} \dfrac{1}{(b-a)(d-c)} & \text{при } a \leqslant x \leqslant b, \quad c \leqslant y \leqslant d, \\ 0 & \text{при } \begin{matrix} x > b, & y > d, \\ x < a, & y < c; \end{matrix} \end{cases}$$

$$F(x,y) = F_1(x) F_2(y), \text{где } F_1(x) = \begin{cases} 0 & \text{при } x \leqslant d, \\ \dfrac{x-a}{b-a} & \text{при } a < x < b, \\ 1 & \text{при } x \geqslant b; \end{cases}$$

$$F_2(y) = \begin{cases} 0 & \text{при } y \leqslant c, \\ \dfrac{y-c}{d-c} & \text{при } c < y < d, \\ 1 & \text{при } y \geqslant d. \end{cases}$$

3.7. 1) $A = 20$; 2) $F(x,y) = \left(\frac{1}{\pi}\operatorname{arctg}\frac{x}{4} + \frac{1}{2}\right)\left(\frac{1}{\pi}\operatorname{arctg}\frac{y}{5} + \frac{1}{2}\right)$. **3.8.** $P(\pi/4 < x;\ \pi/6 < y < \pi/3) = 0{,}11$. **3.9.** $C = 0{,}5$, $F(x,y) = 0{,}5(\sin x + \sin y - \sin(x+y))$. **3.10.** $P(x < 1/2;\ y < 1/3) = 9/16$.

3.11. $F_1(x) = \begin{cases} 0 & \text{при } x \leqslant 0, \\ q_1 = 1 - p_1 & \text{при } 0 < x \leqslant 1, \\ 1 & \text{при } x > 1, \end{cases}$ $F_2(x) = \begin{cases} 0 & \text{при } y \leqslant 0, \\ q_2 = 1 - p_2 & \text{при } 0 < y \leqslant 1 \\ 1 & \text{при } y > 1. \end{cases}$

Функция $F(x,y)$ определяется таблицей:

y \ x	$x \leqslant 0$	$0 \leqslant x \leqslant 1$	$x > 1$
$y \leqslant 0$	0	0	0
$0 < y \leqslant 1$	0	$q_1 q_2$	q_2
$y > 1$		q_1	1

3.12. $f(x,y) = \begin{cases} \cos x \cos y & \text{при } x \geqslant 0, \ y \leqslant \pi/2, \\ 0 & \text{при } x < 0, \ y > \pi/2, \end{cases}$ $f_1(x) = \cos x$,

$f_2(y) = \cos y$. **3.13.** $F(x,y) = \begin{cases} \dfrac{e^{-x(1+y)}}{1+y} - e^{-x} & \text{при } x \geqslant 0,\ y \geqslant 0, \\ 0 & \text{при } x < 0,\ y < 0. \end{cases}$

3.14.

$y \backslash x$	0	1	2
0	$q_1 q_2$	$q_2 p_1$	$q_2 p_1^2$
1	$p_2 q_1$	$p_1 p_2$	$p_1^2 p_2$
2	$p_2^2 q_1$	$p_2^2 p_1$	$p_2^2 p_1^2$

3.15. $f_1(x) = \begin{cases} \dfrac{2}{\pi r^2}\sqrt{r^2 - x^2} & \text{при } |x| \leqslant r, \\ 0 & \text{при } |x| > r, \end{cases}$ $f_2(y) = \begin{cases} \dfrac{2}{\pi r^2}\sqrt{r^2 - x^2} & \text{при } |y| \leqslant r, \\ 0 & \text{при } |y| > r. \end{cases}$

3.16. $f_1(x/y) = \dfrac{2}{\sqrt{\pi}} e^{-(2x + \frac{3}{2}y)^2}$, $f_2(y/x) = \dfrac{3}{\sqrt{\pi}} e^{-(x+3y)^2}$.

3.17. $\lambda = \dfrac{1}{20}$,

x_i	10	20	30
p_i	0,2	0,4	0,4

y_i	20	40	60
p_i	0,3	0,35	0,35

3.18. $y = x_1 + x_2$, $f(y) = y - \dfrac{1}{2}$. **3.19.** $M(X) = M(Y) = 0$; $[K_{ij}] = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$.

3.20. $M(X) = 0{,}45$, $M(Y) = 0{,}55$, $D(X) = 0{,}2475$, $D(Y) = 0{,}3375$. **3.21.** $f(x, y) =$

$= \cos x \cos y$ при $0 \leqslant x \leqslant \pi/2$, $0 \leqslant y \leqslant \pi/2$, $M(X) = M(Y) = \pi/2 - 1$, $[K_{ij}] =$

$= \begin{pmatrix} \pi - 3 & 0 \\ 0 & \pi - 3 \end{pmatrix}$. **3.22.** $P(k) = 1 - e^{-k^2/2}$. **3.24.** 1) 0,25; 2) 0,3413; 3) 0,5 ;

4) 0,1573; 5) 0,0476 . **3.25.** $a = 1/(2\pi\sigma_x\sigma_y)$, $\sigma_x = 2$, $\sigma_y = 1$; СВ X и Y независимы; 0,5 . **3.26.** 1) $P(|y| < x) = 0{,}25$; 2) $P(y < x) = 0{,}5$; 3) $P(y < |x|) = 0{,}75$.

3.27. $f(z) = \dfrac{z}{\sigma^2} e^{-z^2/(2\sigma^2)}$. **3.28.** $P(a < \overline{a}) = 1 - e^{-\pi/4}$, $P(a > \overline{a}) = e^{-\pi/4}$,

$\dfrac{P(a < \overline{a})}{P(a > \overline{a})} = \dfrac{0{,}544}{0{,}456} = 1{,}19$, где $\overline{a}$ – средняя амплитуда.

4

4.1. 0,5 . **4.2.** 0,16 . **4.3.** 62 500 . **4.4.** (0,537; 0,557). **4.5.** $p \geq 0{,}936$. **4.6.** 0,996. **4.7.** 0,94 . **4.8.** 0,909. **4.9.** $p = 0{,}95$, $p \geq 0{,}872$. **4.10.** $P(|x-50|< 0{,}5) \geq 0{,}6$.

