Оглавление

Глава 1. Вероятность случайного события

§ 1. Элементы комбинаторики

Основные понятия комбинаторики

Комбинаторикой (комбинаторным анализом) называют раздел математики, в котором изучаются задачи выбора и расположения элементов из некоторого основного (обычно конечного) множества в соответствии с заданными правилами.

Многие комбинаторные задачи могут быть решены с помощью следующих двух важных правил, называемых правилами суммы и произведения.

Правило суммы. Если объект A может быть выбран из некоторой совокупности объектов m способами, а другой объект B – n способами, то выбрать либо A, либо B можно $m+n$ способами.

Правило произведения. Если объект A может быть выбран из некоторой совокупности объектов m способами и после каждого такого выбора объект B может быть выбран n способами, то пара объектов A и B в указанном порядке может быть выбрана $m \cdot n$ способами.

Эти правила распространяются на случай трёх и большего числа объектов.

Кроме этих правил известны формулы, помогающие решать некоторые типовые задачи, встречающиеся довольно часто. Комбинациям, фигурирующим в этих задачах, присвоены особые названия – размещения, перестановки и сочетания.

Определение 1. *Размещениями* из n элементов по m ($m \leq n$) элементов в каждом называются комбинации, содержащие по m различных элементов, выбранных из данных n элементов, которые отличаются либо составом, либо порядком входящих в них элементов.

Число всех возможных размещений из n элементов по m обозначается символом A_n^m и вычисляется по формуле:

$$\mathrm{A}_n^m = n \cdot (n-1) \cdot \ldots \cdot (n-m+1) = \frac{n\,!}{(n-m)!}\,.$$

Определение 2. *Перестановками* из n различных элементов называются комбинации, состоящие из одних и тех же n элементов и отличающиеся только порядком расположения элементов.

Число всех возможных перестановок из n элементов обозначается символом P_n и вычисляется по формуле:

$$\mathrm{P}_n = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1 = n\,!$$

Нетрудно заметить, что перестановки из n элементов являются частным случаем размещений из n элементов по m, когда $m = n$, т.е. число элементов в комбинации совпадает с числом имеющихся элементов. Таким образом, справедлива формула:

$$P_n = A_n^n .$$

Определение 3. *Сочетаниями* из n элементов по m ($m \le n$) в каждом называются комбинации, содержащие по m различных элементов, выбранных из данных n элементов, которые отличаются составом входящих в них элементов. При этом порядок расположения элементов не играет роли.

Число всех возможных сочетаний из n элементов по m обозначается символом C_n^m и вычисляется по формуле:

$$C_n^m = \frac{n\cdot(n-1)\cdot\ldots\cdot(n-m+1)}{1\cdot2\cdot3\cdot\ldots\cdot m} = \frac{n!}{m!(n-m)!}, \text{ т. е. } C_n^m = \frac{A_n^m}{P_m}.$$

Свойства сочетаний:

$C_n^0 = 1$,

$C_n^m = C_n^{n-m}$,

$C_{n+1}^m = C_n^m + C_n^{m-1}$,

$C_n^0 + C_n^1 + C_n^2 + \ldots + C_n^n = 2^n$.

Определение 4. Размещениями с повторениями из n элементов по m элементов в каждом называются комбинации, содержащие по m возможно повторяющихся элементов, выбранных из данных n элементов, которые отличаются либо составом, либо порядком входящих в них элементов.

Число всех возможных размещений с повторениями из n элементов по m обозначается символом $\overline{A}_n^m$ и вычисляется по формуле:

$$\overline{A}_n^m = n^m.$$

Задачи с решениями

Задача 1.1. На первом блюде лежат 8 апельсинов, на втором – 4 яблока. Сколькими способами можно выбрать один фрукт?

Решение. Один апельсин можно выбрать восемью способами, а одно яблоко – четырьмя. Один фрукт – это либо апельсин, либо яблоко. Воспользуемся правилом суммы: $m = 8$, $n = 4$; число способов выбора одного фрукта $m + n = 12$.

Ответ: 12.

Задача 1.2. Сколько трёхзначных чисел можно составить из цифр 0,1,2,3,4,5,9, если:

а) число записано разными цифрами?

б) цифры в записи числа могут повторяться?

Решение. а) Первую цифру в записи числа можно выбрать шестью способами (ноль не может быть первой цифрой), для выбора второй цифры, отличающейся от первой, существует 6 способов (ноль может быть второй цифрой), а для выбора третьей цифры остаётся 5 способов (две цифры из имеющихся семи поставлены на первое и второе места). Таким образом, согласно правилу произведения получаем $6 \cdot 6 \cdot 5 = 180$ способов составления трёхзначного числа, записанного разными цифрами.

