Е. И. Панченко, Л. А. Галимова

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Омск
Издательство ОмГТУ
2013

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное
образовательное учреждение
высшего профессионального образования
«Омский государственный технический университет»

Е. И. Панченко, Л. А. Галимова

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

*Учебное электронное издание
локального распространения*

Омск
Издательство ОмГТУ
2013

Авторы: Е. И. Панченко, Л. А. Галимова (текст издания)

Рецензенты:
Т. А. Ширшова, канд. пед. наук, доцент кафедры преподавания математики ОмГУ им. Ф. М. Достоевского;
С. В. Матвеева, доцент кафедры высшей математики СибАДИ

Редактор К. В. Муковоз
Компьютерная верстка А. Ю. Углиржа

Рекомендовано редакционно-издательским советом Омского государственного технического университета

Издательство ОмГТУ
644050, Омск, пр. Мира, 11
E-mail: info@omgtu.ru

ОГЛАВЛЕНИЕ

Введение 5

1. Теория вероятностей 6
   1.1. Классическое определение вероятности 7
   1.2. Элементы комбинаторики 7
   1.3. Действия над событиями 11
   1.4. Полная группа событий 13
   1.5. Противоположные события 13
   1.6. Вероятность появления хотя бы одного события 14
   1.7. Теоремы сложения и умножения вероятностей 14
      1.7.1. Первая теорема сложения 14
      1.7.2. Первая теорема умножения 14
   1.8. Условная вероятность 15
      1.8.1. Вторая теорема умножения вероятностей зависимых событий 15
      1.8.2. Вторая теорема сложения 16
   1.9. Формула полной вероятности 17
   1.10. Формула Байеса 18
   1.11. Повторение испытаний. Формула Бернулли 19
   1.12. Наивероятнейшее число появлений события в независимых испытаниях 20
   1.13. Локальная теорема Муавра – Лапласа 21
   1.14. Интегральная теорема Муавра – Лапласа 23
   1.15. Формула Пуассона 25
   1.16. Случайные величины 25
   1.17. Дискретные случайные величины и их числовые характеристики 26
   1.18. Непрерывные случайные величины и их числовые характеристики 29
   1.19. Основные законы распределения 35
   1.20. Простейший поток событий 38

2. Математическая статистика 41
2.1. Теория оценок 41
2.1.1. Полигон относительных частот статистического распределения 43
2.1.2. Числовые характеристики выборки. Выборочная средняя, выборочная дисперсия 45
2.2. Статистическая проверка гипотез 47
3. Задачи для самостоятельной работы 66
3.1. Классическое определение вероятности. Формулы комбинаторики 66
3.2. Действия с событиями 67
3.3. Формула полной вероятности. Формула Байеса 69
3.4. Схема с повторением испытаний 71
3.5. Случайные величины 72
Заключение 75
Библиографический список 76

ВВЕДЕНИЕ

Прежде чем перейти к изложению понятий и формул теории вероятностей, рассмотрим кратко, почему у людей возникла необходимость классифицировать и предсказывать казалось бы совершенно случайные процессы.

В окружающем нас мире происходит большое количество событий, зависящих от случайных условий, которые, в свою очередь, также могут быть случайными.

С древних времен люди часто находились в таких условиях, что необходимость в учете и классификации каких-либо данных заставляла человечество искать методы и способы обработки данных, состоящих из совокупности очень больших объектов.

Кроме этого, людей всегда интересовало, с какой долей достоверности можно предсказать наступление того или иного события, которое, казалось бы, наступает случайно. Также с давних времен люди заметили, что случайные процессы в жизни все-таки подчиняются закономерностям. Например, ученые подбрасывали обычную монетку десятки тысяч раз для того, чтобы такие закономерности выявить.

Одним из основных понятий главы является определение вероятности события, причем подход к этому определению осуществляется по-разному. Так, существует определение вероятности события не только классическое, но геометрическое и статистическое, что показывает взаимосвязь рассматриваемых глав.

Поэтому в издании сначала рассматриваются основные теоретические выкладки, сопровождающиеся простыми примерами, показывающими, как студенты, обучающиеся заочно, смогут приобрести не только теоретические знания, но также умение решать задачи по теории вероятностей и проверить правильность решений, сверив их с приведенными ответами.

Также при необходимости студенты могут рассмотреть лабораторные работы, в которых приведены проверки гипотезы о распределении генеральной совокупности по данным выборки.

1. ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Теория вероятностей – это раздел математики, в котором изучаются закономерности случайных процессов.

Множество явлений и событий совершаются в окружающем нас мире. События взаимосвязаны, причем одни из них являются следствием (исходом) других. Событие считается *невозможным*, если оно заведомо не произойдет ни при каких условиях. Событие считается *достоверным*, если оно наступает всегда в результате некоторого испытания. Событие будет считаться *случайным*, если оно может произойти или не произойти при данном испытании. Принято обозначать события буквами **A**, **B**, **C**, …, а вероятность события – ***p***. Под вероятностью события будем понимать *число*, характеризующее лишь возможность наступления события. Исходя из этого, вероятность невозможного события равна **0**, а достоверного **– 1**, тогда вероятность случайного события – это число, попадающее в интервал от **0** до **1**.

На рис. 1 в виде схемы представлена классификация событий.

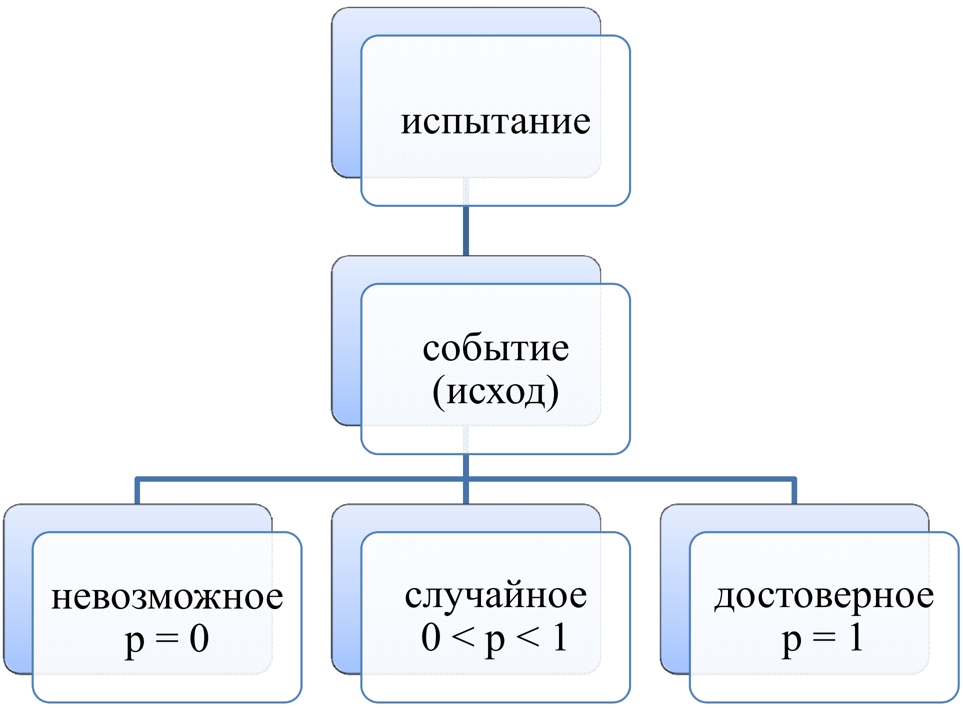


Рис. 1

1.1. Классическое определение вероятности

Определение. Вероятностью случайного события **А** называется отношение числа благоприятных исходов к общему числу несовместных равновозможных исходов:

$$p(\mathrm{A}) = \frac{m_{\mathrm{A}}}{n},$$

где m_{A} – число благоприятных исходов; n – общее число исходов.

Исходы *несовместны*, если при появлении одного из них в одном испытании исключается появление другого.

Под *равновозможными* исходами будем иметь в виду события, которые получаются в результате одинаковых испытаний.

Пример 1

Пусть в ящике находятся 11 одинаковых по размеру и материалу шаров, отличающихся лишь цветом, т. е. 2 белых, 5 черных и 4 красных. Наугад вынимается один шар. Какова вероятность того, что он окажется белым (черным, красным)?

Решение:

Событие А – появление белого шара.

Событие В – появление черного шара.

Событие С – появление красного шара.

Здесь $n = 11$, тогда $m_{\mathrm{A}} = 2, \Rightarrow p(\mathrm{A}) = {}^{2}/_{11}$;

$m_{\mathrm{B}} = 5, \Rightarrow p(\mathrm{B}) = {}^{5}/_{11}$;

$m_{\mathrm{C}} = 4, \Rightarrow p(\mathrm{C}) = {}^{4}/_{11}$.

1.2. Элементы комбинаторики

При решении задач с использованием классического определения вероятности часто используются формулы комбинаторики, которые отвечают на вопрос о количестве комбинаций из n элементов по k, составленных различными способами.

Рассмотрим такие комбинации на примере.

Пример 2

Дано множество D, состоящее из четырех различных цифр, т. е. $D = \{1, 4, 7, 9\}$.

Сколько существует способов выбрать из данного множества подмножеств, состоящих из одного элемента, из двух различных элементов, из трех различных элементов и из четырех элементов?

Составим все возможные подмножества множества D:

1) из одного элемента – таких подмножеств **4**, т. е. $\{1\}$, $\{4\}$, $\{7\}$, $\{9\}$;

2) из двух различных элементов – таких подмножеств **6**, т. е. $\{1; 4\}$, $\{1; 7\}$, $\{1; 9\}$, $\{4; 7\}$, $\{4; 9\}$, $\{7; 9\}$;

3) из трех различных элементов – таких подмножеств **4**, т. е. $\{1; 4; 7\}$, $\{1; 4; 9\}$, $\{1; 7; 9\}$, $\{4; 7; 9\}$;

4) из четырех элементов – одно подмножество $\{1; 4; 7; 9\}$.

Если число элементов в основном множестве D велико, то используют формулу сочетаний из n элементов по k:

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

Подтвердим формулу сочетания следующими вычислениями:

$$C_4^1 = \frac{4!}{1!(4-1)!} = \frac{3! \cdot 4}{1! \cdot 3!} = \frac{4}{1} = 4;$$

$$C_4^2 = \frac{4!}{2!(4-2)!} = \frac{2! \cdot 3 \cdot 4}{2! \cdot 2!} = \frac{12}{2} = 6;$$

$$C_4^3 = \frac{4!}{3!(4-3)!} = \frac{3! \cdot 4}{3! \cdot 1!} = \frac{4}{1!} = 4;$$

$$C_4^4 = \frac{4!}{4!(4-4)!} = \frac{4!}{4! \cdot 0!} = 1.$$

Пример 3

В группе 19 человек. Для участия в конференции выбираются 4 человека. Сколько существует способов это сделать?

Решение:

Каждая четверка из 19 человек – это новый способ отбора делегатов на конференцию, т. е.

$$C_{19}^{4}=\frac{19!}{4!(19-4)!}=\frac{15!\cdot 16\cdot 17\cdot 18\cdot 19}{4!\cdot 15!}=\frac{16\cdot 17\cdot 18\cdot 19}{1\cdot 2\cdot 3\cdot 4}=3876 \text{ способов.}$$

Если же комбинации из n элементов по k отличаются не только составом, но и порядком элементов, то в этом случае используется формула размещения из n по k:

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

Пример 4

Дано множество D = {1; 4; 7; 9}, причем каждому элементу присвоен номер. Цифра 1 имеет номер 1, цифра 4 имеет номер 2, цифра 7 имеет номер 3 и цифра 9 имеет номер 4. Сколько существует способов выбрать из данного множества D подмножеств с различными номерами?

Составим всевозможные подмножества множества D:

1) из одного элемента – 4 способа, т. е. {1}, {4}, {7}, {9};

2) из двух элементов – 12 способов, т. е. {1; 4}, {4; 1}, {1; 7}, {7; 1}, {1; 9}, {9; 1}, {4; 7}, {7; 4}, {4; 9}, {9; 4}, {7; 9}, {9; 7};

3) из трех элементов – 24 способа, т. е. {1; 4; 7}, {4; 1; 7}, {7; 4; 1}, {9; 1; 4}, {1; 7; 4}, {4; 7; 1}, {7; 1; 4}, {9; 4; 1}, {1; 4; 9}, {4; 9; 1}, {7; 1; 9}, {9; 1; 7}, {1; 9; 4}, {4; 1; 9}, {7; 9; 1}, {9; 7; 1}, {1; 7; 9}, {4; 7; 9}, {7; 9; 4}, {9; 4; 7}, {1; 9; 7}, {4; 9; 7}, {7; 4; 9}, {9; 7; 4};

4) из четырех элементов – 24 способа, т. е. {1; 4; 7; 9}, {4; 1; 7; 9}, {7; 1; 4; 9}, {9; 1; 4; 7}, {1; 4; 9; 7}, {4; 1; 9; 7}, {7; 4; 1; 9}, {9; 4; 1; 7}, {1; 7; 4; 9}, {4; 7; 1; 9}, {7; 9; 4; 1}, {9; 1; 7; 4}, {1; 7; 9; 4}, {4; 7; 9; 1}, {7; 9; 1; 4}, {9; 4; 7; 1}, {1; 9; 4; 7}, {4; 9; 1; 7}, {7; 4; 9; 1}, {9; 7; 1; 4}, {1; 9; 7; 4}, {4; 9; 7; 1}, {7; 1; 9; 4}, {9; 7; 4; 1}.

Подтвердим формулу размещения следующими вычислениями:

$$A_4^1=\frac{4!}{(4-1)!}=\frac{3!\cdot 4}{3!}=4,$$

$$A_4^2=\frac{4!}{(4-2)!}=\frac{2!\cdot 3\cdot 4}{2!}=12,$$

$$A_4^3=\frac{4!}{(4-3)!}=\frac{4!}{1!}=24,$$

$$A_4^4=\frac{4!}{(4-4)!}=\frac{4!}{0!}=24.$$

Пример 5

В группе 19 студентов. Необходимо выбрать старосту и профорга. Сколько существует способов это сделать?

Решение:

Так как двойки отличаются и по составу, и по порядку, т. е. являются размещениями из 19 элементов по 2, то их количество находим по формуле

$$A_{19}^2 = \frac{19!}{(19-2)!} = \frac{17!\cdot 18\cdot 19}{17!} = 342 \text{ способа.}$$

Комбинации из n элементов, отличающиеся лишь порядком элементов, будем называть *перестановками*.

Число перестановок из n элементов равно $P_n = n!$

Пример 6

Нужная студенту книга находится в одной из четырех библиотек. Сколько существует способов очередности посещений библиотек?

Решение:

Каждый вариант посещения библиотек отличается только порядком библиотек, т. е. является перестановкой из четырех элементов, поэтому

$$P_4 = 4! = 24 \text{ способа.}$$

Приведем пример непосредственного вычисления вероятности с использованием формул комбинаторики.

Пример 7

В ящике 10 деталей, из них половина окрашена. Наугад берут 2 детали. Найти вероятность того, что обе детали окрашены.

Решение:

А – две выбранные детали окрашены, тогда

$n = C_{10}^2$,

$m_{\text{А}} = C_5^2$,

$p(A) = \frac{m_{\text{А}}}{n} = \frac{C_5^2}{C_{10}^2}$, где $C_5^2 = \frac{5!}{2!\cdot 3!} = \frac{4\cdot 5}{2} = 10$, $C_{10}^2 = \frac{10!}{2!\cdot 8!} = \frac{9\cdot 10}{2} = 45$,

$p(\text{А}) = \frac{10}{45} = \frac{2}{9}$.

Ответ: $p = \frac{2}{9}$.

Часто комбинации бывают *составными*, т. е. они формируются по частям. Это означает, что первая часть комбинации формируется одним из трех описанных выше способов, вторая часть – аналогично первой части и т. д.

Пример 8

В ящике 10 деталей, из них половина окрашена. Наугад берут 3 детали. Найти вероятность того, что из них 2 детали окрашены.

Решение:

А – из трех деталей две окрашены, а одна нет,

$$n = C_{10}^{3},$$

$$m_{\text{А}} = C_5^2 \cdot C_5^1,$$

$$p(A) = m_{\text{А}} / n.$$

Так как в тройке выбранных деталей содержатся *одновременно* 2 окрашенных и 1 неокрашенная детали и в основе комбинаторики лежит принцип умножения, то число благоприятных исходов $m_{\text{А}}$ получается в результате применения числа сочетаний для первой части, состоящей из окрашенных деталей и для второй части, состоящей из неокрашенной детали.

Здесь $n = C_{10}^{3}$,

$$m_{\text{А}} = C_5^2 \cdot C_5^1 = 10 \cdot 5 = 50,$$

$$p(\text{А}) = \frac{50}{120} = \frac{5}{12}.$$

Ответ: $p = \frac{5}{12}$.

1.3. Действия над событиями

Суммой событий А и В является событие С, состоящее в наступлении хотя бы одного из них.

На рис. 2 представлена геометрическая интерпретация суммы двух событий.

Пусть **С = А + В**, тогда С: 1) наступает событие **А** *или*,
2) наступает событие **В** *или*,
3) наступают *и* событие **А**, *и* событие **В**.

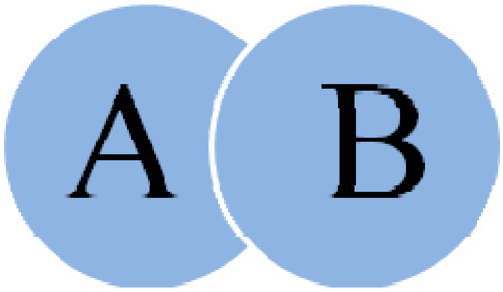


Рис. 2

Например, стрелок стреляет по мишени два раза. Если событие **A** – попадание по мишени при первом выстреле, событие **B** – попадание при втором выстреле, то **(A + B)** – мишень поражена (или первым стрелком, или вторым стрелком, или обоими одновременно).

Замечание. Рассмотрим эту ситуацию для большего количества событий.

Пусть событий в сумме три, тогда наступление хотя бы одного из них означает следующую интерпретацию:

Пусть **D = (A + B + C)**, тогда D:
1) наступает событие **A** *или*,
2) наступает событие **B** *или*,
3) наступает событие **C** *или*,
4) наступают **A** и **B** *или*,
5) наступают **B** и **C** *или*,
6) наступают **A** и **C** *или*,
7) *и* A, *и* B, *и* C.

На рис. 3 приведена геометрическая интерпретация суммы трех событий.

Произведением событий **A** и **B** является событие **C**, состоящее в одновременном (совместном) наступлении и **A**, и **B**.

На рис. 4 представлена геометрическая интерпретация суммы двух событий.

Пусть **C = A·B**, т. е. наступление и **A**, и **B** одновременно.

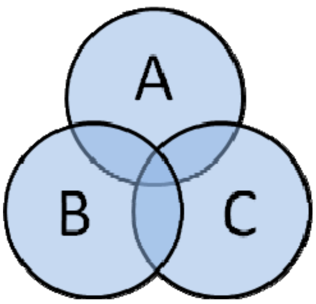


Рис. 3

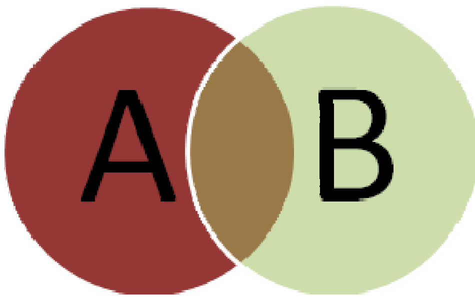


Рис. 4

Например, стрелок стреляет по мишени 2 раза. Если событие **А** – попадание при первом выстреле, событие **В** – попадание при втором выстреле, то $\mathbf{A} \cdot \mathbf{B}$ – это событие, состоящее в том, что мишень поражена дважды.