5

5.1. $m_x(t)=0$, $D_x(t)=\sigma^2$, $K_x(t_1,t_2)=\begin{cases}\sigma^2 & \text{при } t_1=t_2,\\ 0 & \text{при } t_1\neq t_2.\end{cases}$ **5.2.** $m_x(t)=a\cos t$, $D_x(t)=\sigma^2$. **5.3.** $m_x(t)=\sqrt{\pi/2}$, $D_x(t)=(4-\pi)/2$, $K_x(t_1,t_2)=\begin{cases}0 & \text{при } t_1\neq t_2,\\ (4-\pi)/2 & \text{при } t_1=t_2.\end{cases}$

5.4. $m_x(t)=t$, $D_x(t)=1$, $K_x(t_1,t_2)=\begin{cases}0 & \text{при } t_1\neq t_2,\\ 1 & \text{при } t_1=t_2.\end{cases}$ **5.5.** $m_y(t)=2t^3$, $K_y(t_1,t_2)= = 4t_1^3t_2^3e^{-t_1^2-t_2^2}$, $D_y(t)=4t^6e^{-2t^2}$. **5.6.** $m_z(t)=5t^2+20t+1$, $K_z(t_1,t_2)=25t_1^2t_2^2$, $D_z(t)=25t^4$. **5.7.** $m_y(t)=t^2+\frac{4}{3t}(e^t-1)$, $D_y(t)=\frac{4}{9t^2}(1-e^t te^t)^2$; $K_y(t_1,t_2)=\frac{4}{9t_1t_2}(1-e^{t_1}+t_1e^{t_1})(1-e^{t_2}+t_2e^{t_2})$. **5.8.** $m_y(t)=A\cos t$, $D_y(t)=2\alpha D$, $K_y(t_1,t_2)= = 2\alpha e^{-\alpha(t_1-t_2)^2}(2\alpha(t_2-t_1)^2+1)$. **5.9.** $m_x(t)=\sin t$, $K_x(t_1,t_2)=D_1+D_2t_1t_2+ + D_3t_1t_2\sin t_1\sin t_2$, $D_x(t)=D_1+D_2t^2+D_3t^2\sin^2 t$. **5.10.** $X(t)=X_1t+X_2t^2$, $D(X)=2$, $D(X_2)=3$. **5.11.** $Y_1(t)=3t^2+2X_1t+4X_2t^2+6X_3t^3$, $m_{y_1}(t)=3t^2$, $K_{y_1}(t_1,t_2)=4t_1t_2+ + 32t_1^2t_2^2+72t_1^3t_2^3$; $D_{y_1}(t)=4t^2+32t^4+72t^6$. **5.12.** $Z(t)=5t^2+3X_1t\cos 2t+3X_2t\sin 2t$, $m_z(t)=5t^2$, $K_z(t_1,t_2)=18t_1t_2\cos 2(t_2-t_1)$, $D_z(t)=18t^2$. **5.13.** $\frac{2}{\pi}\left(\frac{6{,}864}{\omega^2+121}+ + 0{,}178\left(\frac{0{,}15}{0{,}15^2+(\omega+0{,}36)^2}+\frac{0{,}15}{(\omega-0{,}36)^2+0{,}15^2}\right)\right)$. **5.14.** $\frac{1}{\pi}\left(\frac{0{,}1}{(\omega-0{,}238)^2+0{,}01}+ + \frac{0{,}1}{(\omega+0{,}238)^2+0{,}01}\right)$. **5.15.** $\frac{0{,}2\pi}{\sqrt{0{,}0154}}e^{-\sqrt{0{,}154}\,v_a(T)}$. **5.16.** $m_x(t)=1$, $K_x(t_1, t_2)=\cos\omega_1(t_2-t_1)+2\cos\omega_2(t_2-t_1)$. **5.17.** $m_y(t)=3$, $D_y(t)=4$. **5.18.** 1) $P(k<m)= = e^{-\lambda t}\sum\limits_{k=0}^{m-1}\frac{(\lambda t)^k}{k!}$; 2) $P(2k)=(1+e^{-2\lambda t})/2$. **5.19.** 1) $P(k=0)\cong 0{,}018$; 2) $P(k < 6)\cong 0{,}092$. **5.20.** 1) $P(k=6)=0{,}017$; 2) $P(k<6)=0{,}0184$; 3) $P(k\geq 6)=0{,}9816$.

5.21. $P(k=4) = 0{,}1338$. **5.22.** 1) $\begin{pmatrix} 0{,}37 & 0{,}63 \\ 0{,}36 & 0{,}64 \end{pmatrix}$; 2) $\begin{pmatrix} 0{,}68 & 0{,}32 \\ 0{,}32 & 0{,}68 \end{pmatrix}$.

5.23. 1) $\begin{pmatrix} 4/9 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \end{pmatrix}$; 2) $\begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}$. **5.24.** $P_n = \dfrac{m!}{(m-n)!} \left(\dfrac{\lambda}{\mu}\right)^n P_0$

5.25. Не менее 8 мест. **5.26.** 0,888 заказов.

7

7.1. $]\,12{,}144;\ 13{,}456\,[$. **7.2.** $n = 24$. **7.3.** $n = 25$. **7.4.** $]\,0{,}016;\ 0{,}072\,[$.

9

9.1. $r_в = 0{,}82$. **9.2.** $r_в = 0{,}93$. **9.3.** а) $\overline{y}_x = 0{,}23x + 21{,}78$; б) $\overline{x}_y = 2{,}92y - 27{,}25$. **9.4.** $\overline{y}_x = 68 \cdot 10^{-6}x + 117{,}769$. **9.5.** $\overline{y}_x = -0{,}428x + 0{,}0193$. **9.6.** $\overline{y}_x = -0{,}036x^2 + 2{,}72x + 8{,}607$. **9.7.** $\overline{y}_t = 29{,}82 - 0{,}71t + 0{,}06t^2$. **9.8.** $\overline{y}_x = 2{,}588 + 2{,}065x - 0{,}211x^2$, $\eta_{y/x} = 0{,}99$. **9.9.** $\overline{y}_x = 134x^2$. **9.10.** $\overline{y}_x = 3{,}63/x + 2{,}53$; $\eta_{y/x} = 0{,}88$. **9.11.** Не отвергается, X и Y коррелируемы. **9.12.** Не отвергается, X и Y коррелируемы.