1.4. Полная группа событий

Полной группой событий является множество таких событий, что в результате опыта должно произойти хотя бы одно из них.

Например, при подбрасывании кубика выпадение граней есть полная группа событий. Причем сумма вероятностей событий, образующих полную группу, равна **1**.

Пример 9

На завод в Омск поступают одни и те же детали: из Нижнего Тагила 70 %, из Москвы 10 % и Липецка. Найти вероятность того, что очередная партия прибудет из Липецка.

Решение:

А – партия деталей прибывает из Липецка;

В – партия деталей прибывает из Нижнего Тагила;

С – партия деталей прибывает из Москвы.

Так как $p(\mathrm{B}) = 0{,}7$ и $p(\mathrm{C}) = 0{,}1$, тогда

$$p(\mathrm{A}) = 1 - (p(\mathrm{B}) + p(\mathrm{C})) = 1 - (0{,}7 + 0{,}1) = 0{,}2.$$

Ответ: $p(\mathrm{A}) = 0{,}2$.

1.5. Противоположные события

Два единственно возможных события, образующих полную группу, называются *противоположными*.

Если А – исходное событие, то $\bar{\mathrm{A}}$ – противоположное событие.

Если $p(\mathrm{A}) = \boldsymbol{p}$, а $p(\bar{\mathrm{A}}) = \boldsymbol{q}$, то $\boldsymbol{p} + \boldsymbol{q} = \mathbf{1}$.

Пример 10

Производится выстрел с вероятностью попадания в цель 0,7. Тогда вероятность промаха равна 0,3.

1.6. Вероятность появления хотя бы одного события

Пусть в результате испытания происходит n независимых в совокупности событий, т. е. могут появиться все или некоторые из них.

Вероятности появления всех событий известны, тогда если событие А – появление хотя бы одного из них, то $p(\text{A}) = 1 - q_1 \cdot q_2 \cdot q_3 \cdot \ldots \cdot q_n$.

1.7. Теоремы сложения и умножения вероятностей

1.7.1. Первая теорема сложения

Вероятность появления одного из двух несовместных событий равна сумме вероятностей этих событий:

$$p(\text{A} + \text{B}) = p(\text{A}) + p(\text{B}).$$

Пример 11

В лотерее разыгрывается 1000 билетов, из них на 10 попадают денежные выигрыши, а на 5 – вещевые. Найти вероятность выигрыша в случае приобретения одного билета.

Решение:

Событие А – билет выиграл,

событие В – выигрыш денег,

событие С – выиграна вещь.

Так как события В и С несовместны, то А = В + С, тогда $p(\text{A}) = p(\text{B}) + p(\text{C})$, $p(\text{B}) = {}^{10}/_{1000} = 0{,}01$, $p(\text{C}) = {}^{5}/_{1000} = 0{,}005$, поэтому $p(\text{A}) = 0{,}01 + 0{,}005 = 0{,}015$.

Ответ: $p(\text{A}) = 0{,}015$.

Два события будем считать *независимыми*, если появление одного из них не меняет вероятность появления другого.

1.7.2. Первая теорема умножения

Вероятность произведения двух независимых событий равна произведению вероятностей этих событий.

Пример 12

В двух ящиках находятся детали. В первом ящике – 7 окрашенных и 2 неокрашенных, а в другом – 8 окрашенных и 3 неокрашенных. Из каждого ящика берут по одной детали. Найти вероятность того, что обе детали окрашены.

Решение:

Событие A – вынута неокрашенная деталь из первого ящика,

событие B – вынута неокрашенная деталь из второго ящика.

Если событие C состоит в том, что обе вынутые детали не окрашены, то $C = A \cdot B$, а $p(C) = p(A) \cdot p(B)$. Так как $p(A) = 2/9$ и $p(B) = 3/11$, то $p(C) = \frac{2}{9} \cdot \frac{3}{11} = \frac{2}{33}$.

Ответ: $p(C) = \frac{2}{33}$.

1.8. Условная вероятность

Условной вероятностью $p(A/B)$ называется вероятность события A, вычисленная в предположении, что событие B наступило (или не наступило).

Пример 13

В ящике 2 белых и 4 красных шара. Из него дважды вынимают по одному шару, не возвращая их в ящик. Найти вероятность появления белого шара во втором испытании.

Решение:

A – вынут красный шар,

B – вынут белый шар.

Значит, нам нужно найти $p(B/A)$ и $p(B/B)$, $\Rightarrow p(B/A) = 2/5$, $p(B/B) = 1/5$.

1.8.1. Вторая теорема умножения вероятностей зависимых событий

Вероятность совместного появления двух зависимых событий равна произведению вероятности одного из них на условную вероятность другого, вычисленную в предположении, что первое наступило.

$$p(A \cdot B) = p(A) \cdot p(B/A), \text{ или } p(A \cdot B) = p(B) \cdot p(A/B).$$

Пример 14

Предприятие изготовляет 99 % стандартных изделий, причем 86 % из них первого сорта. Какова вероятность того, что взятое наудачу изделие, изготовленное на этом предприятии, первого сорта?

Решение:

Событие А – изделие стандартное,

событие В – изделие первого сорта.

Эти события зависимы и $p(\text{A}) = 0{,}99$, а $p(\text{B/A}) = 0{,}86$, тогда $p(\text{A} \cdot \text{B}) = p(\text{A}) \cdot p(\text{B/A}) = 0{,}99 \cdot 0{,}86 = 0{,}85$.

Ответ: $p(\text{A} \cdot \text{B}) = 0{,}85$.

1.8.2. Вторая теорема сложения

Вероятность появления хотя бы одного из двух совместных событий равна сумме вероятностей этих событий без вероятности их совместного появления, т. е.

$$p(\text{A} + \text{B}) = p(\text{A}) + p(\text{B}) - p(\text{A} \cdot \text{B}).$$

Если А и В независимы, то $p(\text{A} + \text{B}) = p(\text{A}) + p(\text{B}) - p(\text{A}) \cdot p(\text{B})$.

Если А и В зависимы, то $p(\text{A} + \text{B}) = p(\text{A}) + p(\text{B}) - p(\text{A}) \cdot p(\text{B/A})$.

Пример 15

Найти вероятность того, что наудачу взятое двузначное число окажется кратным либо 2, либо 5.

Решение:

Событие А – число кратно двум,

событие В – число кратно пяти.

События А и В совместны, т. е. среди них есть числа, кратные одновременно *и* 2, *и* 5.

Найдем $p(\text{A} + \text{B})$.

$p(\text{A}) = \dfrac{45}{90} = \dfrac{1}{2} = 0{,}5;$

$p(\text{B}) = \dfrac{18}{90} = 0{,}2;$

$p(\text{A} \cdot \text{B}) = p(\text{A}) \cdot p(\text{B}) = 0{,}2 \cdot 0{,}5 = 0{,}1;$

$p(\text{A} + \text{B}) = p(\text{A}) + p(\text{B}) - p(\text{A} \cdot \text{B}) = 0{,}5 + 0{,}2 - 0{,}1 = 0{,}6.$

Ответ: $p(\text{A} + \text{B}) = 0{,}6$.

Пример 16

По мишени производятся выстрелы из трех орудий. Вероятность попадания при выстреле из первого орудия 0,8, а из второго – 0,7 и из третьего – 0,9.

Найти вероятности следующих событий:

а) A – попали все три орудия в цель;

б) B – попали ровно два орудия;

в) C – попали не менее двух орудий;

г) D – промах всех трех орудий;

д) E – попало хотя бы одно орудие.

Решение:

а) $p(\mathrm{A}) = p_1 \cdot p_2 \cdot p_3 = 0{,}7 \cdot 0{,}8 \cdot 0{,}9 = 0{,}504$;

б) $p(\mathrm{B}) = p_1 \cdot p_2 \cdot q_3 + p_1 \cdot q_2 \cdot p_3 + q_1 \cdot p_2 \cdot p_3 = 0{,}7 \cdot 0{,}8 \cdot 0{,}1 + 0{,}7 \cdot 0{,}2 \cdot 0{,}9 + 0{,}3 \cdot 0{,}8 \cdot 0{,}9 = 0{,}056 + 0{,}126 + 0{,}216 = 0{,}398$;

в) $p(\mathrm{C}) = p(\mathrm{A}) + \mathrm{p}(\mathrm{B}) = 0{,}504 + 0{,}398 = 0{,}796$;

г) $p(\mathrm{D}) = q_1 \cdot q_2 \cdot q_3 = 0{,}3 \cdot 0{,}2 \cdot 0{,}1 = 0{,}006$;

д) $p(\mathrm{E}) = 1 - q_1 \cdot q_2 \cdot q_3 = 1 - 0{,}006 = 0{,}994$.

1.9. Формула полной вероятности

Одним из следствий совместного применения теорем сложения и умножения вероятностей является формула полной вероятности.

Пусть заданы события $\mathrm{H}_1, \mathrm{H}_2, \ldots, \mathrm{H}_n$, образующие полную группу событий, а интересующее нас событие A наступает при появлении одного из этих событий, т. е. либо H_1, либо H_2, …, либо H_n.

Если известны вероятности гипотез $\mathrm{H}_1, \mathrm{H}_2, \ldots, \mathrm{H}_n$, а также условные вероятности события A при этих гипотезах, т. е. $p(\mathrm{A}/\mathrm{H}_1), p(\mathrm{A}/\mathrm{H}_2), \ldots, p(\mathrm{A}/\mathrm{H}_n)$, то вероятность события A находится по формуле

$$p(\mathrm{A}) = p(\mathrm{H}_1) \cdot p(\mathrm{A}/\mathrm{H}_1) + p(\mathrm{H}_2) \cdot p(\mathrm{A}/\mathrm{H}_2) + \ldots + p(\mathrm{H}_n) \cdot p(\mathrm{A}/\mathrm{H}_n).$$

Пример 17

В группе 30 спортсменов, из них 20 лыжников, 6 конькобежцев и 4 бегуна. Вероятность выполнить норму мастера спорта: для лыжника – 0,9, для конькобежца – 0,8, для бегуна – 0,75. Какова вероятность того, что наудачу выбранный спортсмен выполнит норму мастера спорта?

Решение:

Событие А – выбранный спортсмен выполнит норму мастера спорта,

H_1 – выбран лыжник,

H_2 – выбран конькобежец,

H_3 – выбран бегун.

События H_1, H_2, H_3 образуют полную систему событий, причем

$$p(\mathrm{H}_1) = \frac{20}{30} = \frac{2}{3}, \quad p(\mathrm{A}/\mathrm{H}_1) = 0,9;$$

$$p(\mathrm{H}_2) = \frac{6}{30} = \frac{1}{5}, \quad p(\mathrm{A}/\mathrm{H}_2) = 0,8;$$

$$p(\mathrm{H}_3) = \frac{4}{30} = \frac{2}{3}, \quad p(\mathrm{A}/\mathrm{H}_3) = 0,75.$$

Следовательно, выполнены все условия формулы полной вероятности события А, тогда

$$p(\mathrm{A}) = p(\mathrm{H}_1) \cdot p(\mathrm{A}/\mathrm{H}_1) + p(\mathrm{H}_2) \cdot p(\mathrm{A}/\mathrm{H}_2) + p(\mathrm{H}_3) \cdot p(\mathrm{A}/\mathrm{H}_3) =$$
$$= \frac{20}{30} \cdot 0,9 + \frac{6}{30} \cdot 0,8 + \frac{4}{30} \cdot 0,75 = 0,66 \cdot 0,9 + 0,2 \cdot 0,8 + 0,13 \cdot 0,75 =$$
$$= 0,594 + 0,16 + 0,0975 = 0,8515.$$

1.10. Формула Байеса

Формула Байеса позволяет вычислять вероятности гипотез при наличии дополнительной информации. Нам известно, что испытание уже проведено и известен его результат – наступило событие А. Формула показывает, как найти вероятности каждой гипотезы при наступлении события А, т. е. она переоценивает вероятности гипотез после того, как становится известен исход испытания:

$$p(\mathrm{H}_k/\mathrm{A}) = \frac{p(\mathrm{H}_k) \cdot p(\mathrm{A}/\mathrm{H}_k)}{p(\mathrm{H}_1) \cdot p(\mathrm{A}/\mathrm{H}_1) + p(\mathrm{H}_2) \cdot p(\mathrm{A}/\mathrm{H}_2) + \ldots + p(\mathrm{H}_n) \cdot p(\mathrm{A}/\mathrm{H}_n)},$$

где $1 \le k \le n$.

Пример 18

В группе 30 спортсменов, из них 20 лыжников, 6 конькобежцев и 4 бегуна. Вероятность выполнить норму мастера спорта: для лыжников – 0,9, для конькобежца – 0,8, для бегуна – 0,75. Наудачу выбрали спортсмена, и он оказался мастером спорта. Найти вероятность того, что он конькобежец.

Решение:

Событие А – выбранный спортсмен мастер спорта и конькобежец.

Гипотезы: H_1 – выбранный спортсмен лыжник,

H_2 – выбранный спортсмен конькобежец,

H_3 – выбранный спортсмен бегун.

Гипотезы H_1, H_2, H_3 образуют полную группу, причем $p(H_1) = 2/3$, $p(H_2) = 6/30$, $p(H_3) = 4/30$.

Контроль: $p(H_1) + p(H_2) + p(H_3) = \frac{2}{3} + \frac{6}{30} + \frac{4}{30} = 1.$

$$p(A/H_1) = 0{,}9,\ \ p(A/H_2) = 0{,}8,\ \ p(A/H_3) = 0{,}75.$$

$$p(H_2/A) = \frac{p(H_2)\cdot p(A/H_2)}{p(H_1)\cdot p(A/H_1) + p(H_2)\cdot p(A/H_2) + p(H_3)\cdot p(A/H_3)} =$$

$$= \frac{\frac{6}{30}\cdot 0{,}8}{\frac{20}{30}\cdot 0{,}9 + \frac{6}{30}\cdot 0{,}8 + \frac{4}{30}\cdot 0{,}75} = \frac{\frac{6}{30}\cdot 0{,}8}{0{,}6} = \frac{4}{15} = 0{,}266 \approx 0{,}3.$$

Ответ: $p(H_2/A) \approx 0{,}3$.

1.11. Повторение испытаний. Формула Бернулли

В теории вероятностей имеют место такие события, при которых испытания можно повторять (теоретически) бесконечное число раз, и появление или непоявление некоторого события не будет зависеть от исходов предыдущих испытаний.

Таким образом, мы имеем последовательность независимых одинаковых испытаний, причем вероятность появления события А в каждом испытании одна и та же и равна $\boldsymbol{p}$, а вероятность непоявления события А равна $\boldsymbol{q} = 1 - p$.

Такая система независимых испытаний в теории вероятностей носит название схемы Бернулли.

Например, пусть проводится серия из четырех независимых однотипных испытаний. Что означает появление интересующего нас события А в данной серии ровно 3 раза?

Возможные варианты ситуации будут выглядеть следующим образом: AAAĀ, AAĀA, AĀAA, ĀAAA. Такие ситуации можно составить с помощью числа сочетаний из четырех по три, т. е. $C_4^3 = 6$.

При этом комбинация теорем сложения и умножения показывает, что вероятность наступления события А в $\boldsymbol{n}$ испытаниях ровно $\boldsymbol{k}$ раз находится по формуле Бернулли:

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

где в n испытаниях события наступают k раз с вероятностью p, а не наступают $(n - k)$ раз с вероятностью q.

Пример 19

В семье пятеро детей. Найти вероятность того, что среди них два мальчика, если вероятность рождения мальчика равна 0,51.

Решение:

$n = 5$,

$k = 2$,

$p = 0{,}51 \Rightarrow q = 0{,}49$;

Найдем $P_5(2)$.

$$P_5(2) = C_5^2 \cdot 0{,}51^2 \cdot 0{,}49^3 = 10 \cdot 0{,}2601 \cdot 0{,}1176 = 0{,}306.$$

Ответ: $P_5(2) = 0{,}306$.

1.12. Наивероятнейшее число появлений события в независимых испытаниях

Встречаются задачи, в которых наибольший интерес представляет наивероятнейшее число успехов, которому соответствует наибольшая вероятность. Обозначим это число k_0.

Наивероятнейшее число k_0 определяют из двойного неравенства:

$$np - q \le k_0 \le np + p, \text{ где } p \ne 0,\ p \ne 1.$$

При этом:

а) если число $(np + p)$ – дробное, то существует одно наивероятнейшее число k_0;

б) если число $(np + p)$ – целое, то существует два наивероятнейших числа, а именно: k_0 и $k_0 + 1$.

Пример 20

Производится 20 выстрелов из орудия с вероятностью попадания в цель при каждом выстреле 0,8. Найти наивероятнейшее число попаданий в цель.

Решение:

$n = 20$;

$p = 0{,}8 \Rightarrow q = 1 - 0{,}8 = 0{,}2$;

Найдем k_0.

$$20 \cdot 0{,}8 - 0{,}2 \le k_0 \le 20 \cdot 0{,}8 + 0{,}8;$$
$$16 - 0{,}2 \le k_0 \le 16 + 0{,}8;$$
$$15{,}8 \le k_0 \le 16{,}8.$$

Ответ: $k_0 = 16$.

1.13. Локальная теорема Муавра – Лапласа

Если в схеме Бернулли число испытаний $\boldsymbol{n}$ велико, то вероятность появления события ровно $\boldsymbol{k}$ раз приближенно вычисляется по формуле

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \cdot \varphi(x) \text{ (тем точнее, чем больше } n\text{).}$$

Здесь $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-x^2/2}$, $x = \frac{k-np}{\sqrt{npq}}$, причем $\varphi(x)$ называется функцией Гаусса, а ее график – кривой Гаусса.

Для функции $\varphi(x)$ составлены таблицы значений (они находятся, как правило, в так называемых «Приложениях» книг по теории вероятностей). Пользуясь таблицей для нахождения значений функции $\varphi(x)$ [4], приведем график функции Гаусса (рис. 5).

Свойства $\varphi(x)$:

1) функция $\varphi(x)$ четная, т. е. $\varphi(-x) = \varphi(x)$;

2) при $x \ge 4$ можно считать, что $\varphi(x) = 0$.

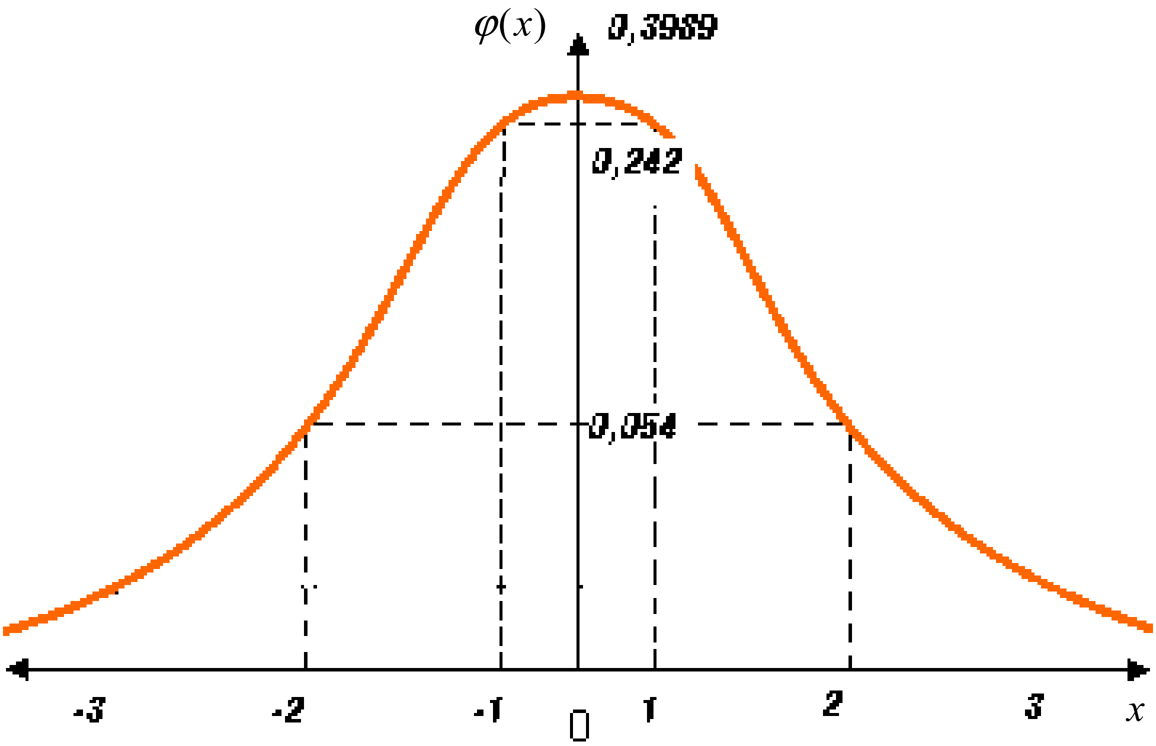


Рис. 5

Пример 21

Стрелок сделал 30 выстрелов с вероятностью попадания при одном выстреле 0,3. Найти вероятность того, что в результате будет 8 попаданий.

Решение:

$n = 30;$

$p = 0{,}3 \Rightarrow q = 0{,}7;$

$k = 6.$

$$P_{30}(8) \approx \frac{1}{\sqrt{30 \cdot 0{,}3 \cdot 0{,}7}} \cdot \varphi(x)\text{, где } x = \frac{8 - 30 \cdot 0{,}3}{\sqrt{6{,}3}} = \frac{-1}{2{,}5} = -0{,}4.$$

$\varphi(-0{,}4) = \varphi(0{,}4)$, так как функция четная.

Находим по таблице [4] значение $\varphi(0{,}4) = 0{,}3683$, тогда

$$P_{30}(8) = 0{,}3683 \cdot \frac{1}{2{,}5} = 0{,}147.$$

Ответ: $P_{30}(8) = 0{,}147$.

1.14. Интегральная теорема Муавра – Лапласа

В тех случаях, когда требуется вычислить вероятность того, что в ***n*** независимых испытаниях событие А появится не менее k_1 раз, но не более k_2, т. е. $P_n(k_1 \le k \le k_2)$ или $P_n(k_1, k_2)$, используют интегральную теорему Муавра – Лапласа:

$$P_n(k_1, k_2) \approx \frac{1}{\sqrt{2\pi}} \int_{x'}^{x''} \mathrm{e}^{-z^2/2}\, dz,$$

где $x' = \dfrac{k_1 - np}{\sqrt{npq}}, x'' = \dfrac{k_2 - np}{\sqrt{npq}}$.

Но так как неопределенный интеграл $\int \mathrm{e}^{-z^2/2} dz$ не выражается через элементарные функции, то значения функции $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \mathrm{e}^{-z^2/2}\, dz$, где значения функции $\Phi(x)$ – функции Лапласа – содержатся в специальных таблицах.

Свойства $\Phi(x)$:

1) $\Phi(0) = 0$;

2) $\Phi(-x) = -\Phi(x)$, т. е. $\Phi(x)$ – нечетная;

3) при $x \ge 3$ можно считать $\Phi(x) \approx 0{,}5$.

Пользуясь таблицей для нахождения значений функции Лапласа [4], приведем ее график (рис. 6).

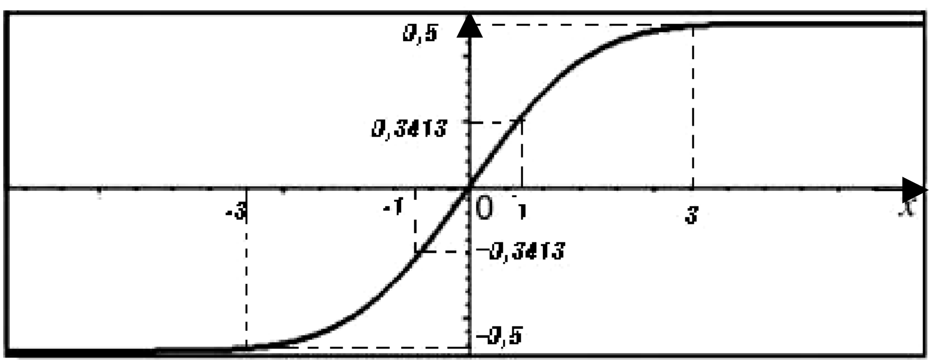


Рис. 6

Пример 22

Средний процент нарушения работы кинескопа телевизора в течение гарантийного срока равен 12. Вычислить вероятность того, что из 46 наблюдаемых телевизоров не менее 36 выдержат гарантийный срок.

Решение:

$n = 46$,

$q = 0{,}12$, т. е. средний процент нарушения работы кинескопа, тогда $p = 1 - 0{,}12 = 0{,}88$;

$36 \le k \le 46$;

Найдем $P_{46}(36, 46)$.

$$P_n(k_1, k_2) \approx \Phi(x'') - \Phi(x');$$

$$x' = \frac{k_1 - np}{\sqrt{npq}} = \frac{36 - 46 \cdot 0{,}88}{\sqrt{46 \cdot 0{,}88 \cdot 0{,}12}} = \frac{36 - 40{,}48}{\sqrt{4{,}86}} = -\frac{4{,}48}{2{,}2} = -2{,}04;$$

$$x'' = \frac{k_2 - np}{\sqrt{npq}} = \frac{46 - 46 \cdot 0{,}88}{\sqrt{46 \cdot 0{,}88 \cdot 0{,}12}} = \frac{46 - 40{,}48}{\sqrt{4{,}86}} = \frac{5{,}52}{2{,}2} = 2{,}5.$$

$\Phi(-2{,}04) = -\Phi(2{,}04) = -0{,}4793$;

$\Phi(2{,}5) = 0{,}4938$;

$P_{46}(36, 46) = 0{,}4938 - (-0{,}4793) = 0{,}4938 + 0{,}4793 = 0{,}9731$.

Ответ: $P_{46}(36, 46) = 0{,}9731$.

Пример 23

Найти вероятность того, что в партии из 800 изделий число изделий высшего сорта заключено между 600 и 700, если вероятность того, что отдельное изделие будет высшего сорта, равна 0,62.

Решение:

$n = 800$;

$p = 0{,}62$;

$q = 1 - 0{,}62 = 0{,}38$;

$600 \le k \le 700$.

$$P_n(k_1, k_2) \approx \Phi(x'') - \Phi(x');$$

$$x' = \frac{k_1 - np}{\sqrt{npq}} = \frac{600 - 800 \cdot 0{,}62}{\sqrt{800 \cdot 0{,}62 \cdot 0{,}38}} = \frac{600 - 496}{\sqrt{188{,}48}} = \frac{104}{13{,}73} = 7{,}57;$$

$$x'' = \frac{k_2 - np}{\sqrt{npq}} = \frac{700 - 800 \cdot 0{,}62}{\sqrt{800 \cdot 0{,}62 \cdot 0{,}38}} = \frac{700 - 496}{\sqrt{188{,}48}} = \frac{204}{13{,}73} = 14{,}85.$$

$$\begin{matrix} \Phi(7{,}57) = 0{,}5 \\ \Phi(14{,}85) = 0{,}5 \end{matrix} \Rightarrow P_{800}(600, 700) = 0{,}5 - 0{,}5 = 0.$$

Ответ: $P_{800}(600, 700) = 0$.

1.15. Формула Пуассона

Формула Пуассона применяется в схеме Бернулли, когда ***n*** – очень большое число, а вероятность появления события в одном испытании очень мала ($p \le 0{,}1$):

$$P_n(k) \approx \frac{\lambda^k}{k!} \cdot \mathrm{e}^{-\lambda}, \text{ где } \lambda = n \cdot p = \text{const}.$$

Замечание. В приложениях книг по теории вероятностей имеются специальные таблицы для нахождения значений функции $\Phi(x)$ [4], позволяющие найти $P_n(k)$, зная $\boldsymbol{k}$ и $\boldsymbol{\lambda}$.

Пример 24

Вероятность попадания в самолет при одном выстреле равна 0,01. Производится 100 выстрелов. Определить вероятность двух попаданий.

Решение:

$n = 100$;

$k = 2$;

$p = 0{,}01$;

$q = 0{,}99$.

В данном случае, хотя ***n*** достаточно велико, но $\sqrt{npq} \approx 1$, формула Лапласа может дать значительную погрешность. Вычислим вероятность по формуле Пуассона:

$$P_{100}(2) \approx \frac{\lambda^k}{k!} \cdot \mathrm{e}^{-\lambda} \approx \frac{1^2}{2!} \cdot \mathrm{e}^{-1} \approx \frac{1}{2\mathrm{e}} \approx 0{,}184.$$

Ответ: $P_{100}(2) \approx 0{,}184$.

1.16. Случайные величины

Под случайной величиной будем понимать величину, которая в результате испытания принимает одно и только одно из возможных значений, наперед неизвестное и зависящее от случайных величин.

Случайные величины могут быть дискретными и непрерывными. Если величина принимает отдельные, изолированные друг от друга значения, то она считается дискретной.

Например, число родившихся мальчиков среди 100 новорожденных – это случайная величина, которая может принять одно из следующих значений: 0, 1, 2, 3, …, 100.

Если же случайная величина принимает достаточно мало отличающихся друг от друга значений, то ее будем считать непрерывной. Например, расстояние, которое пролетит снаряд при выстреле из орудия, может принять одно из любых значений в заданном промежутке.

1.17. Дискретные случайные величины и их числовые характеристики

Дискретная случайная величина задается с помощью закона распределения:

x_1	x_2	…	x_n
p_1	p_2	…	p_n

где $x_1 < x_2 < \ldots < x_n$, $p_1 + p_2 + \ldots + p_n = 1$, так как при испытании хотя бы одно из значений заведомо реализуется.

Если закон распределения подчиняется какой-либо закономерности, например составлен по формуле Бернулли или по формуле Пуассона, то это распределение так и называют: распределение Бернулли, распределение Пуассона.

Пример 25

Производится три выстрела по цели с вероятностью попадания при одном выстреле 2/3. Найти закон распределения числа попаданий в цель.

Решение:

$n = 3$,

$p = \dfrac{2}{3} \Rightarrow q = \dfrac{1}{3}$. В этом случае $P_k = C_n^k p^k q^{n-k}$,

т. е. $P_0 = C_3^0 \left(\dfrac{2}{3}\right)^0 \cdot \left(\dfrac{1}{3}\right)^3 = \dfrac{1}{27}$,

$$P_1 = C_3^1 \left(\frac{2}{3}\right)^1 \cdot \left(\frac{1}{3}\right)^2 = \frac{6}{27},$$

$$P_2 = C_3^2\left(\frac{2}{3}\right)^2 \cdot \left(\frac{1}{3}\right)^1 = \frac{12}{27},$$

$$P_3 = C_3^3\left(\frac{2}{3}\right)^3 \cdot \left(\frac{1}{3}\right)^0 = \frac{8}{27}, \text{ тогда}$$

x	0	1	2	3
p	$\frac{1}{27}$	$\frac{6}{27}$	$\frac{12}{27}$	$\frac{8}{27}$

Контроль: $\frac{1}{27}+\frac{6}{27}+\frac{12}{27}+\frac{8}{27}=\frac{27}{27}=1.$

Числовые характеристики дискретной случайной величины – это математическое ожидание, дисперсия и среднее квадратическое отклонение.

Математическим ожиданием называется сумма произведений всех возможных ее значений на их вероятности:

$$M(X) = x_1 \cdot p_1 + x_2 \cdot p_2 + \ldots + x_n \cdot p_n.$$

Пример 26

x	1	3
p	0,4	0,6

Найдем $M(X)$.

$$M(X) = 1 \cdot 0{,}4 + 3 \cdot 0{,}6 = 0{,}4 + 1{,}8 = 2{,}2.$$

Ответ: $M(X) = 2{,}2$.

Встречаются такие случайные величины, у которых значения величин мало отличаются от значений математического ожидания, например

x	–0,1	0,1
p	0,5	0,5

$$M(X) = -0{,}1 \cdot 0{,}5 + 0{,}1 \cdot 0{,}5 = 0.$$

На практике же часто требуется оценить рассеяние возможных значений величины вокруг ее среднего значения, например в артиллерии важно знать, насколько кучно лягут снаряды вокруг цели, которая должна быть поражена.

Исходя из данных распределения, понятно, что использование среднего значения обычного отклонения величины от математического ожидания не является убедительным.

Поэтому для таких оценок используют среднее значение не самого отклонения, а его квадрата, которое называют дисперсией случайной величины.

Дисперсией (*рассеянием*) *дискретной случайной величины* называется математическое ожидание квадрата отклонения от ее математического ожидания:

$$D(X) = M(x - M(X))^2.$$

$(x - M(X))^2$	$(-0{,}1 - 0)^2$	$(0{,}1 - 0)^2$
p	0,5	0,5

т. е.

$(x - M(X))^2$	0,01	0,01
p	0,5	0,5

$$D(X) = 0{,}01 \cdot 0{,}5 + 0{,}01 \cdot 0{,}5 = 0{,}005 + 0{,}005 = 0{,}01.$$

Можно находить дисперсию не с помощью определения, которое часто дает громоздкие вычисления, а с помощью специальной формулы

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

$$D(X) = (-0{,}1)^2 \cdot 0{,}5 + (0{,}1)^2 \cdot 0{,}5 - 0^2 = 0{,}005 + 0{,}005 = 0{,}01.$$

Еще одной числовой характеристикой дискретной случайной величины служит *среднее квадратическое отклонение*:

$$\sigma(X) = \sqrt{D(X)},$$

характеризующее разброс случайной величины относительно центра распределения с учетом возможных значений и их вероятностей.

1.18. Непрерывные случайные величины и их числовые характеристики

В случае малоотличающихся друг от друга значений случайной величины перечислять их с помощью таблицы нецелесообразно, поэтому непрерывные случайные величины задаются функционально двумя способами:

1) с помощью функции распределения $F(x)$ (интегральная функция распределения),

2) с помощью плотности вероятности $f(x)$ (дифференциальная функция распределения).

Функцией распределения называют функцию $F(x)$, определяющую для каждого значения x вероятность того, что случайная величина X примет значение, меньшее x, т. е. $F(x) = P(X < x)$.

Часто вместо термина «функция распределения» используют термин «интегральная функция распределения».

Функция распределения обладает следующими свойствами:

1. Значения функции распределения принадлежат отрезку [0; 1]:

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

2. Функция распределения есть неубывающая функция:

$$F(x_2) \ge F(x_1), \text{ если } x_2 > x_1.$$

3. Вероятность того, что случайная величина X примет значение, заключенное в интервале (a, b), равна приращению функции распределения на этом интервале:

$$P(a < X < b) = F(b) - F(a).$$

4. $F(-\infty) = 0$; $F(\infty) = 1$.

На рис. 7 графически изображены свойства функции распределения.

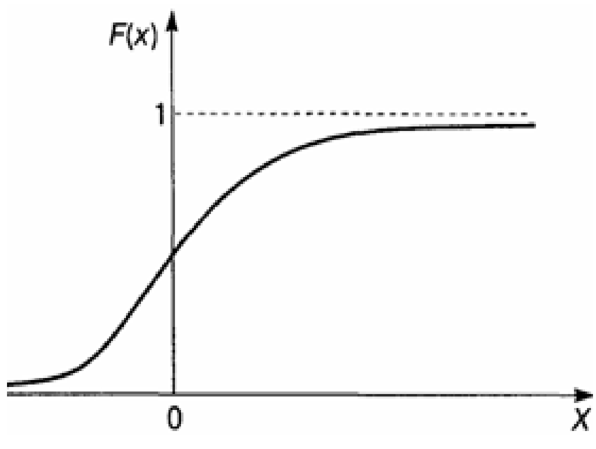


Рис. 7

Пример 27

Случайная величина X задана интегральной функцией:

$$F(x) = \begin{cases} 0, & \text{если } x \le -1, \\ \frac{1}{3}x + \frac{1}{3}, & \text{если } -1 < x \le 2, \\ 1, & \text{если } x > 2. \end{cases}$$

Найти вероятность того, что в результате испытания x примет значение, заключенное в интервале (0; 1).

Решение:

Искомая вероятность равна приращению интегральной функции на интервале (0; 1), т. е. $P(0 < X < 1) = F(1) - F(0)$, а на нем $F(x) = \frac{1}{3}x + \frac{1}{3}$, тогда

$$P(0 < X < 1) = \frac{1}{3} \cdot 1 + \frac{1}{3} - \left[\frac{1}{3} \cdot 0 + \frac{1}{3}\right] = \frac{2}{3} - \frac{1}{3} = \frac{1}{3}.$$

Ответ: $P(0 < X < 1) = \dfrac{1}{3}$.

Функция распределения $F(x)$ имеет смысл и для дискретных случайных величин.

Пример 28

Дискретная случайная величина задана законом распределения:

x	2	6	10
p	0,5	0,4	0,1

Построить график интегральной функции для этой случайной величины.

Решение:

а) прежде, чем строить график такой функции, представим ее аналитически:

если $x \le 2$, то $F(x) = 0$,

если $2 < x \le 6$, то $F(x) = 0{,}5$ (т. е. x может принять значение 2 с вероятностью 0,5),

если $6 < x \le 10$, то $F(x) = 0{,}9$ (т. е. x может принять значение или 2 (с вероятностью 0,5), или 6 (с вероятностью 0,4), а так как события несовместны $\Rightarrow 6 < x \le 10$ принимает значения с вероятностью $0{,}4 + 0{,}5 = 0{,}9$).

Если $x > 0$, то $F(x) = 1$, тогда

$$F(x) = \begin{cases} 0 & \text{при } x \le 2, \\ 0{,}5 & \text{при } 2 < x \le 6, \\ 0{,}9 & \text{при } 6 < x \le 10, \\ 1 & \text{при } x > 10, \end{cases}$$

т. е. это $F(x)$, представленная аналитически для дискретной случайной величины;

б) теперь построим ее график функции (рис. 8).