10

10.1. $y = 8{,}8 - 7{,}11x_1 - 144x_2 + 24x_3 + 0{,}3x_1x_2 - 1{,}75x_1x_2 - 0{,}13x_2x_3 + 12{,}8x_1^2$.
10.2. $y = 10{,}5 - 1{,}1x_1 - 1{,}85x_2 + 2{,}95x_3$.

ПРИЛОЖЕНИЯ

1. Нормальное распределение (плотность распределения вероятностей)

Плотность распределения вероятностей нормированного нормального распределения

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} .$$

x	0	1	2	3	4	5	6	7	8	9
0,0	0, 3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

2. Нормальное распределение (функция распределения)

Значение функции $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz$.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000	0,44	0,1700	0,88	0,3106	1,32	0,4066
0,01	0,0040	0,45	0,1736	0,89	0,3133	1,33	0,4082
0,02	0,0080	0,46	0,1772	0,90	0,3159	1,34	0,4099
0,03	0,0120	0,47	0,1808	0,91	0,3186	1,35	0,4115
0,04	0,0160	0,48	0,1844	0,92	0,3212	1,36	0,4131
0,05	0,0199	0,49	0,1879	0,93	0,3238	1,37	0,4147
0,06	0,0239	0,50	0,1915	0,94	0,3264	1,38	0,4162
0,07	0,0279	0,51	0,1950	0,95	0,3289	1,39	0,4177
0,08	0,0319	0,52	0,1985	0,96	0,3315	1,40	0,4192
0,09	0,0359	0,53	0,2019	0,97	0,3340	1,41	0,4207
0,10	0,0398	0,54	0,2054	0,98	0,3365	1,42	0,4222
0,11	0,0438	0,55	0,2088	0,99	0,3389	1,43	0,4236
0,12	0,0478	0,56	0,2123	1,00	0,3413	1,44	0,4251
0,13	0,0517	0,57	0,2157	1,01	0,3438	1,45	0,4265
0,14	0,0557	0,58	0,2190	1,02	0,3461	1,46	0,4279
0,15	0,0596	0,59	0,2224	1,03	0,3485	1,47	0,4292
0,16	0,0636	0,60	0,2257	1,04	0,3508	1,48	0,4306
0,17	0,0675	0,61	0,2291	1,05	0,3531	1,49	0,4319
0,18	0,0714	0,62	0,2324	1,06	0,3554	1,50	0,4332
0,19	0,0753	0,63	0,2357	1,07	0,3577	1,51	0,4345
0,20	0,0793	0,64	0,2389	1,08	0,3599	1,52	0,4357
0,21	0,0832	0,65	0,2422	1,09	0,3621	1,53	0,4370
0,22	0,0871	0,66	0,2454	1,10	0,3643	1,54	0,4382
0,23	0,0910	0,67	0,2486	1,11	0,3665	1,55	0,4394
0,24	0,0948	0,68	0,2517	1,12	0,3686	1,56	0,4406
0,25	0,0987	0,69	0,2549	1,13	0,3708	1,57	0,4418
0,26	0,1026	0,70	0,2580	1,14	0,3729	1,58	0,4429
0,27	0,1064	0,71	0,2611	1,15	0,3749	1,59	0,4441
0,28	0,1103	0,72	0,2642	1,16	0,3770	1,60	0,4452
0,29	0,1141	0,73	0,2673	1,17	0,3790	1,61	0,4463
0,30	0,1179	0,74	0,2703	1,18	0,3810	1,62	0,4474
0,31	0,1217	0,75	0,2734	1,19	0,3830	1,63	0,4484
0,32	0,1255	0,76	0,2764	1,20	0,3849	1,64	0,4495
0,33	0,1293	0,77	0,2794	1,21	0,3869	1,65	0,4505
0,34	0,1331	0,78	0,2823	1,22	0,3883	1,66	0,4515
0,35	0,1368	0,79	0,2852	1,23	0,3907	1,67	0,4525
0,36	0,1406	0,80	0,2881	1,24	0,3925	1,68	0,4535
0,37	0,1443	0,81	0,2910	1,25	0,3944	1,69	0,4545
0,38	0,1480	0,82	0,2939	1,26	0,3962	1,70	0,4554
0,39	0,1517	0,83	0,2967	1,27	0,3980	1,71	0,4564
0,40	0,1554	0,84	0,2995	1,28	0,3997	1,72	0,4573
0,41	0,1591	0,85	0,3023	1,29	0,4015	1,73	0,4582

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,42	0,1628	0,86	0,3051	1,30	0,4032	1,74	0,4591
0,43	0,1664	0,87	0,3078	1,31	0,4049	1,75	0,4599
1,76	0,4608	1,97	0,4756	2,36	0,4909	2,76	0,4971
1,77	0,4616	1,98	0,4761	2,38	0,4913	2,78	0,4973
1,78	0,4625	1,99	0,4767	2,40	0,4918	2,80	0,4974
1,79	0,4633	2,00	0,4772	2,42	0,4922	2,82	0,4976
1,80	0,4641	2,02	0,4783	2,44	0,4927	2,84	0,4977
1,81	0,4649	2,04	0,4793	2,46	0,4931	2,86	0,4979
1,82	0,4656	2,06	0,4803	2,48	0,4934	2,88	0,4980
1,83	0,4664	2,08	0,4812	2,50	0,4938	2,90	0,4981
1,84	0,4671	2,10	0,4821	2,52	0,4941	2,92	0,4982
1,85	0,4678	2,12	0,4830	2,54	0,4945	2,94	0,4984
1,86	0,4686	2,14	0,4838	2,56	0,4948	2,96	0,4985
1,87	0,4693	2,16	0,4846	2,58	0,4951	2,98	0,4986
1,88	0,4699	2,18	0,4854	2,60	0,4953	3,00	0,49865
1,89	0,4706	2,20	0,4861	2,62	0,4956	3,20	0,49931
1,90	0,4713	2,22	0,4868	2,64	0,4959	3,40	0,49966
1,91	0,4719	2,24	0,4875	2,66	0,4961	3,60	0,499841
1,92	0,4726	2,26	0,4881	2,68	0,4963	3,80	0,499928
1,93	0,4732	2,28	0,4887	2,70	0,4965	4,00	0,499968
1,94	0,4738	2,30	0,4893	2,72	0,4967	4,50	0,499997
1,95	0,4744	2,32	0,4898	2,74	0,4969	5,00	0,499997
1,96	0,4750	2,34	0,4904				

3. Основные статистические распределения случайных величин

Распределение χ^2 (хи-квадрат)

Рассмотрим СВ Y, распределенную по нормальному закону с математическим ожиданием $M(Y) = a$ и средним квадратичным отклонением σ.