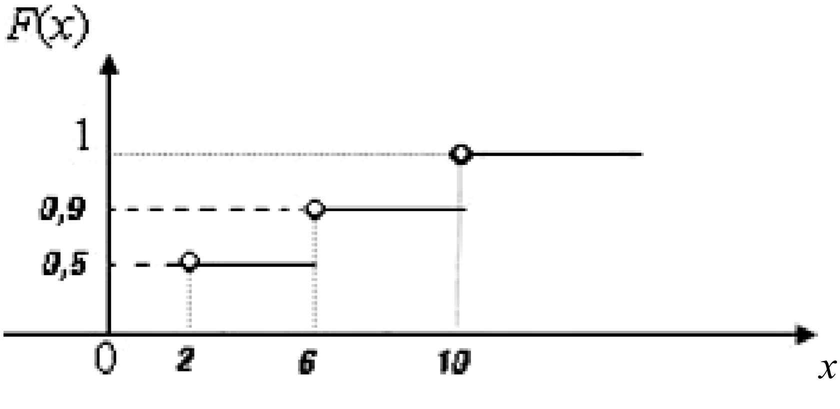


Рис. 8

Кружок означает, что в этом месте отсутствует точка на графике.

Дифференциальная функция $f(x)$ (плотность вероятности) является одной из форм закона распределения случайной величины, которая существует только для непрерывных случайных величин.

На рис. 9 приведен график дифференциальной функции распределения.

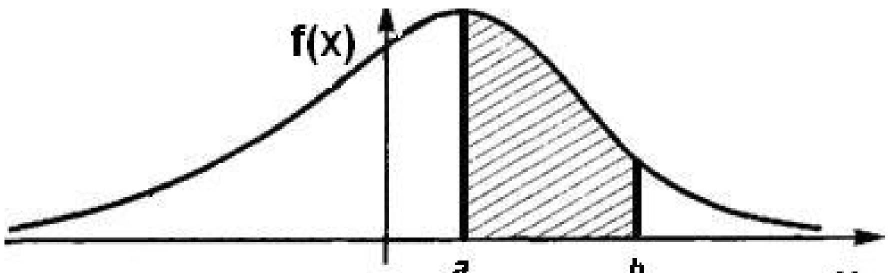


Рис. 9

Дифференциальная функция обладает следующими свойствами:

1) $f(x)$ неотрицательная, т. е. $f(x) \geq 0$,

2) вероятность попадания непрерывной случайной величины в промежуток $[a; b]$ равна определенному интегралу от ее плотности в пределах от a до b, т. е. $P(a \leq X \leq b) = \int_a^b f(x)\,dx = S$ (площадь заштрихованной фигуры),

3) площадь бесконечной фигуры, ограниченной графиком плотности $f(x)$ и осью абсцисс, равна 1.

Связь между $f(x)$ и $F(x)$:

1) $F(x) = \int\limits_{-\infty}^{x} f(x)\,dx$;

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

Рассмотрим данную связь на примерах.

Пример 29

Случайная величина X задана интегральной функцией:

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

Найти дифференциальную функцию распределения.

Решение:

Так как $f(x) = (F(x))'$, то

$$f(x) = \begin{cases} 0, & \text{если } x \le 0, \\ 1, & \text{если } 0 < x \le 1, \\ 0, & \text{если } x > 1. \end{cases}$$

Пример 30

Случайная величина X задана дифференциальной функцией распределения:

$$f(x) = \begin{cases} 0, & \text{если } x \le 0, \\ \frac{1}{2}\sin x, & \text{если } 0 < x \le \pi, \\ 0, & \text{если } x > \pi. \end{cases}$$

Найти: а) интегральную функцию распределения;

б) вероятность того, что в результате испытания случайная величина примет значение, заключенное в интервале (0; π/4).

Решение:

а) так как $f(x) = (F(x))'$, $\Rightarrow F(x) = \int\limits_{-\infty}^{x} f(x)\,dx$, тогда если:

1) $x < 0$, то $f(x) = 0$, $\Rightarrow F(x) = 0$;

2) $0 < x < \pi$, $\Rightarrow$ $F(x) = \int\limits_{-\infty}^{x} f(x)\,dx = \int\limits_{-\infty}^{x} \frac{1}{2}\sin x\,dx = \int\limits_{-\infty}^{0} 0\,dx + \int\limits_{0}^{x} \frac{1}{2}\sin x\,dx =$

$$= -\frac{1}{2}\cos x - \left[-\frac{1}{2}\cos 0\right] = \frac{1}{2} - \frac{1}{2}\cos x;$$

3) $x > \pi$, то $F(x) = \int\limits_{-\infty}^{0} 0\,dx + \int\limits_{0}^{\pi} \frac{1}{2}\sin x\,dx + \int\limits_{\pi}^{x} 0\,dx = -\frac{1}{2}\cos x\Big|_0^{\pi} =$

$$= -\frac{1}{2}\cos\pi - \left[-\frac{1}{2}\cos 0\right] = -\frac{1}{2}\cdot(-1) + \frac{1}{2}\cdot 1 = \frac{1}{2} + \frac{1}{2} = 1.$$

Интегральная функция распределения выглядит аналитически следующим образом:

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

б) $P\left(0 < X < \frac{\pi}{4}\right) = \frac{1}{2}\int\limits_{0}^{\pi/4} \sin x\,dx = -\frac{1}{2}\cos x\Big|_0^{\pi/4} = -\frac{1}{2}\cos\frac{\pi}{4} - \left(-\frac{1}{2}\cos 0\right) =$

$$= -\frac{1}{2}\cdot\frac{\sqrt{2}}{2} + \frac{1}{2} = -\frac{\sqrt{2}}{4} + \frac{1}{2} = \frac{2-\sqrt{2}}{4}.$$

Ответ: а) $F(x) = \begin{cases} 0, & \text{если } x \le 0, \\ \frac{1}{2}(1-\cos x), & \text{если } 0 < x \le \pi, \\ 1, & \text{если } x > \pi; \end{cases}$

б) $P\left(0 < X < \frac{\pi}{4}\right) = \frac{2-\sqrt{2}}{4}.$

Числовые характеристики непрерывной случайной величины имеют такой же смысл и сохраняют все свои свойства, как и в дискретном случае.

Находятся они по следующим формулам:

$M(X) = \int_{-\infty}^{+\infty} x\cdot f(x)\,dx$ – математическое ожидание непрерывной случайной величины;

$D(X) = \int_{-\infty}^{+\infty} [x - M(X)]^2\cdot f(x)\,dx$ – дисперсия непрерывной случайной величины, найденная по определению;

$D(X) = \int_{-\infty}^{+\infty} x^2\cdot f(x)\,dx - [M(x)]^2$ – специальная формула для вычисления дисперсии непрерывной случайной величины;

$\sigma(X) = \sqrt{D(X)}$ – среднее квадратическое отклонение непрерывной случайной величины, если возможные значения принадлежат всей числовой оси.

Пример 31

Найти $M(X)$, $D(X)$, $\sigma(X)$ случайной величины X, заданной интегральной функцией распределения:

$$F(x) = \begin{cases} 0, & \text{если } x \le 0, \\ x^2, & \text{если } 0 < x \le 1, \\ 1, & \text{если } x > 1; \end{cases}$$

Решение:

Найдем дифференциальную функцию распределения

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

Так как интервал (0; 1) задан по условию задачи, тогда

$$M(X) = \int_0^1 2x \cdot x\, dx = \int_0^1 2x^2\, dx = \left.\frac{2x^3}{3}\right|_0^1 = \frac{2}{3}.$$

$$D(X) = \int_0^1 [x - M(x)]^2 f(x)\, dx = \int_0^1 \left(x - \frac{2}{3}\right)^2 \cdot 2x\, dx = \int_0^1 \left(x^2 - \frac{4}{3}x + \frac{4}{9}\right) \cdot 2x\, dx =$$

$$= \int_0^1 \left(2x^3 - \frac{8}{3}x^2 + \frac{8}{9}x\right) dx = \left.\left(\frac{2}{4}x^4 - \frac{8}{3}\cdot\frac{x^3}{3} + \frac{8}{9}\cdot\frac{x^2}{2}\right)\right|_0^1 = \left.\left(\frac{x^4}{2} - \frac{8}{9}x^3 + \frac{4}{9}x^2\right)\right|_0^1 =$$

$$= \frac{1}{2} - \frac{8}{9} + \frac{4}{9} = \frac{9 - 16 + 8}{18} = \frac{1}{18} \approx 0,06.$$

$$\sigma(X) = \sqrt{0,06} \approx 0,244.$$

Ответ: $M(X) = 2/3$, $D(X) \approx 0,06$, $\sigma(X) \approx 0,244$.

1.19. Основные законы распределения

Перечислим законы распределения:

1. Биномиальный закон (закон Бернулли).
2. Закон Пуассона.
3. Нормальный закон (закон Гаусса).
4. Равномерный закон.
5. Показательный закон.

1. Биномиальный закон распределения дискретной случайной величины.

Дискретная случайная величина X имеет биномиальное распределение (или подчинена биномиальному закону распределения) с параметрами (n, p), если она принимает значения 0, 1, 2, …, n и вероятности этих значений находятся по формуле

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

где $k = 0, 1, 2, \ldots, n$; $q = 1 - p$; $p \in (0; 1)$.

Приведем числовые характеристики этого распределения:

$$M(X) = n \cdot p, \quad D(X) = n \cdot p \cdot q, \; \sigma(X) = \sqrt{n \cdot p \cdot q}\,.$$

См. пример 25.

2. Закон Пуассона. Дискретная случайная величина X имеет распределение Пуассона с параметром $\lambda = M(X)$, если она принимает любые целые неотрицательные значения и вероятности значений находятся по формуле

$$P_k = \frac{\lambda^k \mathrm{e}^{-\lambda}}{k!},$$

где $k = 0, 1, 2, \ldots$; $\lambda = n \cdot p$.

Приведем значения числовых характеристик $M(X) = \lambda$, $D(X) = \lambda$, $\sigma(X) = \sqrt{\lambda}$, если $n \to \infty$, а $p \to 0$, тогда λ сколь угодно мала, что встречается нечасто. Поэтому закон Пуассона в теории вероятностей носит название закона *редких явлений*.

3. Нормальный закон. Непрерывная случайная величина X имеет нормальное распределение с параметрами a, σ, если ее плотность вероятности задается формулой

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{\frac{(x-a)^2}{2\sigma^2}}.$$

Приведем числовые характеристики этого распределения

$$M(X) = a, \quad D(X) = \sigma^2, \quad \sigma(X) = \sqrt{D(X)}, \quad \text{тогда } F(x) = \frac{1}{2} + \Phi\left(\frac{x-a}{\sigma}\right),$$

где $\Phi(x)$ – функция Лапласа.

Вероятность попадания случайной величины X в заданный интервал $(\alpha; \beta)$ находится следующим образом:

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

Вероятность заданного отклонения от математического ожидания находится как

$$P(|x - a| < \varepsilon) = 2\Phi\left(\frac{\varepsilon}{\sigma}\right).$$

Значения нормальной случайной величины X с практической достоверностью лежат в интервале ($a \pm 3\sigma$). Это утверждение носит название «правило трех сигм» (рис. 10):

$$P(|x - a| < 3\sigma) = 2\Phi\left(\frac{3\sigma}{\sigma}\right) = 2\Phi(3) \approx 1.$$

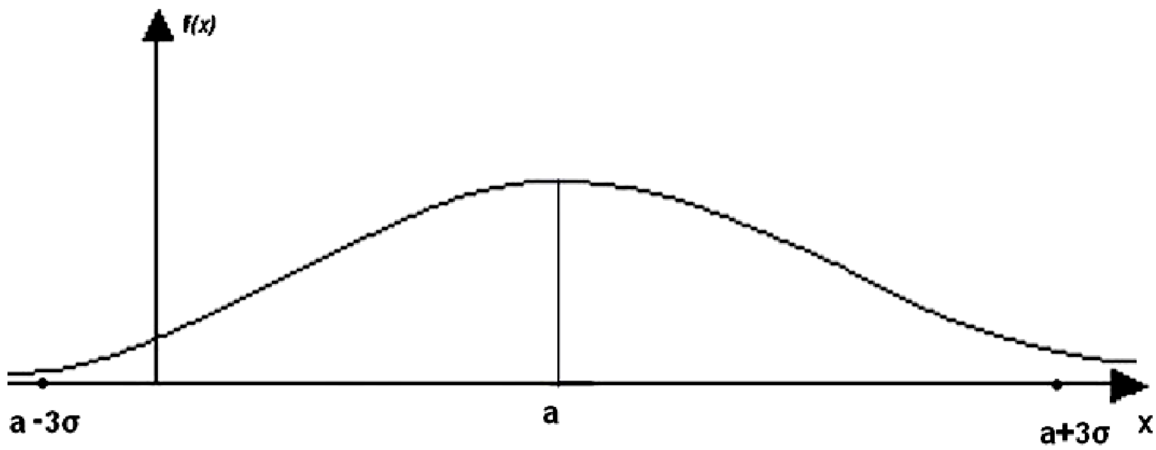


Рис. 10

Пример 32

Длина изготовляемой автоматом детали представляет собой случайную величину, распределенную по нормальному закону с параметрами

$M(X)$ = 15 см, $\sigma(X)$ = 0,2 см. Найти: 1) вероятность брака, если допускаемые размеры детали должны быть (15 ± 0,3) см; 2) какую точность детали можно гарантировать с вероятностью 0,97?

Решение:

1) $P(|x - 15| < 0{,}3) = 2\Phi(0{,}3/0{,}2) = 2\Phi(1{,}5) = 2 \cdot 0{,}4332 = 0{,}8662$.

Вероятность брака: $q = 1 - 0{,}8664 = 0{,}1336$.

2) $P(|x - a| < \varepsilon) = 2\Phi(\varepsilon/\sigma) = 0{,}97$ (по условию) $\Rightarrow \Phi(\varepsilon/\sigma) = 0{,}97/2 = 0{,}485$, тогда $\varepsilon/\sigma = 2{,}17$ – по таблице значений функции Лапласа $\Phi(x)$ [4] нашли значение аргумента функции.

$\varepsilon = 2{,}17 \cdot \sigma = 2{,}17 \cdot 0{,}2 = 0{,}434$ см.

Ответ: с вероятностью 0,97 размеры изготовляемой детали будут составлять (15 ± 0,434) см.

4. Равномерный закон. Непрерывная случайная величина X имеет равномерное распределение на отрезке $[a; b]$, если ее плотность вероятности задается формулой

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

График функции $f(x)$ представлен на рис. 11.

$$F(x) = \begin{cases} 0, & \text{если } x < a, \\ \dfrac{x-a}{b-a}, & \text{если } a \le x \le b, \\ 1, & \text{если } x > b. \end{cases}$$

График интегральной функции $F(x)$ приведен на рис. 12.

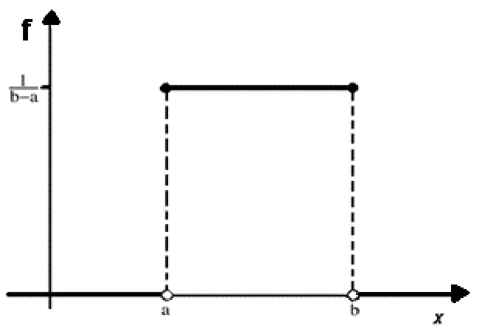


Рис. 11

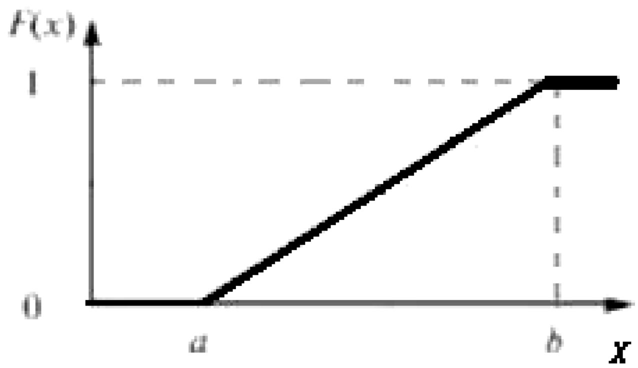


Рис. 12

$$M(X)=\frac{a+b}{2},\quad D(X)=\frac{(b-a)^2}{12},\quad \sigma(X)=\frac{b-a}{\sqrt{12}}.$$

Пример 33

Цена деления шкалы амперметра равна 0,1 А. Показания округляются до ближайшего целого деления. Найти вероятность того, что при отсчете будет сделана ошибка, превышающая 0,02 А.

Решение:

Ошибка округления отсчета – случайная величина X, которая распределена равномерно в интервале между двумя соседними целыми делениями. Плотность равномерного распределения $f(x) = 1/(b - a)$, где $(b - a)$ – длина интервала, в котором заключены возможные значения X, вне интервала $f(x) = 0$.

У нас $(b - a) = 0{,}1, \Rightarrow f(x) = \dfrac{1}{0{,}1} = 10.$

Таким образом, если величина будет заключена в интервале (0,02; 0,08), то $P(a < X < b) = \int_a^b f(x)\,dx,$ тогда

$$P(0{,}02 < X < 0{,}08) = \int_{0{,}02}^{0{,}08} 10\,dx = 10x\Big|_{0{,}02}^{0{,}08} = 0{,}8 - 0{,}2 = 0{,}6.$$

Ответ: при отсчете будет сделана ошибка, превышающая 0,02 А с вероятностью 0,6.

5. Показательный (экспоненциальный) закон распределения. Непрерывная случайная величина X имеет показательный (или экспоненциальный) закон распределения, если ее плотность вероятности имеет вид

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

где λ – параметр распределения.

График функции $f(x)$ представлен на рис. 13.

Функция распределения показательного распределения имеет вид

$$F(x)=\begin{cases}1-e^{-\lambda x} & \text{при } x\ge 0,\\ 0 & \text{при } x<0.\end{cases}$$

График функции $F(x)$ приведен на рис. 14.

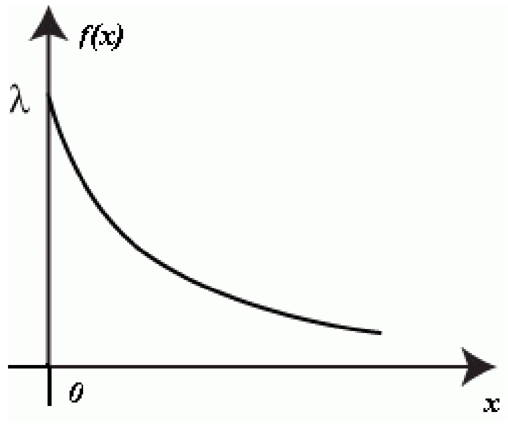


Рис. 13

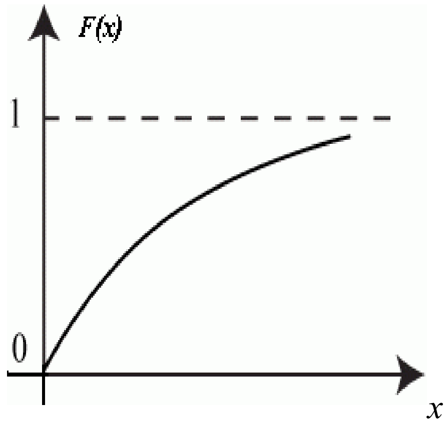


Рис. 14

Вероятность попадания в интервал $(a; b)$ непрерывной случайной величины X, распределенной по показательному закону, находится по формуле

$$P(a < X < b) = e^{-\lambda a} - e^{-\lambda b}.$$

Числовые характеристики: $M(X) = \frac{1}{\lambda}$; $D(X) = \frac{1}{\lambda^2}$; $\sigma(X) = \frac{1}{\lambda}$.