Случайная величина $U = (Y - a)/\sigma$, называемая *стандартизованной СВ*, распределена по нормальному закону с параметрами $M(U) = 0$ и $\sigma_u = 1$. Квадрат стандартизованной величины $U^2 = \left(\dfrac{Y-a}{\sigma}\right)^2 = \chi^2$ называется *случайной величиной* χ^2 *(хи-квадрат) с одной степенью свободы.*

Рассмотрим n независимых СВ $Y_1, Y_2, \ldots, Y_n$, распределенных по нормальному закону с математическими ожиданиями $a_1, a_2, \ldots, a_n$ и средними квадратичными отклонениями $\sigma_1, \sigma_2, \ldots, \sigma_n$. Образуем для каждой из этих СВ стандартизованную случайную величину $U_i = (Y_i - a_i)/\sigma_i$, $i = 1, 2, \ldots, n$.

Сумма квадратов стандартизованных переменных

$$\chi^2 = U_1^2 + U_2^2 + \ldots + U_n^2 = \frac{(Y_1 - a_1)^2}{\sigma_1^2} + \frac{(Y_2 - a_2)^2}{\sigma_2^2} + \ldots + \frac{(Y_n - a_n)^2}{\sigma_n^2}$$

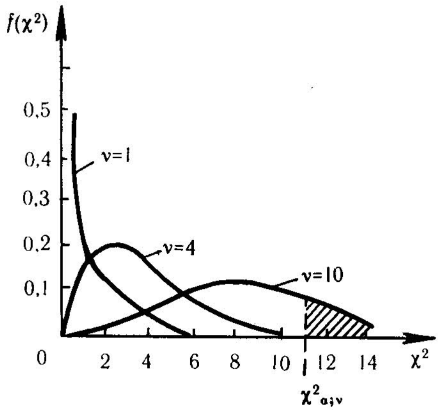


Рис. 1

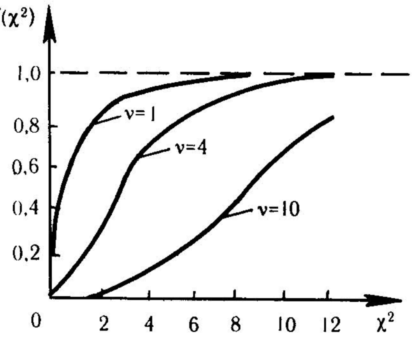


Рис. 2

называется *случайной величиной* χ^2 *с* $\nu = n$ *степенями свободы*. (В статистических таблицах число степеней свободы принято обозначать буквой ν.)

Плотность распределения вероятностей СВ χ^2 имеет вид

$$f(\chi^2) = \begin{cases} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} (\chi^2)^{\nu/2-1} e^{-\chi^2/2} & \text{при } \chi^2 > 0, \\ 0 & \text{при } \chi^2 < 0, \end{cases}$$

где $\Gamma(\alpha) = \int\limits_0^\infty x^{\alpha-1} e^{-x} dx$ – *гамма-функция* (интеграл Эйлера второго рода). Таким образом, распределение χ^2 зависит от одного параметра ν – числа степеней свободы.

Функция распределения χ^2 имеет вид $F(\chi^2) = P(\chi^2 < \chi_0^2)$:

$$F(\chi^2) = \begin{cases} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} \int\limits_0^{\chi_0^2} (\chi^2)^{\nu/2-1} e^{-\chi^2/2} d(\chi^2) & \text{при } \chi^2 > 0, \\ 0 & \text{при } \chi^2 < 0. \end{cases}$$

На рис. 1 и 2 изображены графики плотности распределения вероятностей и функции χ^2-распределения.

На практике часто используются квантили χ^2-распределения $\chi^2_{\alpha,\nu}$. *Квантилем* $\chi^2_{\alpha,\nu}$, отвечающим заданному уровню вероятности α, называется такое значение $\chi^2 = \chi^2_{\alpha,\nu}$, при котором

$$P(\chi^2 > \chi^2_{\alpha,\nu}) = \int\limits_{\chi^2_{\alpha,\nu}}^{\infty} f(\chi^2)\, d(\chi^2) = \alpha .$$

С геометрической точки зрения, нахождение квантиля $\chi^2_{\alpha,\nu}$ заключается в таком выборе значения $\chi^2 = \chi^2_{\alpha,\nu}$, при котором площадь заштрихованной криволинейной трапеции (см. рис. 1) была бы равной α. В прил. 5 приведены значения квантилей $\chi^2_{\alpha,\nu}$ в зависимости от числа ν степеней свободы и вероятностей α.

Распределение Стьюдента

Важную роль при нахождении оценок параметров гипотетического распределения СВ X играет распределение Стьюдента (*t-распределение*).

Пусть $Y, Y_1, \ldots, Y_n$ – независимые СВ, имеющие нормальное распределение с параметрами

$$M(Y) = M(Y_1) = \ldots = M(Y_n) = 0 \text{ и } \sigma_y = \sigma_{y_1} = \ldots = \sigma_{y_n} = 1.$$

Случайная величина

$$t = \frac{Y}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} Y_i^2}} = \frac{Y}{\sqrt{\frac{1}{n}\chi_n^2}}, \tag{1}$$

являющаяся функцией нормально распределенных СВ, имеет *распределение Стьюдента*, а ее плотность

$$f(t) = S(t, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}, \quad -\infty < t < \infty. \tag{2}$$

В формуле (2) буквой ν обозначено число слагаемых в подкоренном выражении соотношения (1), т.е. $\nu = n$.