Пример 34

Найти дисперсию и среднее квадратическое отклонение показательного распределения, заданного плотностью вероятности:

$$f(x) = \begin{cases} 10e^{-10x}, & \text{если } x \geq 0, \\ 0, & \text{если } x < 0. \end{cases}$$

Решение:

$D(X) = \frac{1}{\lambda^2}$; $\sigma(X) = \frac{1}{\lambda}$, а так как $\lambda = 10$, $\Rightarrow$

$D(X) = \frac{1}{100} = 0{,}01$, $\sigma(X) = \frac{1}{10} = 0{,}1$.

Ответ: $D(X) = 0{,}01$; $\sigma(X) = 0{,}1$.

1.20. Простейший поток событий

Формула Пуассона и показательный закон распределения имеют продолжение, именуемое пуассоновским потоком событий.

Потоком событий будем называть последовательность событий A_1, A_2, …, A_m, …, которые наступают в случайные моменты времени.

Пуассоновский поток является простейшим, если:

1) вероятность появления k событий за промежуток времени t есть функция, зависящая только от k и t;

2) вероятность появления k событий в любом промежутке времени не зависит от того, появились ли события в моменты времени, предшествующие началу рассматриваемого промежутка;

3) вероятность появления двух или более событий на элементарном участке Δt сравнительно мала по сравнению с вероятностью появления одного события.

Таким образом, интенсивность потока событий λ приобретает в качестве сомножителя время t, и формула Пуассона примет вид:

$$P_k(t) = \frac{(\lambda t)^k}{k!} = \mathrm{e}^{-\lambda t},$$

где λ – интенсивность потока, т. е. среднее число событий, которые появляются в единицу времени.

Пример 35

Среднее число вызовов, поступающих на АТС в одну минуту, равно 2. Найти вероятность того, что за 4 минуты поступит меньше 3 вызовов.

Решение:

$\lambda = 2$,

$t = 4$, $\Rightarrow \lambda t = 8$, тогда P_4 $(k < 3)$ будет равна

$$P_4(k<3) = \frac{8^2}{2!}\mathrm{e}^{-8} + \frac{8^1}{1!}\mathrm{e}^{-8} + \frac{8^0}{0!}\mathrm{e}^{-8} \approx 0{,}0137.$$

Ответ: P_4 $(k < 3) = 0{,}0137$.

2. МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

2.1. Теория оценок

Математическая статистика – наука о способах получения выводов из данных опыта.

Пусть существует совокупность однородных объектов, например 1000 деталей.

Иногда обследуют каждый объект совокупности относительно признака. Но если исходная совокупность, называемая генеральной совокупностью, содержит большое число объектов, то провести сплошное обследование бывает невозможно.

Тогда из всей совокупности отбирают ограниченное число объектов и подвергают их изучению. Такая совокупность называется выборочной или просто *выборкой*.

Если из 1000 деталей отобрано для исследования 100, то эти 100 объектов и есть выборка.

Объем совокупности – это число объектов совокупности.

Например, объем генеральной совокупности: $N = 1000$, тогда объем выборки: $n = 100$.

Замечание. Часто генеральная совокупность содержит конечное, но достаточно большое число объектов. Тогда для обеспечения теоретических выводов допускают, что генеральная совокупность состоит из бесчисленного множества объектов, ведь увеличение объема генеральной совокупности практически не сказывается на результатах обработки данных.

Таким образом, если элементы выборки располагаются в порядке возрастания, т. е. $x_1, x_2, \ldots, x_n$, то она называется вариационным рядом, где $x_1, x_2, \ldots$ – варианты ряда.

Число $R = x_{\max} - x_{\min}$ называется размахом варьирования выборки.

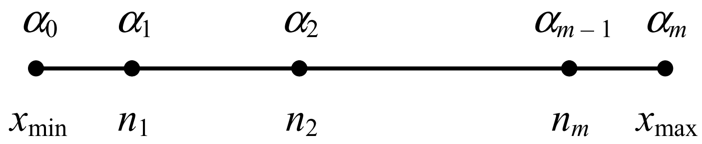
Медианой называется число: $m_e = x_{k+1}$, если вариационный ряд имеет нечетное число вариант; $m_e = (x_k + x_{k+1})/2$, если вариационный ряд имеет четное число вариант.

Пример 1

а) дан вариационный ряд 2 3 5 6 7, тогда $m_e = 5$; $R = 7 - 2 = 5$;

б) дан вариационный ряд 2 3 5 6 7 9, тогда $m_e = (5 + 6)/2 = 5{,}5$; $R = 9 - 2 = 7$.

Если разделить отрезок



на некоторое число интервалов одинаковой длины (R / m) и подсчитать число элементов выборки, попадающих в каждый интервал (n_1, n_2, …, n_m), тогда $n_1 + n_2 + \ldots + n_m = n$.

Исходя из всего сказанного, числа n_1, n_2, …, n_m называются частотами попадания в интервал, тогда статистическое выборочное распределение случайной величины X представлено следующим образом:

$$\begin{pmatrix} \alpha_0 - \alpha_1 & \alpha_1 - \alpha_2 & \ldots & \alpha_{n-1} - \alpha_m \\ n_1/n & n_2/n & \ldots & n_m/n \end{pmatrix},$$

т. е. это перечень наблюденных значений величины X и их относительных частот:

$$\begin{pmatrix} x_1 & x_2 & \ldots & x_m \\ n_1/n & n_2/n & \ldots & n_m/n \end{pmatrix}.$$

Модой M_0 называют варианту, которая имеет наибольшую частоту.

Например, дано статистическое выборочное распределение

$$\begin{matrix} 1 & 4 & 7 & 9 \\ 5 & 1 & 20 & 6 \end{matrix} \Rightarrow M_0 = 7,$$

так как максимальная частота равна 20.

Замечание. Сумма относительных частот статистического распределения равна 1, т. е.

$$\frac{n_1}{n} + \frac{n_2}{n} + \ldots + \frac{n_m}{n} = \frac{n_1 + n_2 + \ldots + n_m}{n} = \frac{n}{n} = 1.$$

Пример 2

Результаты 20 наблюдений над количественным признаком генеральной совокупности представлены следующим образом:

$$\begin{matrix} \text{значение признака, } x_i \\ \text{частота, } n_i \end{matrix} \begin{pmatrix} 2 & 6 & 10 & 12 & 14 \\ 1 & 5 & 7 & 3 & 4 \end{pmatrix}.$$

Найти статистическое распределение выборки.

Решение:

Так как $n = 1 + 5 + 7 + 3 + 4 = 20$, то

$$\begin{matrix} x_i \\ w_i \end{matrix} \begin{pmatrix} 2 & 6 & 10 & 12 & 14 \\ \dfrac{1}{20} & \dfrac{5}{20} & \dfrac{7}{20} & \dfrac{3}{20} & \dfrac{4}{20} \end{pmatrix}.$$

Контроль: сумма относительных частот равна 1.

2.1.1. Полигон относительных частот статистического распределения

Иногда удобнее построить статистическое распределение в виде графика, где на оси абсцисс откладываются значения вариант, а на оси ординат – относительные частоты.

Пример 3

$$\begin{matrix} x_i \\ w_i \end{matrix} \begin{pmatrix} 1,5 & 3,5 & 5,5 & 7,5 \\ 0,1 & 0,2 & 0,4 & 0,3 \end{pmatrix}.$$

Построить полигон относительных частот.

Решение:

Полигон относительных частот представлен на рис. 15.

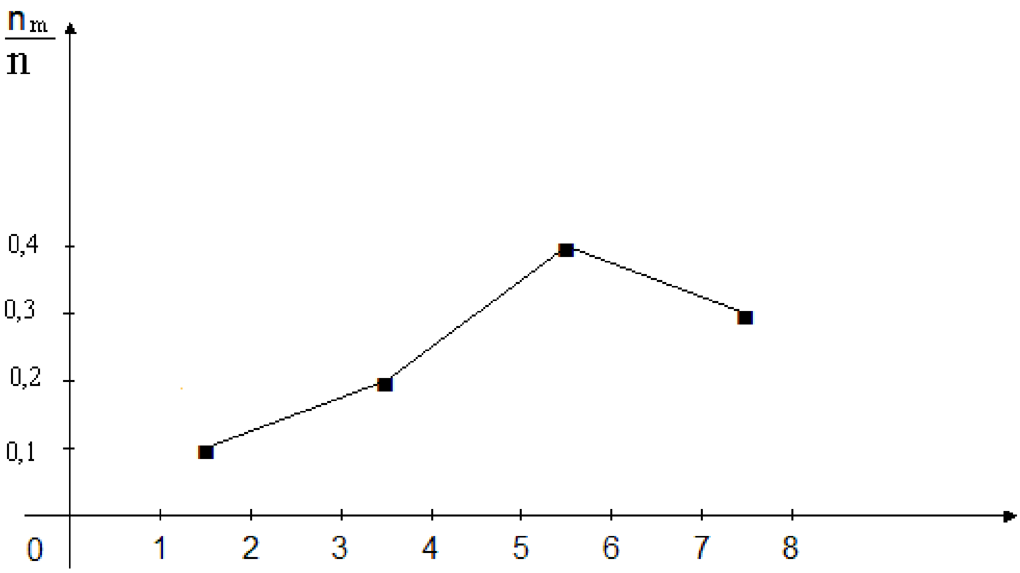


Рис. 15

Замечание. Если на оси ординат откладываются не относительные частоты, а просто частоты наблюдаемых значений, то график называется полигоном частот.

Если признак распределен непрерывно, то целесообразнее строить гистограмму относительных частот (или просто частот), которая представляет собой выборочный аналог плотности вероятности случайной величины.

В этом случае за единицу масштаба на оси абсцисс примем длину интервала $h = R / m$. На рис. 16 представлена гистограмма относительных частот.

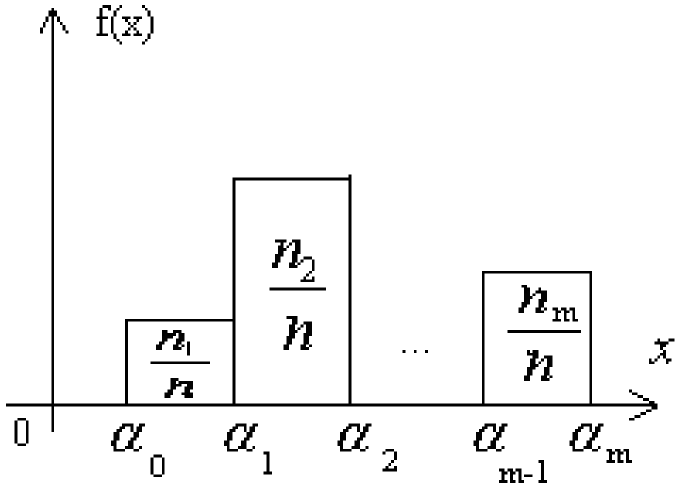


Рис. 16

Площадь построенной ступенчатой фигуры равна 1.

Пример 4

$h = 5$	h_m	h_m / n
5–10	4	0,8 = 4/5
10–15	6	1,2 = 6/5
15–20	16	3,2 = 16/5
20–25	36	7,2 = 36/5
25–30	24	4,8 = 24/5
30–35	10	2 = 10/5
35–40	4	0,8 = 4/5

Построить гистограмму относительных частот.

Решение:
На рис. 17 изображена гистограмма относительных частот.

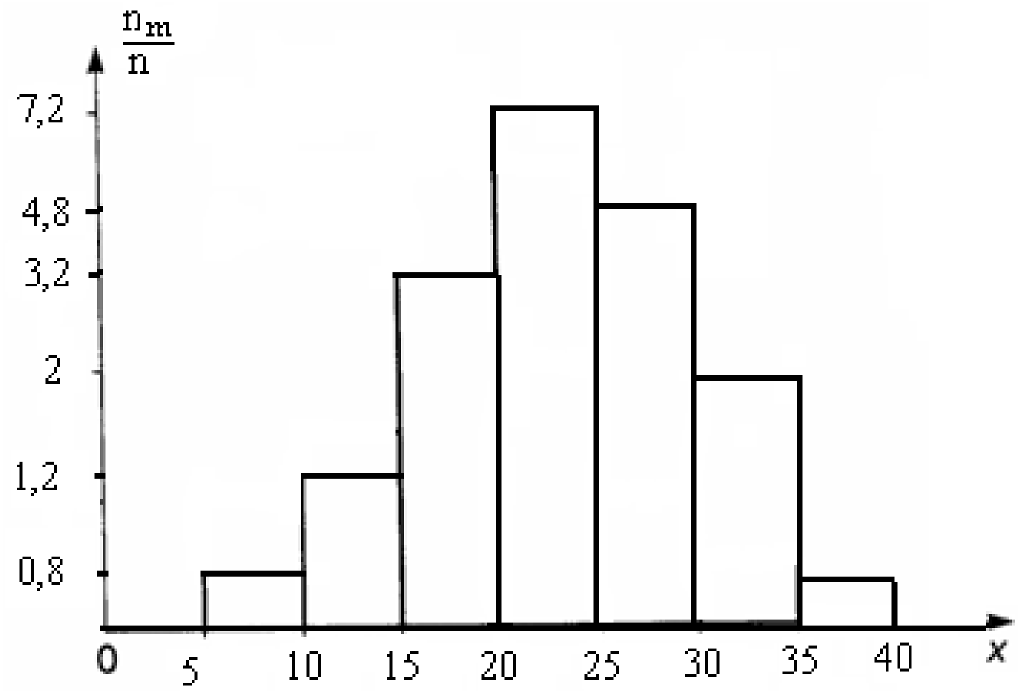


Рис. 17

Площадь построенной ступенчатой фигуры равна 1.

2.1.2. Числовые характеристики выборки. Выборочная средняя, выборочная дисперсия

Выборочная средняя

Определение. Выборочной средней $\overline{x}$ называется среднее арифметическое значений признака выборочной совокупности, т. е. если значения $x_1, x_2, \ldots, x_n$ различны, то

$$\overline{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}.$$

Если же значения признака имеют соответствующие частоты, причем $n_1, n_2, \ldots, n_k$ в сумме равны n, т. е.

$$\begin{matrix} x_1 & x_2 & \ldots & x_k \\ n_1 & n_2 & \ldots & n_k \end{matrix}, \quad \text{тогда} \quad \overline{x} = \frac{x_1 \cdot n_1 + x_2 \cdot n_2 + \ldots + x_k \cdot n_k}{n}.$$

Пример 5

Найти среднюю по данным выборки:

$$\begin{array}{ccccc} x_i & 1 & 2 & 3 & 4 \\ n_i & 20 & 15 & 10 & 5 \end{array} \quad \Rightarrow \quad \overline{x} = \frac{20\cdot 1 + 15\cdot 2 + 10\cdot 3 + 5\cdot 4}{20+15+10+5} = \frac{100}{50} = 2.$$

Ответ: $\overline{x} = 2$.

Выборочная дисперсия

Определение. Выборочной дисперсией D называют среднее арифметическое квадратов отклонений наблюдаемых значений признака от их среднего значения.

Если имеется перечень значений признака, то $D = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$.

Если даны соответствующие частоты значений признака, т. е.

$$\begin{array}{cccc} x_1 & x_2 & \dots & x_k \\ n_1 & n_2 & \dots & n_k \end{array}, \quad \text{то} \quad D = \sum_{i=1}^{k} n_i (x_i - \overline{x})^2.$$

Для вычисления дисперсии можно пользоваться формулой $D = \overline{x^2} - (\overline{x})^2$.

Пример 6

Найти дисперсию данного распределения: $\begin{array}{ccccc} x_i & 1 & 2 & 3 & 4 \\ n_i & 20 & 15 & 10 & 5 \end{array}$.

Решение:

$$\overline{x} = \frac{20\cdot 1 + 15\cdot 2 + 10\cdot 3 + 5\cdot 4}{20+15+10+5} = \frac{100}{50} = 2.$$

$$\overline{x^2} = \frac{20\cdot 1^2 + 15\cdot 2^2 + 10\cdot 3^2 + 5\cdot 4^2}{50} = 5, \text{ тогда } D = \overline{x^2} - (\overline{x})^2 = 5 - 2^2 = 1.$$

Ответ: $D = 1$.

Среднее квадратическое отклонение

Эта характеристика имеет тот же смысл, что и в теории вероятностей, т. е. $\sigma = \sqrt{D}$.

Кроме перечисленных характеристик, существуют также выборочный коэффициент асимметрии A^* и коэффициент эксцесса E^*, где

$$A^* = \frac{1}{\sigma^3}\sum_{i=1}^{k}(x_i - \overline{x})^3 P_i, \quad E^* = \frac{1}{\sigma^4}\sum_{i=1}^{k}(x_i - \overline{x})^4 P_i - 3.$$

2.2. Статистическая проверка гипотез

Пример 7

1. Исходные данные.

Дана выборка из генеральной совокупности случайной величины X.

Данные представлены в табл. 1.

Таблица 1

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

Выборка содержит 51 наблюдаемое значение, поэтому выборка имеет объем $n = 51$.

2. Построение вариационного ряда.

Операция расположения значений случайной величины по неубыванию называется ранжированием. Последовательность элементов $x^{(1)} \le x^{(2)} \le \le \ldots \le x^{(k)}$ называется вариационным рядом, элементы которого называют вариантами.

Проранжировав статистические данные, получаем вариационный ряд (табл. 2).

Таблица 2

№		№		№		№		№	
1	0	11	2	21	2	31	3	41	4
2	0	12	2	22	2	32	3	42	4
3	0	13	2	23	2	33	3	43	4
4	1	14	2	24	3	34	3	44	4
5	1	15	2	25	3	35	3	45	4
6	1	16	2	26	3	36	3	46	4
7	1	17	2	27	3	37	3	47	4
8	1	18	2	28	3	38	3	48	4
9	1	19	2	29	3	39	3	49	5
10	1	20	2	30	3	40	4	50	5
								51	6

3. Построение интервального вариационного ряда.

Опытные данные объединим в группы так, чтобы в каждой отдельной группе значения вариант были одинаковы, и тогда можно определить число, показывающее, сколько раз встречается соответствующая варианта в определенной (соответствующей) группе.

Численность отдельной группы сгруппированного ряда опытных данных называется выборочной частотой соответствующей варианты $x^{(i)}$ и обозначается n_i. При этом $\sum_{i=1}^{k} n_i = n$, где n – объем выборки.

Отношение выборочной частоты данной варианты к объему выборки называется относительной выборочной частотой P_i^*, т. е. $P_i^* = n_i / n$, где индекс i – номер варианты.

Согласно теореме Бернулли имеем $P_i^* \to P_i$ при ($n \to \infty$), т. е. $\sum_{i=1}^{n} P_i = 1$, $\Rightarrow \sum_{i=1}^{n} P_i^* = 1$.

Интервальным вариационным рядом распределения называется упорядоченная совокупность частичных интервалов значений случайной величины с соответствующими им частотами или относительными частотами.