Можно показать, что математическое ожидание и дисперсия СВ t следующие:

$$M(t) = 0, \quad D(t) = \nu/(\nu-2), \qquad \nu > 2.$$

Из формулы (2) видно, что распределение СВ t зависит только от одного параметра – ν. При $\nu = 1$ распределение Стьюдента является *распределением Коши*, плоность распределения вероятностей которого задается выражением

$$f(t) = \frac{1}{\pi(1+t^2)}.$$

Если $\nu \to \infty$, то

$$(1 + t^2/\nu)^{-(\nu+1)/2} \to e^{-t^2/2}, \quad \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma(\nu/2)\sqrt{\pi\nu}} \to \frac{1}{\sqrt{2\pi}}.$$

Это означает, что последовательности плотностей распределения вероятностей СВ, имеющих распределение Стьюдента, при неограниченном увеличении числа степеней свободы сходятся к стандартизованному нормальному распределению.

В прил. 5 приведены значения квантилей распределения Стьюдента $t_{\alpha/2;\nu}$ в зависимости от числа ν степеней свободы и заданного уровня вероятности α, найденные из решения уравнения

$$P(|t| > t_{\alpha/2;\nu}) = 2\int_{t_{\alpha/2;\nu}}^{\infty} f(t)\,dt = \alpha.$$

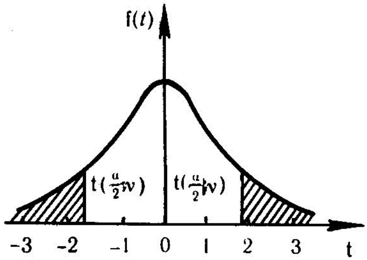


Рис. 3

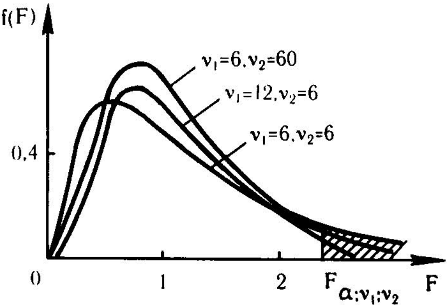


Рис. 4

С геометрической точки зрения нахождение квантилей $t_{\alpha/2;\,\nu}$ заключается в таком выборе значения $t = t_{\alpha/2;\,\nu}$, при котором суммарная площадь заштрихованных на рис.3 криволинейных трапеций равна α.

Распределение Фишера (F-распределение)

Это распределение употребляется при сравнении дисперсий нормальных распределений, вычисленных на основании выборочных данных.

Пусть СВ $X_1, X_2, \ldots, X_m, Y_1, Y_2, \ldots, Y_n$ независимы и имеют нормальное распределение с параметрами $M(X_i) = M(Y_j) = 0$, $D(X_i) = D(Y_j) = 1$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$. Тогда безразмерная СВ

$$F = n \sum_{i=1}^{m} X_i^2 / m \sum_{j=1}^{n} Y_j^2$$

имеет *распределение Фишера*, т.е. ее плотность распределения вероятностей

$$f(F) = \begin{cases} \dfrac{\Gamma\left(\dfrac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\dfrac{\nu_1}{2}\right)\Gamma\left(\dfrac{\nu_2}{2}\right)} \left(\dfrac{\nu_1}{\nu_2}\right)^{\nu_1/2} \dfrac{F^{\nu_1/2-1}}{\left(1+\dfrac{\nu_1}{\nu_2}F\right)^{(\nu_1+\nu_2)/2}} & \text{при } F > 0, \\ 0 & \text{при } F < 0, \end{cases} \tag{3}$$

где $\nu_1 = m$ – число степеней свободы числителя; $\nu_2 = n$ – число степеней свободы знаменателя.

Из формулы (3) следует, что распределение СВ F зависит от двух параметров – чисел $\nu_1 = m$ и $\nu_2 = n$ степеней свободы. График плотности распределения вероятностей F-распределения изображен на рис. 4.

В прил. 6 приведены значения квантилей F-распределения $F_{\alpha;\,\nu_1;\,\nu_2}$ в зависимости от числа степеней свободы ν_1 и ν_2. Значения квантилей найдены из решения уравнения

$$P(F > F_{\alpha;\,\nu_1;\,\nu_2}) = \int_{F_{\alpha;\,\nu_1;\,\nu_2}}^{\infty} f(F)\,dF = \alpha,$$

т.е. площадь заштрихованной на рис. 4 фигуры равна заданному уровню вероятности.

4. Таблица критических значений $t_{\alpha;\,\nu}$, удовлетворяющих условию

$$P(t \geqslant t_{\alpha;\,\nu}) = \int_{t_{\alpha;\,\nu}}^{\infty} p(t, \nu)\,dt = \alpha$$

ν \ α	0,40	0,30	0,20	0,10	0,050	0,025	0,010	0,005	0,001	0,0005
1	0,325	0,727	1,376	3,078	6,314	12,71	31,82	63,66	318,3	636,6
2	0,289	0,617	1,061	1,886	2,920	4,303	6,965	9,925	22,33	31,60
3	0,277	0,584	0,978	1,638	2,353	3,182	4,541	5,841	10,22	12,94
4	0,271	0,569	0,941	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	0,267	0,559	0,920	1,476	2,015	2,571	3,365	5,032	5,893	6,859
6	0,265	0,553	0,906	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	0,263	0,549	0,896	1,415	1,895	2,365	2,998	3,499	4,785	5,405
8	0,262	0,546	0,889	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	0,261	0,543	0,883	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	0,260	0,542	0,879	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	0,260	0,540	0,876	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	0,259	0,539	0,873	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	0,259	0,538	0,870	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	0,258	0,537	0,868	1,345	1,761	2,145	2,624	3,977	3,787	4,140
15	0,258	0,536	0,866	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	0,258	0,535	0,865	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	0,257	0,534	0,863	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	0,257	0,534	0,862	1,330	1,734	2,101	2,552	2,878	3,611	3,922
19	0,257	0,533	0,861	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	0,257	0,533	0,860	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	0,257	0,532	0,859	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	0,256	0,532	0,858	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	0,256	0,532	0,858	1,319	1,714	2,069	2,500	2,807	3,485	3,767
24	0,256	0,531	0,857	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	0,256	0,531	0,856	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	0,256	0,531	0,856	1,315	1,706	2,056	2,479	2,779	3,435	3,707
27	0,256	0,531	0,855	1,314	1,703	2,052	2,473	2,771	3,421	3,690
28	0,256	0,530	0,855	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	0,256	0,530	0,854	1,311	1,699	2,045	2,462	2,756	3,396	3,659
30	0,256	0,530	0,854	1,310	1,697	2,042	2,457	2,750	3,385	3,646
40	0,255	0,529	0,851	1,303	1,684	2,021	2,423	2,704	3,307	3,551
50	0,255	0,528	0,849	1,298	1,676	2,009	2,403	2,678	3,262	3,495
60	0,254	0,527	0,848	1,296	1,671	2,000	2,390	2,660	3,232	3,460
80	0,254	0,527	0,846	1,292	1,664	1,990	2,374	2,639	3,195	3,415
100	0,254	0,526	0,845	1,290	1,660	1,984	2,365	2,626	3,174	3,389
200	0,254	0,525	0,843	1,286	1,653	1,972	2,345	2,601	3,131	3,339
500	0,253	0,525	0,842	1,283	1,648	1,965	2,334	2,586	3,106	3,310
∞	0,253	0,524	0,842	1,282	1,645	1,960	2,326	2,576	3,090	3,291

5. Таблица значений $\chi^2_{\alpha,\nu}$

ν \ α	0,20	0,10	0,05	0,02	0,01	0,001
1	1,642	2,706	3,841	5,412	6,635	10,827
2	3,219	4,605	5,991	7,824	9,210	13,815
3	4,642	6,251	7,815	9,837	11,345	16,266
4	5,989	7,779	9,488	11,668	13,277	18,467
5	7,289	9,236	11,070	13,388	15,086	20,515
6	8,558	10,645	12,592	15,033	16,812	22,457
7	9,803	12,017	14,067	16,622	18,475	24,322
8	11,030	13,362	15,507	18,168	20,090	26,125
9	12,242	14,684	16,919	19,679	21,666	27,877
10	13,442	15,987	18,307	21,161	23,209	29,588
11	14,631	17,275	19,675	22,618	24,725	31,264
12	15,812	18,549	21,026	24,054	26,217	32,909
13	16,985	19,812	22,362	25,472	27,688	34,528
14	18,151	21,064	23,685	26,783	29,141	36,123
15	19,311	22,307	24,996	28,259	30,578	37,697
16	20,465	23,542	26,296	29,633	32,000	39,252
17	21,615	24,769	27,587	30,995	33,409	40,790
18	22,760	25,989	28,869	32,346	34,805	42,312
19	23,900	27,204	30,144	33,687	36,191	43,820
20	25,038	28,412	31,410	35,020	37,566	45,315
21	26,171	29,615	32,671	36,343	38,932	46,797
22	27,301	30,813	33,924	37,659	40,289	48,268
23	28,429	32,007	35,172	38,968	41,638	49,728
24	29,553	33,196	36,415	40,270	42,980	51,179
25	30,675	34,382	37,652	41,566	44,314	52,620
26	31,795	35,563	38,885	42,856	45,642	54,052
27	32,912	36,741	40,113	44,140	46,963	55,476
28	34,027	37,916	41,337	45,419	48,278	56,893
29	35,139	39,087	42,557	46,693	49,588	58,302
30	36,250	40,256	43,773	47,962	50,892	59,703

6. Таблица критических значений $F_{\alpha;\nu_1;\nu_2}$ для $\alpha = 0{,}05$ распределения Фишера $P(F \geqslant F_{\alpha;\nu_1;\nu_2}) = 0{,}05$

ν_2 \ ν_1	1	2	3	4	5	6	8	12	24	∞
1	161,45	199,50	215,72	224,57	230,17	233,97	238,89	243,91	249,04	254,32
2	18,512	18,999	19,163	19,248	19,298	19,329	19,371	19,414	19,453	19,496
3	10,129	9,552	9,276	9,118	9,014	8,941	8,844	8,744	8,638	8,527
4	7,710	6,945	6,591	6,388	6,257	6,164	6,041	5,912	5,774	5,628
5	6,607	5,786	5,410	5,192	5,050	4,950	4,818	4,678	4,527	4,365
6	5,987	5,143	4,756	4,534	4,388	4,284	4,147	4,000	3,841	3,669
7	5,591	4,737	4,347	4,121	3,972	3,866	3,725	3,574	3,410	3,230
8	5,317	4,459	4,067	3,838	3,688	3,580	3,438	3,284	3,116	2,928
9	5,117	4,256	3,863	3,633	3,482	3,374	3,230	3,073	2,900	2,707
10	4,965	4,103	3,708	3,478	3,326	3,217	3,072	2,913	2,737	2,538
11	4,844	3,982	3,587	3,357	3,204	3,094	2,948	2,788	2,609	2,405
12	4,747	3,885	3,490	3,259	3,106	2,999	2,848	2,686	2,505	2,296
13	4,667	3,805	3,410	3,179	3,025	2,915	2,767	2,604	2,420	2,207
14	4,600	3,739	3,344	3,112	2,958	2,848	2,699	2,534	2,349	2,131
15	4,543	3,683	3,287	3,056	2,901	2,790	2,641	2,475	2,288	2,066
16	4,494	3,634	3,239	3,007	2,853	2,741	2,591	2,424	2,235	2,010
17	4,451	3,592	3,197	2,965	2,810	2,699	2,548	2,381	2,190	1,961
18	4,414	3,555	3,160	2,928	2,773	2,661	2,510	2,342	2,150	1,917
19	4,381	3,522	3,127	2,895	2,740	2,629	2,477	2,308	2,114	1,878
20	4,351	3,493	3,098	2,866	2,711	2,599	2,447	2,278	2,083	1,843
21	4,325	3,467	3,072	2,840	2,685	2,573	2,421	2,250	2,054	1,812
22	4,301	3,443	3,049	2,817	2,661	2,549	2,397	2,226	2,028	1,783
23	4,279	3,422	3,028	2,795	2,640	2,528	2,375	2,203	2,005	1,757
24	4,260	3,403	3,009	2,777	2,621	2,508	2,355	2,183	1,984	1,733
25	4,242	3,385	2,991	2,759	2,603	2,490	2,337	2,165	1,965	1,711
26	4,225	3,369	2,975	2,743	2,587	2,474	2,321	2,148	1,947	1,691
27	4,210	3,354	2,961	2,728	2,572	2,459	2,305	2,132	1,930	1,672
28	4,196	3,340	2,947	2,714	2,558	2,445	2,292	2,118	1,915	1,654
29	4,183	3,328	2,934	2,702	2,545	2,432	2,278	2,104	1,901	1,638
30	4,171	3,316	2,922	2,690	2,534	2,421	2,266	2,092	1,887	1,622
40	4,085	3,232	2,839	2,606	2,449	2,336	2,180	2,004	1,793	1,509
60	4,001	3,151	2,758	2,525	2,368	2,254	2,097	1,918	1,700	1,389
120	3,920	3,072	2,680	2,447	2,290	2,175	2,106	1,834	1,608	1,254
∞	3,841	2,996	2,605	2,372	2,214	2,098	1,938	1,752	1,517	1,000