Для построения интервального вариационного ряда выполняем следующие действия:

1) находим размах варьирования $R = x_{\max} - x_{\min}$. Имеем $R = 6 - 0$;

2) определяем длину частичного интервала Δ – шаг разбиения по формуле Стерджеса: $\Delta \approx R / (1 + 3{,}322 \cdot \lg n) = R / K$, где n – объем выборки, K – число частичных интервалов. Так как $n = 51$, то $K = 1 + 3{,}322 \cdot \lg 51 \approx 7$, $\Delta \approx 6/7 \approx 1$;

3) определяем начало первого частичного интервала $x_{нач} = x_{\min} - \Delta/2 = 0 - 1/2$. Выбираем $x_{нач} = -0{,}5$.

После разбиения на частичные интервалы просматриваем ранжированную выборку и определяем, сколько значений признака попало в каждый частичный интервал, включая в него те значения, которые больше либо равны нижней границы и меньше верхней границы. Строим интервальный вариационный ряд (табл. 3).

Таблица 3

$[x_i; x_{i+1})$	[–0,5; 0,5)	[0,5; 1,5)	[1,5; 2,5)	[2,5; 3,5)	[3,5; 4,5)	[4,5; 5,5)	[5,5; 6,5)
n_i	3	7	13	16	9	8	1

4. Построение гистограммы.

Гистограммой частот (частостей) называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длины Δ, а высоты равны отношению n_i / Δ – плотность частоты (или P_i^* / Δ – плотность частости).

Для построения гистограммы строим вспомогательную табл. 4.

Таблица 4

i	Разряды $[x_i; x_{i+1})$	n_i	$P_i^* = \dfrac{n_i}{n}$	$f^*(x) = h^* = \dfrac{P_i^*}{\Delta}$	$\overline{x}_i = \dfrac{x_i + x_{i+1}}{2}$
1	[–0,5; 0,5)	3	0,0588	0,0588	0
2	[0,5; 1,5)	7	0,1373	0,0588	1
3	[1,5; 2,5)	13	0,2549	0,2549	2
4	[2,5; 3,5)	16	0,3137	0,3137	3
5	[3,5; 4,5)	9	0,1765	0,1765	4
6	[4,5; 5,5)	2	0,0392	0,0392	5
7	[5,5; 6,5)	1	0,0196	0,0196	6
Контроль		51	1		

По данным табл. 4 строим гистограмму частостей (рис. 18).

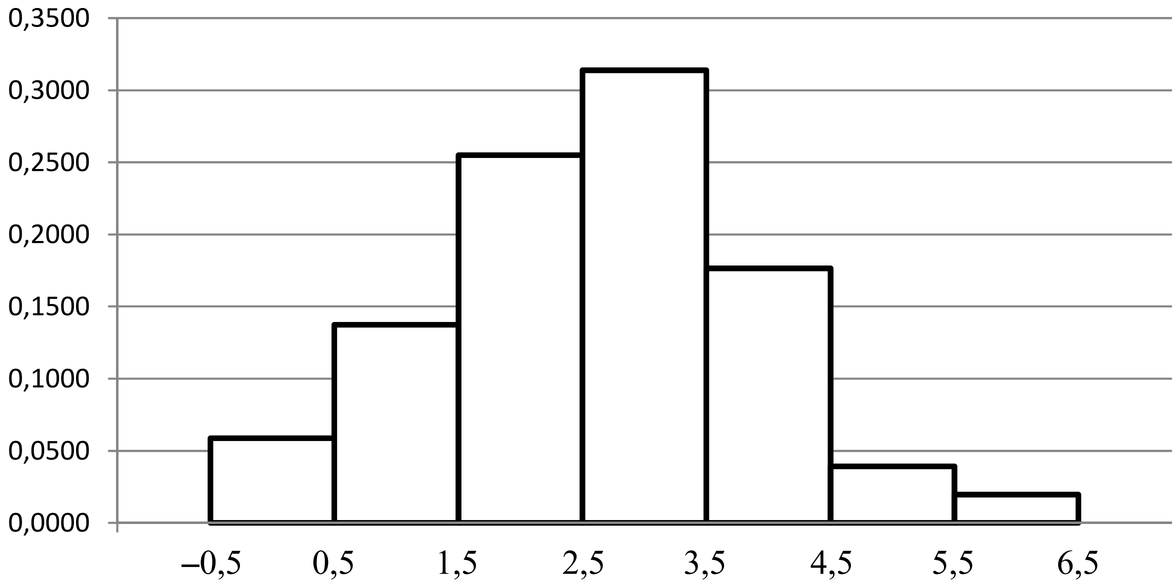


Рис. 18

Гистограмма частостей является статистическим аналогом дифференциальной функции распределения (плотности) $f(x)$ случайной величины X. Площадь гистограммы частостей равна 1.

5. Нахождение числовых характеристик выборки.

Рассчитаем статистическое среднее по формуле

$$x_{\text{В}} = \sum_{i=1}^{k} \overline{X}_i \cdot P_i^* = 0 \cdot 0{,}0588 + 1 \cdot 0{,}1373 + \ldots + 6 \cdot 0{,}0196 \approx 2{,}6.$$

Вычислим статистическую дисперсию:

$$D_{\text{В}} = \sum_{i=1}^{k} (\overline{X}_i)^2 \cdot P_i^* - (x_{\text{В}})^2 =$$
$$= 0^2 \cdot 0{,}0588 + 1^2 \cdot 0{,}1373 + \ldots + 6^2 \cdot 0{,}0196 - 2{,}6^2 \approx 1{,}6894.$$

Определим среднее квадратическое отклонение:

$$\sigma^* = \sqrt{D_{\text{В}}} = \sqrt{1{,}6894} \approx 1{,}3.$$

Найдем выборочный коэффициент асимметрии:

$$A^* = \frac{1}{\sigma^3} \sum_{i=1}^{k} (\overline{X}_i - x_{\text{В}})^3 P_i^* =$$
$$= \frac{(0-2{,}6)^3 \cdot 0{,}0588 + (1-2{,}6)^3 \cdot 0{,}1373 + \ldots + (6-2{,}6)^3 \cdot 0{,}0196}{1{,}3^3} \approx 0{,}0262.$$

Вычислим выборочный коэффициент эксцесса:

$$E^* = \frac{1}{\sigma^4} \sum_{i=1}^{k} (\overline{X}_i - x_{\text{В}})^4 P_i^* - 3 =$$

$$= \frac{(0-2,6)^4 \cdot 0,0588 + (1-2,6)^4 \cdot 0,1373 + \ldots + (6-2,6)^4 \cdot 0,0196}{1,3^4} - 3 \approx 0,0097\,.$$

6. Выдвижение гипотезы о нормальном законе распределения генеральной совокупности.

По виду гистограммы, значениям коэффициентов асимметрии и эксцесса можно предположить, что случайная величина имеет нормальное распределение.

Выдвигаем гипотезу о нормальном законе распределения генеральной совокупности.

7. Оценка числовых характеристик и параметров закона распределения.

Выборочное среднее $x_{\text{В}} = \sum_{i=1}^{k} \overline{X}_i \cdot P_i^* \approx 2,6$ – несмещенная и состоятельная оценка математического ожидания.

Исправленная выборочная дисперсия $S^2 = D_{\text{В}} \cdot n / (n - 1)$ – несмещенная и состоятельная оценка дисперсии, т. е.

$$S^2 = \frac{51}{51-1} \cdot 1,6894 \approx 1,718.$$

Выдвинута гипотеза о распределении генеральной совокупности X по нормальному закону с плотностью

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{1}{2} \cdot \frac{(x-a)^2}{\sigma}}, \;\; x \in (-\infty; +\infty),$$

$$\text{или} \quad f(x) = \frac{1}{\sigma} \varphi\left(\frac{x-a}{\sigma}\right), \quad \text{где} \quad \varphi(t) = \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{t^2}{2}}.$$

Найдем оценки a и σ теоретического закона:

$$a \approx x_{\text{В}} = 2,6; \; \sigma \approx S = S^2 = \sqrt{1,718} \approx 1,31.$$

Заменяя a и σ этими оценками, получаем теоретический закон распределения:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{(x-2,6)^2}{2 \cdot 1,31^2}}.$$

Так как для случайной величины, имеющей нормальное распределение, функцию плотности можно представить в виде $f(x)=\frac{1}{\sigma}\varphi\left(\frac{x-a}{\sigma}\right)$, где $\varphi(t)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2}$, то для нашего случая имеем: $f(x)=\frac{1}{1{,}31}\varphi\left(\frac{x-2{,}6}{1{,}31}\right)$.

8. Нахождение доверительного интервала для математического ожидания.

Полагая, что генеральная совокупность X имеет нормальное распределение, построим доверительный интервал для математического ожидания при неизвестном среднем квадратическом отклонении.

В курсе математической статистики доказывается, что случайная величина в случае выборки из нормальной совокупности имеет распределение Стьюдента с (n – 1)-й степенью свободы, не зависящее от параметров генеральной совокупности. Доверительный интервал с выбранной надежностью γ, покрывающий математическое ожидание, будет иметь вид $(\bar{X}_{\text{в}} - t_\gamma S/\sqrt{n-1};\ \bar{X}_{\text{в}} + t_\gamma S/\sqrt{n-1})$. Для нахождения точности оценки $\delta = t_\gamma S/\sqrt{n-1}$ значение t_γ определяется по таблице Стьюдента [4] по заданной надежности γ и числу степеней свободы $k = (n - 1)$.

Построим 95%-й доверительный интервал для оценки математического ожидания. Тогда имеем γ = 0,95, число степеней свободы k = 51 – 1 = 50, уровень значимости $\alpha = 1 - \gamma = 0{,}05$. По таблице распределения Стьюдента [4] находим t_γ = 2,009. Точность оценки $\delta = t_\gamma S/\sqrt{n-1} = 2{,}009 \cdot 1{,}31/\sqrt{50} \approx 0{,}372$.

Получаем доверительный интервал для оценки математического ожидания нормально распределенной генеральной совокупности: (2,6 – 0,372; 2,6 + 0,372). Окончательно имеем (2,228; 2,972).

9. Проверка гипотезы о законе распределения генеральной совокупности по критерию Пирсона.

В качестве меры расхождения между статистическим и гипотетическим (теоретическим) распределениями возьмем критерий Пирсона $k = \chi^2$.

Мера расхождения в этом критерии определяется равенством

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - nP_i)^2}{n_i},$$

где k – число интервалов (у нас k = 7); n_i – эмпирические частоты (число элементов в i-м интервале); n – объем выборки (у нас n = 51); nP_i – теоретические частоты; P_i – теоретические вероятности попадания значений случайной величины в i-й интервал.

Найдем теоретические вероятности P_i, подставив $a = 2{,}6$, $\sigma = 1{,}31$ (табл. 5).

Таблица 5

i	$t_i = \dfrac{\overline{x}_i - a}{\sigma}$	$\varphi(t_i)$	$f(\overline{x}_i) = \dfrac{\varphi(t_i)}{\sigma}$	$p_i = f(\overline{x}_i) \cdot \Delta$
1	–1,98	0,0562	0,0429	0,0429
2	–1,22	0,01895	0,1447	0,1447
3	–0,46	0,3589	0,2740	0,2740
4	0,31	0,3802	0,2902	0,2902
5	1,07	0,2251	0,1718	0,1718
6	1,83	0,0748	0,0571	0,0571
7	2,6	0,0136	0,0104	0,0104
				$\sum = 0{,}9911$

Дальнейшие вычисления, необходимые для определения расчетного значения выборочной статистики χ^2, проведем в табл. 6.

Таблица 6

$[x_i; x_{i+1})$	[–0,5; 0,5)	[0,5; 1,5)	[1,5; 2,5)	[2,5; 3,5)	[3,5; 4,5)	[4,5; 5,5)	[5,5; 6,5)
n_i	3	7	13	16	9	2	1
P_i	0,0429	0,1447	0,274	0,2902	0,1718	0,0571	0,0104
$nP_i = 51P_i$	2,1879	7,3775	13,9724	14,8017	8,7634	2,9121	0,5295

В рассматриваемом эмпирическом распределении имеются частоты, меньшие 5. При использовании критерия Пирсона такие интервалы целесообразно объединять с соседними. После объединения интервалов с низкой степенью частоты получим вспомогательную табл. 7.

Таблица 7

$[x_i; x_{i+1})$	[–0,5; 1,5)	[1,5; 2,5)	[2,5; 3,5)	[3,5; 6,5)	$\sum$
n_i	10	13	16	12	51
nP_i	9,5654	13,9724	14,8017	12,205	50,5445
$(n_i - nP_i)^2$	0,1889	0,9456	1,4359	0,0420	–
$\dfrac{(n_i - nP_i)^2}{nP_i}$	0,0197	0,0677	0,0970	0,0034	$\chi^2_{\text{набл}} = 0,1879$

Случайная величина χ^2 (мера расхождения) независимо от вида закона распределения генеральной совокупности при $n \geq 50$ имеет распределение χ^2 с числом степеней свободы $r = k - \ell - 1$, где k – число интервалов после объединения; ℓ – число параметров распределения, определенных по выборке. В нашем примере $k = 4$, $\ell = 2$ (так как функция плотности распределения зависит только от двух параметров a и σ), $r = 4 - 2 - 1 = 1$.

Зададим уровень значимости $\alpha = 0,05$ и найдем по таблице значений χ^2 [4] критическое значение для $\alpha = 0,05$ и $r = 1$. Имеем $\chi^2_{0,05;\,1} = 3,8$. Так как $\chi^2_{\text{набл}} < \chi^2_{0,05;\,1}$, то предполагаемая гипотеза о нормальном законе распределения генеральной совокупности не противоречит опытным данным и принимается на уровне значимости α.

Пример 8

1. Исходные данные.

Дана выборка из генеральной совокупности случайной величины X. Данные представлены в табл. 8.

Таблица 8

8,0	12,5	15,4	6,9	11,4	7,2	10,5	11,5
17,7	13,6	15,1	13,4	17,9	18,6	9,8	12,6
14,9	7,3	16,5	15,5	12,9	11,0	16,8	18,4
12,8	11,4	13,5	16,2	14,3	12,1	12,2	18,1
10,9	7,9	17,9	18,6	10,5	13,7	10,3	17,2
13,5	17,7	6,7	17,1	16,4	7,1	16,9	14,2
11,3	15,2	15,8	12,3	9,9	15,6	18,9	14,2
8,2	11,5	18,6	19,0				

Выборка содержит 60 наблюдаемых значений, поэтому выборка имеет объем $n = 60$.

2. Построение вариационного ряда.

Проранжировав статистические данные, получаем вариационный ряд (табл. 9).

Таблица 9

№		№		№		№		№		№	
1	6,7	11	10,3	21	12,1	31	13,6	41	15,6	51	17,7
2	6,9	12	10,5	22	12,2	32	13,7	42	15,8	52	17,9
3	7,1	13	10,5	23	12,3	33	14,2	43	16,2	53	17,9
4	7,2	14	10,9	24	12,5	34	14,2	44	16,4	54	18,1
5	7,3	15	11	25	12,6	35	14,3	45	16,5	55	18,4
6	7,9	16	11,3	26	12,8	36	14,9	46	16,8	56	18,6
7	8	17	11,4	27	12,9	37	15,1	47	16,9	57	18,6
8	8,2	18	11,4	28	13,4	38	15,2	48	17,1	58	18,6
9	9,8	19	11,5	29	13,5	39	15,4	49	17,2	59	18,9
10	9,9	20	11,5	30	13,5	40	15,5	50	17,7	60	19

3. Построение интервального вариационного ряда.

Для построения вариационного интервального ряда выполняем следующие действия:

1) находим размах выборки $R = x_{\max} - x_{\min}$. Имеем $R = 19 - 6{,}7 = 12{,}3$.

2) определяем длину частичного интервала Δ – шаг разбиения по формуле Стерджеса: $\Delta \approx R / (1 + 3{,}322 \cdot \lg n) = R / K$, где n – объем выборки, K – число частичных интервалов. Так как $n = 60$, то $K = 1 + 3{,}322 \cdot \lg 60 \approx 7$, $\Delta \approx 12{,}3/7 \approx 2$;

3) определяем начало первого частичного интервала $x_{нач} = x_{\min} - \Delta/2 = 6{,}7 - 2/2$. Выбираем $x_{нач} = 5{,}7$.

После разбиения на частичные интервалы просматриваем ранжированную выборку и определяем, сколько значений признака попало в каждый частичный интервал, включая в него те значения, которые больше либо равны нижней границы и меньше верхней границы. Строим интервальный вариационный ряд (табл. 10).

Таблица 10

$[x_i; x_{i+1})$	[5,7; 7,7)	[7,7; 9,7)	[9,7; 11,7)	[11,7; 13,7)	[13,7; 15,7)	[15,7; 17,7)	[17,7; 19,7)
n_i	5	3	12	11	10	8	11

4. Построение гистограммы.

Для построения гистограммы строим вспомогательную табл. 11.

Таблица 11

i	Разряды $[x_i; x_{i+1})$	n_i	$P_i^* = \dfrac{n_i}{n}$	$f^*(x) = h^* = \dfrac{P_i^*}{\Delta}$	$\overline{x}_i = \dfrac{x_i + x_{i+1}}{2}$
1	[5,7; 7,7)	5	0,0833	0,04165	6,7
2	[7,7; 9,7)	3	0,05	0,025	8,7
3	[9,7; 11,7)	12	0,2	0,1	10,7
4	[11,7; 13,7)	11	0,1833	0,09165	12,7
5	[13,7; 15,7)	10	0,1667	0,08335	14,7
6	[15,7; 17,7)	8	0,1334	0,0667	16,7
7	[17,7; 19,7)	11	0,1833	0,09165	18,7
Контроль		60	1		

По данным табл. 11 строим гистограмму частостей (рис. 19).

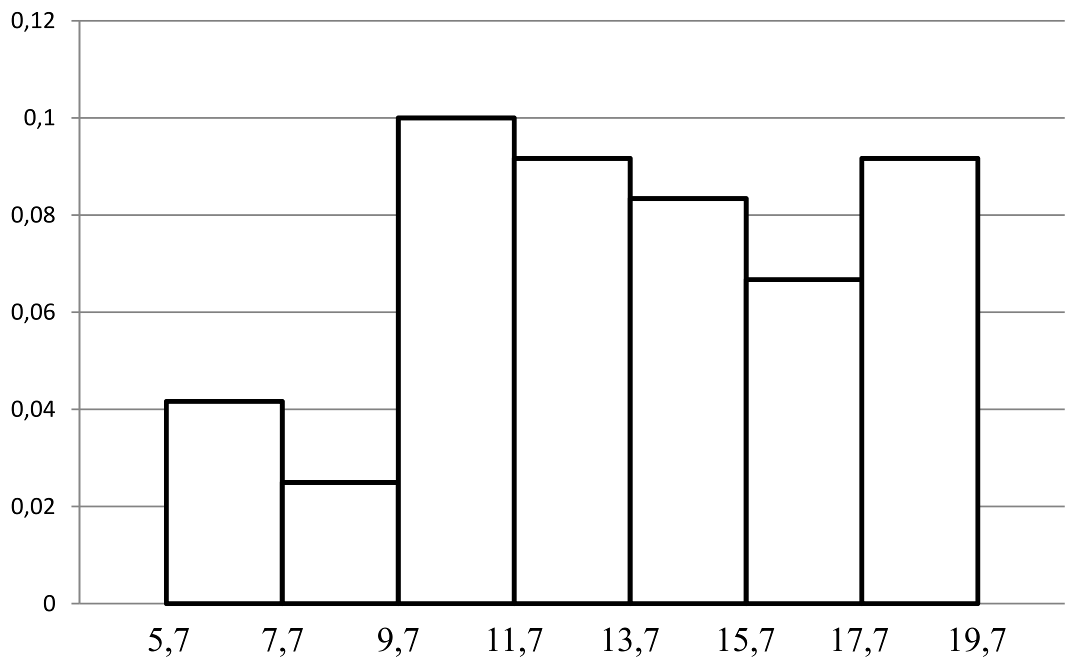


Рис. 19

5. Нахождение числовых характеристик выборки.