7. Доверительные интервалы для σ

Нижние ν_1 и верхние ν_2 границы доверительного интервала

$$\nu_1 S < \sigma < \nu_2 S \quad \left(S = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2}\right).$$

p / $\nu = n-1$	0,99		0,98		0,95		0,90	
	ν_1	ν_2	ν_1	ν_2	ν_1	ν_2	ν_1	ν_2
1	0,356	159	0,388	79,8	0,446	31,9	0,510	15,9
2	0,434	14,1	0,466	9,97	0,521	6,28	0,578	4,40
3	0,483	6,47	0,514	5,11	0,566	3,73	0,620	2,92
4	0,519	4,39	0,549	3,67	0,599	2,87	0,649	2,37
5	0,546	3,48	0,576	3,00	0,624	2,45	0,672	2,090
6	0,569	2,98	0,597	2,62	0,644	2,202	0,690	1,916
7	0,588	2,66	0,616	2,377	0,661	2,035	0,705	1,797
8	0,604	2,440	0,631	2,205	0,675	1,916	0,718	1,711
9	0,618	2,277	0,644	2,076	0,688	1,826	0,729	1,645
10	0,630	2,154	0,656	1,977	0,699	1,755	0,739	1,593
11	0,641	2,056	0,667	1,898	0,708	1,698	0,748	1,550
12	0,651	1,976	0,677	1,833	0,717	1,651	0,755	1,515
13	0,660	1,910	0,685	1,779	0,725	1,611	0,762	1,485
14	0,669	1,854	0,693	1,733	0,732	1,577	0,769	1,460
15	0,676	1,806	0,700	1,694	0,739	1,548	0,775	1,437
16	0,683	1,764	0,707	1,659	0,745	1,522	0,780	1,418
17	0,690	1,727	0,713	1,629	0,750	1,499	0,785	1,400
18	0,696	1,695	0,719	1,602	0,756	1,479	0,790	1,385
19	0,702	1,666	0,725	1,578	0,760	1,460	0,794	1,370
20	0,707	1,640	0,730	1,556	0,765	1,444	0,798	1,358
21	0,712	1,617	0,734	1,536	0,769	1,429	0,802	1,346
22	0,717	1,595	0,739	1,519	0,773	1,416	0,805	1,335
23	0,722	1.576	0,743	1,502	0,777	1,402	0,809	1,326
24	0,726	1,558	0,747	1,487	0,781	1,391	0,812	1,316
25	0,730	1,541	0,751	1,473	0,784	1,380	0,815	1,308
26	0,734	1,526	0,755	1,460	0,788	1,371	0,818	1,300
27	0,737	1,512	0,758	1,448	0,791	1,361	0,820	1,293
28	0,741	1,499	0,762	1,436	0,794	1,352	0,823	1,286
29	0,744	1,487	0,765	1,426	0,796	1,344	0,825	1,279
30	0,748	1,475	0,768	1,417	0,799	1,337	0,828	1,274
40	0,774	1,390	0,792	1,344	0,821	1,279	0,847	1,228
50	0,793	1,336	0,810	1,297	0,837	1,243	0,861	1,199
60	0,808	1,299	0,824	1,265	0,849	1,217	0,871	1,179
70	0,820	1,272	0,835	1,241	0,858	1,198	0,879	1,163
80	0,829	1,250	0,844	1,222	0,866	1,183	0,886	1,151
90	0,838	1,233	0,852	1,207	0,873	1,171	0,892	1,141
100	0,845	1,219	0,858	1,195	0,878	1,161	0,897	1,133
200	0,887	1,15	0,897	1,13	0,912	1,11	0,925	1,09

8. Таблица значений $t_\nu = t(\nu, n)$

n \ ν	0,95	0,99	0,999	n \ ν	0,95	0,99	0,999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	3,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	4,729	3,600
9	2,31	3,36	5,04	40	2,023	2,708	4,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,991	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

9. Таблица значений $q = q(\nu, n)$

n \ ν	0,95	0,99	0,999	n \ ν	0,95	0,99	0,999
5	1,37	2,67	5,64	20	0,37	0,58	0,88
6	1,09	2,01	3,88	25	0,32	0,49	0,73
7	0,92	1,62	2,98	30	0,28	0,43	0,63
8	0,80	1,38	2,42	35	0,26	0,38	0,56
9	0,71	1,20	2,06	40	0,24	0,35	0,50
10	0,65	1,08	1,80	45	0,22	0,32	0,46
11	0,59	0,98	1,60	50	0,21	0,30	0,43
12	0,55	0,90	1,45	60	0,188	0,269	0,38
13	0,52	0,83	1,33	70	0,174	0,245	0,34
14	0,48	0,78	1,23	80	0,161	0,226	0,31
15	0,46	0,73	1,15	90	0,151	0,211	0,29
16	0,44	0,70	1,07	100	0,143	0,198	0,27
17	0,42	0,66	1,01	150	0,115	0,160	0,221
18	0,40	0,63	0,96	200	0,099	0,136	0,185
19	0,39	0,60	0,92	250	0,089	0,120	0,162

10. Таблица значений вероятностей $P(\chi^2 \geqslant \kappa_q^2)$

κ_q^2 \ k	1	2	3	4	5	6	7	8
1	0,3173	0,6065	0,8013	0,9098	0,9626	0,9856	0,9948	0,9982
2	1574	3679	5724	7358	8491	9197	9598	9810
3	0833	2231	3916	5578	7000	8088	8850	9344
4	0455	1353	2615	4060	5494	6767	7798	8571
5	0254	0821	1718	2873	4159	5438	6600	7576
6	0143	0498	1116	1991	3062	4232	5398	6472
7	0081	0302	0719	1359	2206	3208	4289	5366
8	0047	0183	0460	0916	1562	2381	3326	4335
9	0027	0111	0293	0611	1091	1736	2527	3423
10	0016	0067	0186	0404	0752	1247	1886	2650
11	0009	0041	0117	0266	0514	0884	1386	2017
12	0005	0025	0074	0174	0348	0620	1006	1512
13	0003	0015	0046	0113	0234	0430	0721	1119
14	0002	0009	0029	0073	0146	0296	0512	0818
15	0001	0006	0018	0047	0104	0203	0360	0591
16	0001	0003	0011	0030	0068	0138	0251	0424
17	0000	0002	0007	0019	0045	0093	0174	0301
18		0001	0004	0012	0029	0062	0120	0212
19		0001	0003	0008	0019	0042	0082	0149
20		0000	0002	0005	0013	0028	0056	0103
21		0000	0001	0003	0008	0018	0038	0071
22		0000	0001	0002	0005	0012	0025	0049
23		0000	0000	0001	0003	0008	0017	0034
24		0000	0000	0001	0002	0005	0011	0023
25		0000	0000	0001	0001	0003	0008	0016
26		0000	0000	0000	0001	0002	0005	0010
27		0000	0000	0000	0001	0001	0003	0007
28		0000	0000	0000	0000	0001	0002	0005
29		0000	0000	0000	0000	0001	0001	0003
30		0000	0000	0000	0000	0000	0001	0002

11. Таблица критических значений критерия Кохрена G_k

n	m							
	4	5	6	7	8	9	10	17
				$P = 0{,}95$				
3	0,798	0,746	0,707	0,677	0,653	0,633	0,617	0,547
4	0,684	0,629	0,589	0,560	0,537	0,518	0,502	0,437
5	0,589	0,544	0,507	0,478	0,456	0,439	0,424	0,365
6	0,532	0,480	0,445	0,418	0,398	0,382	0,368	0,314
7	0,480	0,431	0,397	0,373	0,354	0,338	0,326	0,276
8	0,438	0,391	0,359	0,336	0,319	0,304	0,293	0,246
10	0,373	0,331	0,303	0,282	0,267	0,254	0,244	0,203
12	0,326	0,288	0,262	0,244	0,230	0,219	0,210	0,174
15	0,276	0,242	0,219	0,203	0,191	0,182	0,144	0,143
20	0,220	0,192	0,174	0,160	0,150	0,142	0,136	0,111
				$P = 0{,}99$				
3	0,883	0,834	0,793	0,761	0,734	0,711	0,691	0,606
4	0,781	0,721	0,676	0,641	0,613	0,590	0,570	0,488
5	0,696	0,633	0,588	0,553	0,526	0,504	0,485	0,409
6	0,626	0,564	0,520	0,487	0,461	0,440	0,423	0,353
7	0,569	0,508	0,466	0,435	0,411	0,391	0,375	0,310
8	0,521	0,463	0,423	0,393	0,370	0,352	0,337	0,278
10	0,447	0,393	0,357	0,331	0,311	0,295	0,281	0,230
12	0,392	0,343	0,310	0,286	0,268	0,254	0,242	0,196
15	0,332	0,288	0,259	0,239	0,223	0,210	0,200	0,161
20	0,265	0,229	0,205	0,188	0,175	0,165	0,157	0,125

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

1. *Булдык Г.М.* Теория вероятностей и математическая статистика. – Мн.: Выш.шк., 1989.
2. *Вентцель Е.С., Овчаров Л.А.* Прикладные задачи теории вероятностей. – М.: Радио и связь, 1983.
3. *Герасимович А.И.* Математическая статистика. – Мн.: Выш. шк., 1985.
4. *Гмурман В.Е.* Теория вероятностей и математическая статистика. – М.: Высш. шк., 1977.
5. *Гурский Е.И.* Теория вероятностей с элементами математической статистики. – М.: Высш. шк., 1971.
6. *Ермаков С.С.* и др. Математическая теория планирования эксперимента. – М.: Наука, 1983.
7. *Жевержеев В.Ф., Кальницкий Л.А., Сапогов Н.А.* Специальный курс высшей математики для втузов. – М.: Высш. шк., 1970.
8. *Микулик Н.А.* Динамические системы с реактивными звеньями. – Мн.: Выш. шк., 1985.
9. *Микулик Н.А., Рейзина Г.Н.* Руководство к решению технических задач по теории вероятностей и математической статистике. – Мн.: Выш. шк., 1977.
10. *Пугачев В.С.* Теория вероятностей и математическая статистика. – М.: Наука, 1979.
11. Сборник задач по математике для втузов: Специальные курсы / Э.А. Вуколов, А.В. Ефимов, В.Н. Земсков и др.; Под ред А.В. Ефимова. – М.: Наука, 1984.
12. Сборник задач по теории вероятностей, математической статистике и теории случайных функций / Под ред. А.А. Свешникова. – М.: Наука, 1970.
13. *Смирнов Н.В., Дунин-Барковский И.В.* Курс теории вероятностей и математической статистики для технических приложений. – М.: Наука, 1965.
14. *Чистяков В.П.* Курс теории вероятностей. – М.: Наука, 1982.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

ОГЛАВЛЕНИЕ

Справочное издание

Микулик Николай Александрович
Рейзина Галина Николаевна

РЕШЕНИЕ ТЕХНИЧЕСКИХ ЗАДАЧ ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ

Справочное пособие

Заведующий редакцией *Л.Д. Духвалов*
Редактор *М.С. Молчанова*
Художественный редактор *Ю.С. Сергачев*
Технический редактор *Л.И. Счисленок*
Корректор *В.В. Неверко*
Оператор *И.В. Скубий*

ИБ № 3125

Подписано в печать 30.08.91 г. Формат 60х90/16. Бумага кн.журн. Гарнитура Пресс-Роман
Офсет. печать. Усл. печ. л. 10,25. Усл. кр.-отт. 10,625. Уч.-изд.л. 11,26. Тираж 3 700 экз.
Заказ 5603. Цена 2 р. 30 к.

Издательство "Вышэйшая школа" Государственного комитета БССР по печати. 220048 Минск, проспект Машерова, 11.

Типография "Победа", 222310, г. Молодечно, ул. Тавлая, 11