Рассчитаем статистическое среднее по формуле

$$x_{\text{В}} = \sum_{i=1}^{k} \bar{X}_i \cdot P_i^* = 6{,}7 \cdot 0{,}0833 + 8{,}7 \cdot 0{,}05 + \ldots + 18{,}7 \cdot 0{,}1833 \approx 13{,}567.$$

Вычислим статистическую дисперсию:

$$D_{\text{В}} = \sum_{i=1}^{k} (\bar{X}_i)^2 \cdot P_i^* - (x_{\text{В}})^2 =$$
$$= 6{,}7^2 \cdot 0{,}0833 + 8{,}7^2 \cdot 0{,}05 + \ldots + 18{,}7^2 \cdot 0{,}1833 - 13{,}567^2 \approx 13{,}247.$$

Определим среднее квадратическое отклонение:

$$\sigma^* = \sqrt{D_{\text{В}}} = \sqrt{13{,}247} \approx 3{,}64.$$

Найдем выборочный коэффициент асимметрии:

$$A^* = \frac{1}{\sigma^3} \sum_{i=1}^{k} (\bar{X}_i - x_{\text{В}})^3 P_i^* =$$
$$= \frac{(6{,}7 - 13{,}567)^3 \cdot 0{,}0833 + (8{,}7 - 13{,}567)^3 \cdot 0{,}05 + \ldots + (18{,}7 - 13{,}567)^3 \cdot 0{,}1833}{3{,}64^3} = -0{,}003629.$$

Вычислим выборочный коэффициент эксцесса:

$$E^* = \frac{1}{\sigma^4} \sum_{i=1}^{k} (\bar{X}_i - x_{\text{В}})^4 P_i^* - 3 =$$
$$= \frac{(6{,}7 - 13{,}567)^4 \cdot 0{,}0833 + (18{,}7 - 13{,}567)^4 \cdot 0{,}05 + \ldots + (18{,}7 - 13{,}567)^4 \cdot 0{,}1833}{3{,}64^4} - 3 = 0{,}0119264.$$

6. Выдвижение гипотезы о законе распределения генеральной совокупности.

По виду гистограммы можно предположить, что случайная величина имеет равномерное распределение.

Выдвигаем гипотезу о равномерном законе распределения генеральной совокупности.

7. Оценка числовых характеристик и параметров закона распределения.

Выборочное среднее $x_{\text{В}} = \sum_{i=1}^{k} \bar{X}_i \cdot P_i^* \approx 13{,}567$ – несмещенная и состоятельная оценка математического ожидания.

Исправленная выборочная дисперсия $S^2 = D_{\text{В}} \cdot n / (n - 1)$ – несмещенная и состоятельная оценка дисперсии, т. е.

$$S^2 = \frac{60}{60-1} \cdot 13{,}247 \approx 13{,}472.$$

Выдвинута гипотеза о распределении генеральной совокупности X по нормальному закону с плотностью

$$f(x)=\begin{cases} 0, & x\notin (a;b), \\ \dfrac{1}{b-a}, & x\in (a;b). \end{cases}$$

Найдем оценки a^* и b^* – оценки параметров a и b теоретического закона крайних значений выборки, которые находятся из системы уравнений

$$\begin{cases} x_{\text{В}}=(a+b)/2; \\ S^2=(b-a)^2/12. \end{cases}$$

Имеем $(a + b)/2 = 13{,}567$; $(b – a)^2/12 = 13{,}472$.

Следовательно, $a^* = 7{,}21$; $b^* = 19{,}92$.

Заменяя a и b оценками a^* и b^*, получаем теоретический закон распределения

$$f(x)=\begin{cases} 0, & x\notin (7{,}21;19{,}92). \\ 0{,}0787, & x\in (7{,}21;19{,}92). \end{cases}$$

8. Проверка гипотезы о законе распределения генеральной совокупности по критерию Пирсона.

В качестве меры расхождения между статистическим и гипотетическим (теоретическим) распределениями возьмем критерий Пирсона $k = \chi^2$.

Мера расхождения в этом критерии определяется равенством

$$\chi^2=\sum\nolimits_{i=1}^{k}\frac{(n_i-nP_i)^2}{n_i},$$

где k – число интервалов (у нас $k = 7$); n_i – эмпирические частоты (число элементов в i-м интервале); n – объем выборки (у нас $n = 60$); nP_i – теоретические частоты; P_i – теоретические вероятности попадания значений случайной величины в i-й интервал.

В рассматриваемом эмпирическом распределении имеются частоты, меньшие 5. При использовании критерия Пирсона такие интервалы целесообразно объединять с соседними. После объединения интервалов с низкой степенью частоты получим вспомогательную табл. 12.

Таблица 12

$[x_i; x_{i+1})$	[5,7; 9,7)	[9,7; 11,7)	[11,7; 13,7)	[13,7; 15,7)	[15,7; 17,7)	[17,7; 19,7)	$\sum$
n_i	8	12	11	10	8	11	60

Находим теоретические вероятности P_i по формуле

$$P_i = \int_{x_{i-1}}^{x_i} f(x)\,dx = \int_{x_{i-1}}^{x_i} \frac{1}{b^* - a^*}\,dx = \frac{x_i - x_{i-1}}{12,71},$$

$$P_1 = \int_{5,7}^{9,7} \frac{1}{12,71}\,dx = \frac{9,7 - 5,7}{12,71} = 0,315,$$

$$P_2 = \int_{9,7}^{11,7} \frac{1}{12,71}\,dx = \frac{11,7 - 9,7}{12,71} = 0,1574,$$

$$P_3 = \int_{11,7}^{13,7} \frac{1}{12,71}\,dx = \frac{13,7 - 11,7}{12,71} = 0,1574,$$

$$P_4 = \int_{13,7}^{15,7} \frac{1}{12,71}\,dx = \frac{15,7 - 13,7}{12,71} = 0,1574,$$

$$P_5 = \int_{15,7}^{17,7} \frac{1}{12,71}\,dx = \frac{17,7 - 15,7}{12,71} = 0,1574,$$

$$P_6 = \int_{17,7}^{19,7} \frac{1}{12,71}\,dx = \frac{19,7 - 17,7}{12,71} = 0,1574.$$

Дальнейшие вычисления, необходимые для определения расчетного значения выборочной статистики χ^2, проведем в табл. 13.

Таблица 13

$[x_i; x_{i+1})$	[5,7; 9,7)	[9,7; 11,7)	[11,7; 13,7)	[13,7; 15,7)	[15,7; 17,7)	[17,7; 19,7)	$\sum$
n_i	8	12	11	10	8	11	60
$nP_i = 60P_i$	18,88	9,444	9,444	9,444	9,444	9,444	–
$(n_i - nP_i)^2$	100	6,533	2,421	0,309	2,085	2,421	–
$\frac{(n_i - nP_i)^2}{nP_i}$	5,29	0,692	0,256	0,033	0,221	0,256	$\chi^2 = 6,75$

Случайная величина χ^2 (мера расхождения) независимо от вида закона распределения генеральной совокупности при $n \geq 50$ имеет распределение χ^2 с числом степеней свободы $r = k - \ell - 1$, где k – число интервалов после объединения; ℓ – число параметров распределения, определенных по выборке. В нашем примере $k = 6$, $\ell = 2$ (так как функция плотности распределения зависит только от двух параметров a и b), $r = 6 - 2 - 1 = 3$.

Зададим уровень значимости $\alpha = 0{,}05$ и найдем по таблице значений χ^2 [4] критическое значение для $\alpha = 0{,}05$ и $r = 3$. Имеем $\chi^2_{0,05;\,3} = 7{,}8$. Так как $\chi^2_{\text{набл}} < \chi^2_{0,05;\,3}$, то предполагаемая гипотеза о равномерном законе распределения генеральной совокупности не противоречит опытным данным и принимается на уровне значимости α.

Пример 9

1. Исходные данные.

Дана выборка из генеральной совокупности случайной величины X. Данные представлены в табл. 14.

Таблица 14

2,79	7,42	42,9	2,06	12,4	39,9	30,6	13,7	26	23,2
17,5	76,9	25,8	16,6	10,1	59,5	1,63	25,5	12,1	18,6
26,1	31,3	38,2	2,23	13,6	5,2	1,37	3,03	38,2	6,25
70,1	8,45	2,8	50,6	0,459	8,6	2,61	8,27	9,04	27,5
1,91	9,47	10,1	12,3	58,4	26,1	30,5	29,8	7,97	10,3

Выборка содержит 50 наблюдаемых значений, поэтому выборка имеет объем $n = 50$.

2. Построение вариационного ряда.

Проранжировав статистические данные, получаем вариационный ряд (табл. 15).

Таблица 15

№		№		№		№		№	
1	0,459	11	5,2	21	10,3	31	25,5	41	38,2
2	1,37	12	6,25	22	12,1	32	25,8	42	38,2
3	1,63	13	7,42	23	12,3	33	26	43	39,9
4	1,91	14	7,97	24	12,4	34	26,1	44	40,1
5	2,06	15	8,27	25	13,6	35	26,1	45	42,9
6	2,23	16	8,45	26	13,7	36	27,5	46	50,6
7	2,61	17	8,6	27	16,6	37	29,8	47	58,4
8	2,79	18	9,04	28	17,5	38	30,5	48	52,5
9	2,8	19	9,47	29	18,6	39	30,6	49	70,1
10	3,03	20	10,1	30	23,2	40	31,3	50	76,9

3. Построение гистограммы частот.

Для построения гистограммы строим вспомогательную табл. 16.

Таблица 16

Разряды $[x_i; x_{i+1})$	n_i	$P_i^* = \frac{n_i}{n}$	$f^*(x) = h^* = \frac{P_i^*}{\Delta}$	$\bar{x}_i = \frac{x_i + x_{i+1} + 1}{2}$
[0; 10)	19	$0{,}38 = \frac{19}{50}$	0,038	5
[10; 20)	10	$0{,}2 = \frac{10}{50}$	0,02	15
[20; 30)	8	$0{,}16 = \frac{8}{50}$	0,016	25
[30; 40)	6	$0{,}12 = \frac{6}{50}$	0,012	35
[40; 50)	2	$0{,}04 = \frac{2}{50}$	0,004	45
[50; 60)	3	$0{,}06 = \frac{3}{50}$	0,006	55
[60; 70)	0	0	0	65
[70; 80)	2	$0{,}04 = \frac{2}{50}$	0,004	75
	50	1		

Рассмотрим отрезок [0; 80] при $\Delta = 10$, $x_{\text{нач}} = 0$.

По данным табл. 16 строим гистограмму частот (рис. 20).

Гистограмма частостей является статистическим аналогом дифференциальной функции распределения $f(x)$ случайной величины X.

Площадь гистограммы частостей равна 1.

4. Нахождение числовых характеристик выборки.

Рассчитаем статистическое среднее по формуле

$$x_{\text{В}} = \sum_{i=1}^{k} \bar{X}_i \cdot P_i^* = 0{,}46 \cdot 0{,}22 + 11{,}46 \cdot 0{,}32 + 22{,}46 \cdot 0{,}18 + 33{,}46 \cdot 0{,}12 +$$
$$+ 44{,}46 \cdot 0{,}06 + 55{,}46 \cdot 0{,}06 + 66{,}46 \cdot 0{,}02 + 77{,}46 \cdot 0{,}02 =$$
$$= 0{,}1012 + 3{,}667 + 4{,}042 + 4{,}015 + 2{,}667 + 3{,}327 + 1{,}329 + 1{,}549 = 20{,}697.$$

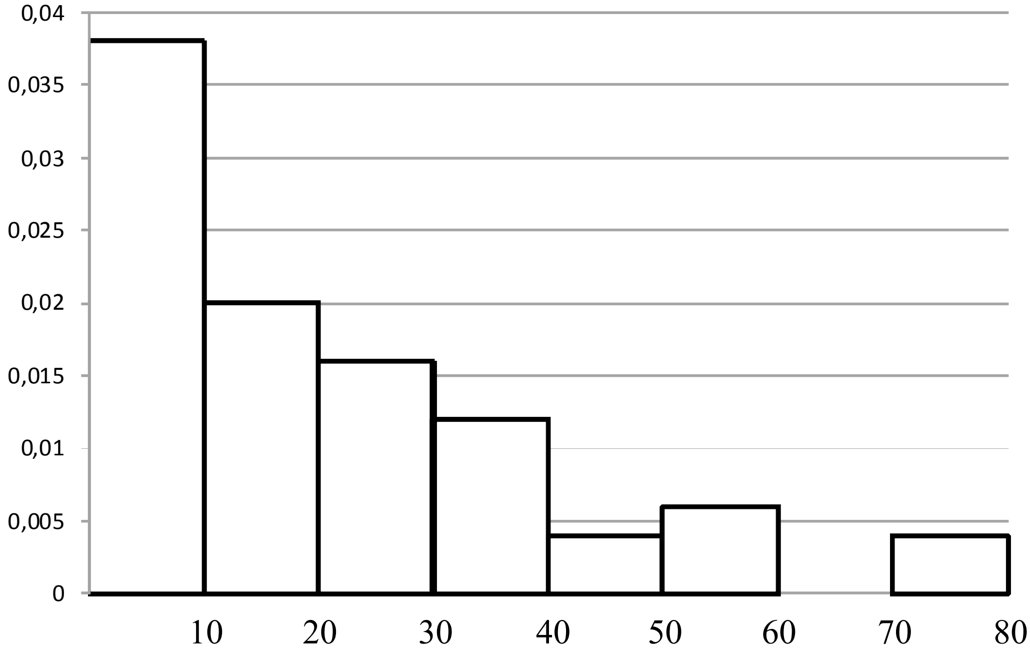


Рис. 20

Вычислим статистическую дисперсию:

$$D_{\text{В}} = \sum_{i=1}^{k} (\bar{X}_i)^2 \cdot P_i^* - (x_{\text{В}})^2 =$$

$$= (0{,}46)^2 \cdot 0{,}22 + (11{,}46)^2 \cdot 0{,}32 + (22{,}46)^2 \cdot 0{,}18 + (33{,}46)^2 \cdot 0{,}12 + (44{,}46)^2 \cdot 0{,}06 +$$
$$+ (55{,}46)^2 \cdot 0{,}06 + (66{,}46)^2 \cdot 0{,}02 + (77{,}46)^2 \cdot 0{,}02 - (20{,}697)^2 =$$
$$= 0{,}046 + 42{,}026 + 90{,}8 + 134{,}35 + 118{,}6 + 184{,}54 + 88{,}33 + 120 - 428{,}365 =$$
$$= 778{,}692 - 428{,}365 = 350{,}327.$$

Определим среднее квадратическое отклонение:

$$\sigma^* = \sqrt{D_{\text{В}}} = \sqrt{350{,}327} \approx 18{,}71.$$

Найдем выборочный коэффициент асимметрии:

$$A^* = \frac{1}{\sigma^3} \sum_{i=1}^{k} (\bar{X}_i - x_{\text{В}})^3 P_i^* =$$

$$= \frac{(0{,}46 - 20{,}697)^3 \cdot 0{,}22 + (11{,}46 - 20{,}697)^3 \cdot 0{,}32 + (22{,}46 - 20{,}697)^3 \cdot 0{,}18}{18{,}71^3} +$$

$$+ \frac{(33{,}46 - 20{,}697)^3 \cdot 0{,}12 + (44{,}46 - 20{,}697)^3 \cdot 0{,}06 + (55{,}46 - 20{,}697)^3 \cdot 0{,}06}{18{,}71^3} +$$

$$+ \frac{\left(66{,}46 - 20{,}697\right)^3 \cdot 0{,}02 + \left(77{,}46 - 20{,}697\right)^3 \cdot 0{,}02}{18{,}71^3} =$$

$$= \frac{-8287,78 \cdot 0,22 + (-788,12) \cdot 0,32 + 5,479 \cdot 0,18 + 2079,018 \cdot 0,12}{18,71^3} +$$

$$+ \frac{13\,418,49 \cdot 0,06 + 42\,009 \cdot 0,06 + 95\,839,26 \cdot 0,02 + 182\,892,55 \cdot 0,02}{18,71^3} =$$

$$= \frac{-1823,31 - 252,198 + 0,986 + 249,48 + 805,1 + 2520,5 + 1916,78 + 3657,85}{18,71^3} =$$

$$= \frac{7075,189}{18,71^3} = \frac{7075,189}{6549,699} \approx 1,08.$$

Вычислим выборочный коэффициент эксцесса:

$$E^* = \frac{1}{\sigma^4} \sum_{i=1}^{k} (\overline{X}_i - x_{\text{В}})^4 P_i^* - 3 =$$

$$= \frac{(0,46 - 20,697)^4 \cdot 0,22 + (11,46 - 20,697)^4 \cdot 0,32 + (22,46 - 20,697)^4 \cdot 0,18}{(18,71)^4} +$$

$$+ \frac{(33,46 - 20,697)^4 \cdot 0,12 + (44,46 - 20,697)^4 \cdot 0,06 + (55,46 - 20,697)^4 \cdot 0,06}{(18,71)^4} +$$

$$+ \frac{(66,46 - 20,697)^4 \cdot 0,02 + (77,46 - 20,697)^4 \cdot 0,02}{(18,71)^4} - 3 =$$

$$= \frac{36\,898,37 + 2329,55 + 1,738 + 3184,14 + 19\,131,82 + 87\,623,42 + 87\,717,8 + 207\,630}{122\,544,87} - 3 =$$

$$= 3,627 – 3 = 0,627.$$

5. Выдвижение гипотезы о законе распределения генеральной совокупности.

По виду гистограммы можно предположить, что случайная величина имеет показательное распределение.

Найдем $\overline{x} = \sum_{i=1}^{8} \overline{x}_i \cdot p_i = 25,2$ = 25,2.

Примем в качестве оценки параметра предполагаемого показательного распределения λ = 1/$\overline{x}$ = 1/25,2 = 0,039 ≈ 0,04, ⇒ плотность предполагаемого показательного распределения имеет вид

$$f(x) = \begin{cases} 0,04 \cdot \mathrm{e}^{-0,04x}, & \text{если } x \ge 0, \\ 0, & \text{если } x < 0. \end{cases}$$

Найдем вероятности попадания X в каждый из интервалов по формуле

$P_i = P(x_i < X < x_{i+1}) = e^{-\lambda x_i} - e^{-\lambda x_{i+1}}$, т. е.

$$P_1 = P(0 < X < 10) = e^{-0,04\cdot 0} - e^{-0,04\cdot 10} = e^0 - e^{-0,4} = 1 - \frac{1}{e^{0,4}} = 1 - 0,672 = 0,328;$$

$$P_2 = P(10 < X < 20) = e^{-0,04\cdot 10} - e^{-0,04\cdot 20} = \frac{1}{e^{0,4}} - \frac{1}{e^{0,8}} = 0,672 - 0,452 = 0,22;$$

$$P_3 = P(20 < X < 30) = 0,452 - \frac{1}{e^{1,2}} = 0,452 - 0,304 = 0,148;$$

$$P_4 = P(30 < X < 40) = 0,304 - \frac{1}{e^{1,6}} = 0,304 - 0,204 = 0,1;$$

$$P_5 = P(40 < X < 50) = 0,204 - \frac{1}{e^{2}} = 0,204 - 0,137 = 0,067;$$

$$P_6 = P(50 < X < 60) = 0,137 - \frac{1}{e^{2,4}} = 0,137 - 0,092 = 0,045;$$

$$P_7 = P(60 < X < 70) = 0,092 - \frac{1}{e^{2,8}} = 0,092 - 0,0619 = 0,0301;$$

$$P_8 = P(70 < X < 80) = 0,0619 - \frac{1}{e^{3,2}} = 0,0619 - 0,0416 = 0,0203.$$

Контроль: $\sum_{i=1}^{8} p_i = 0,9584 \approx 1.$

Найдем теоретические частоты по формуле $n_i' = 50 \cdot p_i$;

$n_1' = 50 \cdot 0,328 = 16,4;$

$n_2' = 50 \cdot 0,22 = 11;$

$n_3' = 50 \cdot 0,148 = 7,4;$

$n_4' = 50 \cdot 0,1 = 5;$

$n_5' = 50 \cdot 0,067 = 3,35;$

$n_6' = 50 \cdot 0,045 = 2,25;$

$n_7' = 50 \cdot 0,0301 = 1,505;$

$n_8' = 50 \cdot 0,0203 = 1,015.$

Объединим соседние интервалы с низкой степенью частоты, меньшие 5, и сведем интервалы в табл. 17.

Таблица 17

$[x_i; x_{i+1})$	[0; 10)	[10; 20)	[20; 30)	[30; 40)	[40; 50)	[50; 60)	[60; 70)	[70; 80)
n_i	19	10	8	6	2	3	0	2
P_i	0,328	0,22	0,148	0,1	0,067	0,045	0,0301	0,0203
nP_i	16,4	11	7,4	5	3,35	2,25	1,505	1,015

Сравним эмпирические и теоретические частоты с помощью критерия Пирсона.

Для этого составим расчетную табл. 18, причем объединим малочисленные частоты (2 + 3 + 0 + 2 = 7) и соответствующие им теоретические частоты (3,35 + 2,25 + 1,505 + 1,1015 = 8,12).

Таблица 18

$[x_i; x_{i+1})$	[0; 10)	[10; 20)	[20; 30)	[30; 40)	[40; 80)	$\sum$
n_i	19	10	8	6	7	50
$nP_i = n_i'$	16,4	11	7,4	5	8,12	47,92
$(n_i - nP_i)^2$	6,76	1	0,36	1	1,254	–
$\dfrac{(n_i - nP_i)^2}{nP_i}$	0,412	0,09	0,0486	0,2	0,154	$\chi^2 = 0{,}9046$

Следовательно, $\chi^2_{\text{набл}} = 0{,}9046$.

По таблице критических точек распределения для χ^2 [4] находим уровень значимости $\alpha = 0{,}05$ и число степеней свободы S. Число интервалов $k = S - 2 = 5 - 2 = 3$. Находим критическую точку правосторонней критической области $\chi^2_{\text{кр}}(0{,}05; 3) = 7{,}8$. Так как $\chi^2_{\text{набл}} < \chi^2_{\text{кр}}$, следовательно, распределение показательное. При использовании критерия Пирсона число степеней свободы

$$k = S - 1 - r,$$

где r – число параметров, оцениваемых по выборке (у нас $r = 1$ при единственном значении λ).

3. ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

3.1. Классическое определение вероятности. Формулы комбинаторики

1. Среди 17 студентов, из которых 8 девушек, разыгрывается 7 билетов. Какова вероятность того, что среди обладателей окажется четыре девушки.

Ответ: 0,302.

2. Из 20 сбербанков 10 расположены за чертой города. Для обследования случайным образом отобрано 5 сбербанков. Какова вероятность того, что среди отобранных окажется в черте города три сбербанка?

Ответ: 0,348.

3. Из колоды 36 карт берут наугад три карты. Какова вероятность того, что они одинакового цвета?

Ответ: 0,114.

4. Комиссия по качеству раз в месяц проверяет качество продуктов в двух из 30 магазинов, среди которых находятся и два известных вам магазина. Какова вероятность того, что в течение месяца они оба будут проверены?

Ответ: 0,00229.

5. Из 20 акционерных обществ (АО) четыре являются банкротами. Гражданин приобрел по одной акции шести АО. Какова вероятность того, что среди купленных акций две окажутся акциями банкротов?

Ответ: 0,281.

6. На склад привезли 50 ящиков комплектующих изделий для одного из видов ЭВМ, но среди них оказалось четыре ящика комплектующих для другого вида ЭВМ. Наудачу взяли шесть ящиков. Найти вероятность того, что в одном из этих шести ящиков окажутся некомплектные детали.

Ответ: 0,345.

7. В ящике мастера 20 годных деталей и 4 дефектных. Какова вероятность того, что из двух наугад взятых деталей хотя бы одна дефектная?

Ответ: 0,758.

8. В лотерее 1000 билетов, из них 100 выигрышных. Какова вероятность того, что из пяти купленных билетов: а) нет ни одного выигрышного; б) хотя бы один выигрышный?

Ответ: а) 0,602; б) 0,398.

9. Какова вероятность угадать четырехзначное число, если оно начинается цифрой 5 и остальные цифры разные.

Ответ: 0,00033.

10. Из урны, в которой имеется 10 белых и 5 черных шаров, берут наугад два шара. Найти вероятность того, что оба шара белые?

Ответ: 0,429.

3.2. ДЕЙСТВИЯ С СОБЫТИЯМИ

1. Вероятность правильного оформления счета на предприятии составляет 0,95. Во время аудиторской проверки были взяты два счета. Какова вероятность того, что только один из них оформлен правильно?

Ответ: 0,095.

2. Три исследователя, независимо друг от друга, производят измерения некоторой физической величины. Вероятность того, что первый исследователь допустит ошибку при считывании показаний прибора, равна 0,1. Для второго и третьего исследователей эта вероятность соответственно равна 0,15 и 0,2. Найти вероятность того, что при однократном измерении хотя бы один из исследователей допустит ошибку.

Ответ: 0,388.

3. Покупатель может приобрести акции двух компаний А и В. Надежность первой оценивается экспертами на уровне 90 %, а второй – 80 %. Чему равна вероятность того, что обе компании в течение года не станут банкротами?

Ответ: 0,72.

4. В первом ящике 5 белых, 7 черных и 3 красных шара, во втором ящике 7 белых, 2 черных и 4 красных шара. Из каждого ящика наугад вынима-

ются по одному шару. Найти вероятность того, что оба выбранных шара – одного цвета.

Ответ: 0,07; 0,179; 0,06.

5. На двадцати одинаковых жетонах написаны двадцать двузначных чисел от 11 до 30. Жетоны помещены в пакет и тщательно перемешаны. Какова вероятность вынуть жетон с номером, кратным 4 или 7?

Ответ: 0,35.

6. Рабочий обслуживает три станка. Вероятности нарушения нормальной работы в течение часа равны: для первого станка – 0,04; для второго – 0,02; для третьего – 0,025. Найти вероятность того, что в течение часа: а) лишь один из станков не будет работать нормально; б) не менее чем один станок будет работать нормально.

Ответ: а) 0,079; б) 0,99.

7. Вероятность того, что студент первого курса перейдет на второй, равна 0,9, а вероятность того, что студент первого курса окончит институт, равна 0,8. Какова вероятность того, что студент второго курса окончит институт?

Ответ: 0,72.

8. Вероятность для компании, занимающейся строительством терминалов для аэропортов, получить контракт в стране А равна 0,4; вероятность выиграть его в стране В равна 0,3. Вероятность того, что контракты будут заключены и в стране А, и в стране В, равна 0,12. Чему равна вероятность того, что компания получит контракт хотя бы в одной стране?

Ответ: 0,42.

9. Для некоторой местности среднее число пасмурных дней в июне равно 6. Какова вероятность того, что первого и второго июня будет ясная погода?

Ответ: 0,64.

10. Вероятность того, что книга имеется в первой библиотеке, равна 0,5; во второй – 0,7; в третьей – 0,4. Какова вероятность наличия книги хотя бы в одной библиотеке?

Ответ: 0,91.

3.3. Формула полной вероятности. Формула Байеса

1. С первого автомата на сборку поступает 20 % деталей, со второго – 30 %, с третьего – 50 %. Первый автомат дает в среднем 0,2 % брака, второй – 0,3 %, третий – 0,1 %. Найти вероятность того, что поступившая на сборку деталь бракованная.

Ответ: 0,18.

2. Результаты исследований показали, что 70 % женщин позитивно реагируют на изучаемый круг ситуаций, в то время как 40 % мужчин реагируют на них негативно. 15 женщин и 5 мужчин заполнили анкеты, в которых отразили свое отношение к предлагаемым ситуациям. Случайно извлеченная анкета содержит негативную реакцию. Чему равна вероятность того, что ее заполнял мужчина?

Ответ: 0,307.

3. В обувную мастерскую для ремонта приносят сапоги и туфли в соотношении 2:3. Вероятность качественного ремонта для сапог равна 0,9, а для туфель – 0,85. Проведена проверка качества одной пары обуви. Оказалось, что эта пара обуви отремонтирована качественно. Какова вероятность того, что это туфли?

Ответ: 0,586.

4. Таможенный досмотр автомашин осуществляют два инспектора. В среднем из каждых 100 машин 45 проходит через первого инспектора. Вероятность того, что при досмотре машина, соответствующая таможенным правилам, не будет задержана, составляет 0,95 у первого инспектора и 0,85 у второго. Машина, соответствующая таможенным правилам, не была задержана. Найти вероятность того, что она прошла досмотр у первого инспектора.

Ответ: 0,475.

5. Вероятность того, что клиент банка не вернет заем в период экономического роста, равна 0,04, а в период экономического кризиса – 0,13. Предположим, что вероятность того, что начнется период экономического роста, равна 0,65. Чему равна вероятность того, что случайно выбранный клиент банка не вернет полученный кредит?

Ответ: 0,072.

6. В специализированную больницу поступают в среднем 50 % больных с заболеванием А, 30 % – с заболеванием В, 20 % – с заболеванием С. Вероятность полного излечения болезни А равна 0,7, для болезней В и С эти вероятности соответственно равны 0,8 и 0,9. Пациент выздоровел и был выписан. Найти вероятность того, что этот больной страдал заболеванием А.

Ответ: 0,45.

7. На предприятии, изготавливающем замки, первый цех производит 35 %, второй 32 %, третий 33 % всех замков. Брак составляет соответственно 4 %, 3 % и 2 %. Найти вероятность того, что случайно выбранный замок является дефектным.

Ответ: 0,0302.

8. Вероятность того, что новый товар будет пользоваться спросом на рынке, если конкурент не выпустит в продажу аналогичный продукт, равна 0,67. Вероятность того, что товар будет пользоваться спросом при наличии на рынке конкурирующего товара, равна 0,42. Вероятность того, что конкурирующая фирма выпустит аналогичный товар на рынке в течение интересующего нас периода, равна 0,35. Чему равна вероятность того, что товар будет иметь успех?

Ответ: 0,583.

9. Изделие проверяется на стандартность одним из двух товароведов. Вероятность того, что изделие попадет к первому товароведу, равна 0,55, а ко второму – 0,45. Вероятность того, что стандартное изделие будет признано стандартным первым товароведом, равна 0,9, а вторым – 0,98. Стандартное изделие при проверке было признано стандартным. Найти вероятность того, что это изделие проверил второй товаровед.

Ответ: 0,471.

10. Эксперт по туризму, нанятый компанией, организующей круизы, предположил, что вероятность того, что корабль будет полон в течение сезона, равна 0,92, если доллар не подорожает по отношению к рублю, и с вероятностью 0,75, если доллар подорожает. По оценкам экономистов, вероятность того, что в течение сезона доллар подорожает по отношению к рублю, равна 0,23. Чему равна вероятность того, что билеты на все круизы будут проданы?

Ответ: 0,8809.

3.4. Схема с повторением испытаний

1. Вероятность того, что расход электроэнергии на протяжении одних суток не превысит установленной нормы, равна 0,75. Найти вероятность того, что в ближайшие 6 суток расход электроэнергии в течение четырех суток не превысит нормы.

Ответ: 0,3.

2. Оптовая база снабжает 10 магазинов, от каждого из которых поступает заявка на очередной день с вероятностью 0,4 независимо от заявок других магазинов. Найти наивероятнейшее число заявок в день и вероятность получения этого числа заявок.

Ответ: $k_0 = 4$; 0,251.

3. Производится 19 выстрелов из винтовки. Вероятность попадания в цель при каждом выстреле равна 0,8. Найти наивероятнейшее число попаданий в цель.

Ответ: 15; 16.

4. Вероятность изготовления детали высшего сорта на данном станке равна 0,4. Найти вероятность того, что среди наудачу взятых 26 деталей половина окажется высшего сорта.

Ответ: 0,093.

5. Стрелок сделал 30 выстрелов с вероятностью попадания при отдельном выстреле 0,3. Найти вероятность того, что при этом будет 8 попаданий.

Ответ: 0,15.

6. Средний процент нарушения работы кинескопа в течение гарантийного срока равен 12. Вычислить вероятность того, что из 46 наблюдаемых телевизоров не менее 36 выдержат гарантийный срок.

Ответ: 0,973.

7. На склад поступает продукция трех фабрик, причем изделия первой фабрики составляют 30 %, второй – 32 % и третьей – 38 %. В продукции первой фабрики 60 % изделий высшего сорта, второй – 25 %, третьей – 50 %. Найти вероятность того, что среди 300 наудачу взятых со склада изделий число изделий высшего сорта заключено между 130 и 170.

Ответ: 0,72.

8. Вероятность попадания в самолет при одном выстреле равна 0,01. Производится 100 выстрелов. Определить вероятность двух попаданий.

Ответ: 0,184.

9. Известно, что в принятой для сборки партии из 1000 деталей имеются 4 дефектных. Найти вероятность того, что среди 50 наудачу взятых деталей нет дефектных.

Ответ: 0,819.

10. Игральную кость бросают 80 раз. Найти с вероятностью 0,99 границы, в которых будет заключено число m выпадений шестерки.

Ответ: $5 \le m \le 22$.

3.5. Случайные величины

1. В партии из шести деталей имеется четыре стандартных. Наудачу отобраны три детали. Составить закон распределения дискретной случайной величины X – числа стандартных деталей среди отобранных.

Ответ:

x	1	2	3
P	0,2	0,6	0,2

2. Предприятие С по выпуску бытовой техники изготовляет холодильники с современным уровнем производства. Для проверки качества выбирают по номерам 5 холодильников из N произведенных. Вероятность того, что изготовленный холодильник отвечает требованиям международного стандарта, равна 0,9. Составить закон распределения случайной величины X – числа отвечающих требованиям международного стандарта холодильников среди отобранных.

Ответ:

x_i	0	1	2	3	4	5
p	0,00001	0,00045	0,0081	0,0729	0,32805	0,59049

3. Найти дисперсию и среднее квадратическое отклонение дискретной случайной величины X, заданной законом распределения:

X	–5	2	3	4
p	0,4	0,3	0,1	0,2

Ответ: $D(X) = 15,21$; $\sigma = \sqrt{15,21}$.

4. Охотник, имевший 4 патрона, стреляет в цель до первого попадания (или пока не израсходует все патроны). Найти математическое ожидание и дисперсию числа израсходованных патронов, если вероятность попадания при каждом выстреле равна 0,25.

Ответ: $M(X) = 2{,}73$; $D(x) = 1{,}54$.

5. Случайная величина X задана функцией распределения

$$F(x) = \begin{cases} 0 & \text{при } x \le 0, \\ \dfrac{x^2}{9} & \text{при } 0 < x \le 3, \\ 1 & \text{при } x > 3. \end{cases}$$

Найти плотность распределения $f(x)$.

Ответ: $f(x) = \begin{cases} 0 & \text{при } x \le 0, \\ \dfrac{2}{9}x & \text{при } 0 < x \le 3, \\ 0 & \text{при } x > 3. \end{cases}$

6. Задана плотность распределения непрерывной случайной величины X:

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

Найти функцию распределения $F(x)$.

Ответ: $F(x) = \begin{cases} 0 & \text{при } x \le 0, \\ 1 - \cos x & \text{при } 0 < x \le \pi/2, \\ 0 & \text{при } x > \pi/2. \end{cases}$

7. Плотность распределения непрерывной случайной величины X имеет вид:

$$f(x) = \begin{cases} 0 & \text{при } x \le -\pi/2, \\ c \cdot \cos x & \text{при } -\pi/2 < x \le \pi/2, \\ 0 & \text{при } x > \pi/2. \end{cases}$$

Найти постоянный параметр c.

Ответ: 0,5.

8. Случайная величина X в интервале (0, 5) задана плотностью распределения $f(x) = \frac{2}{25}x$; вне этого интервала $f(x) = 0$. Найти дисперсию X.

Ответ: $D(X) = 1\frac{7}{18}$.

9. Случайная величина X задана функцией распределения

$$F(x)=\begin{cases}0 & \text{при } x \le 2,\\ 0{,}5x-1 & \text{при } 2 < x \le 4,\\ 1 & \text{при } x > 4.\end{cases}$$

Найти вероятность того, что в результате испытания X примет значение, заключенное в интервале (2, 3).

Ответ: 0,5.

10. Найти математическое ожидание и дисперсию случайной величины X, заданной функцией распределения

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

Ответ: $M(X) = 2/3$; $D(x) = 1/18$.

ЗАКЛЮЧЕНИЕ

Изложенный материал позволяет, не обращаясь ко многим источникам, не только самостоятельно изучить объемный материал, но и проверить полученные знания при решении задач, которые приведены с ответами.

Поскольку материал изложен последовательно, то он может быть использован также и преподавателями при проведении занятий по указанным главам.

Подборка задач может быть использована как материал для контрольных и самостоятельных работ.

Авторы в течение нескольких лет вели дисциплину «Теория вероятностей» на многих специальностях дневного, вечернего и заочного обучения, поэтому акценты расставлены в представленном пособии исходя из практического опыта.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Бродский, Я. С. Статистика. Вероятность. Комбинаторика / Я. С. Бродский. – М. : Оникс, 2008. – 544 с.

2. Ивченко, Г. И. Введение в математическую статистику : учебник / Г. И. Ивченко, Ю. И. Медведев. – М. : Изд-во ЛКИ, 2010. – 600 с.

3. Лагутин, М. Б. Наглядная математическая статистика : учеб. пособие / М. Б. Лагутин. – М. : БИНОМ. Лаборатория знаний, 2009. – 472 с.

4. Письменный, Д. Т. Конспект лекций по теории вероятностей, математической статистике и случайным процессам / Д. Т. Письменный. – 3-е изд. – М. : Изд-во, 2008. – 288 с.

5. Попов, В. А. Руководство к решению задач по теории вероятностей и математической статистике / В. А. Попов, М. Х. Бренерман. – Казань : Изд-во КГУ, 2008. – 119 с.