б) Если цифры в записи числа могут повторяться, то имеем 6 способов выбора первой цифры и по 7 способов выбора каждой из следующих цифр. Количество таких чисел $6 \cdot 7 \cdot 7 = 294$.

Ответ: а) 180; б) 294.

Задача 1.3. Студенты изучают 6 различных дисциплин. Если ежедневно в расписание включается по 3 различных дисциплины, то сколькими способами могут быть распределены занятия в день?

Решение. Различные комбинации трёх дисциплин, выбранных из шести, составляют расписание на один день. При этом они различаются либо составом дисциплин, либо их порядком. Поэтому искомое число определяется формулой числа размещений: $A_6^3 = 6 \cdot 5 \cdot 4 = 120$.

Ответ: 120.

Задача 1.4. Сколько шестизначных чётных чисел можно составить из цифр 1,3,4,5,7,9, если в каждом из этих чисел ни одна цифра не повторяется?

Решение. Чтобы число было чётным, последняя его цифра (число единиц) должна быть чётной. Из заданных цифр только одна чётная – это 4. Поэтому последней цифрой искомого числа может быть только 4. Остальные пять цифр могут стоять на первых пяти местах в любом порядке. Значит, задача сводится к нахождению числа перестановок из пяти элементов: $P_5 = 5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120$.

Ответ: 120.

Задача 1.5. Сколько шестизначных чётных чисел можно составить из цифр 1,3,4,5, если цифры в записи числа могут повторяться?

Решение. Чтобы число было чётным, последняя его цифра (число единиц) должна быть чётной. Из заданных цифр только одна чётная – это 4. Поэтому последней цифрой искомого числа может быть только 4. Остальные пять цифр могут быть любыми из предложенных, причём могут повторяться. Значит, задача сводится к нахождению числа размещений с повторениями из четырёх элементов по пять в каждом: $\bar{A}_4^5 = 4^5 = 1024$.

Ответ: 1024.

Задача 1.6. Сколькими способами можно выбрать 3 книги из 10 книг по математике, имеющихся в библиотеке?

Решение. Искомое число способов равно числу сочетаний из 10 элементов по 3 элемента в каждом, так как интересующие нас комбинации из трёх книг отличаются друг от друга только содержащимися в них книгами, а порядок расположения книг в этих комбинациях роли не играет. Следовательно, находим: $C_{10}^3 = \frac{10 \cdot 9 \cdot 8}{1 \cdot 2 \cdot 3} = 120$.

Ответ: 120.

Задача 1.7. Сколько трёхзначных чётных чисел можно составить из цифр 0,1,2,3,4,5,6, если цифры в записи числа могут повторяться?

Решение. При составлении трёхзначного числа из данных цифр в качестве первой цифры (числа сотен) можно взять любую цифру, кроме 0. Значит, есть шесть возможностей выбора первой цифры. В качестве второй цифры (числа десятков) можно выбрать любую из данных в условии цифр. Значит, есть семь возможностей выбора второй цифры. В качестве последней цифры (числа единиц) можно взять любую из цифр 0,2,4,6. Значит, есть четыре возможности выбора третьей цифры. Следовательно, согласно правилу произведения находим количество способов составления числа, удовлетворяющего условию задачи: $6 \cdot 7 \cdot 4 = 168$.

*Ответ:*168.

Задача 1.8. Сколько различных чисел можно составить из цифр 4 и 5, если количество цифр в записи числа не более пяти и не менее трёх?

Решение. По условию задачи количество цифр в записи числа не более пяти и не менее трёх. Значит, их либо три, либо четыре, либо пять.

Если число, записанное четвёрками и пятёрками, содержит три цифры, то таких чисел будет: $\overline{A}_2^3 = 2^3 = 8$.

Если число, записанное четвёрками и пятёрками, содержит четыре цифры, то таких чисел будет: $\overline{A}_2^4 = 2^4 = 16$.

Если число, записанное четвёрками и пятёрками, содержит пять цифр, то таких чисел будет: $\overline{A}_2^5 = 2^5 = 32$.

Следовательно, согласно правилу суммы, находим количество способов составления числа, удовлетворяющего условию задачи: $8+16+32 = 56$.

Ответ: 56.

Задачи

1.1.Вычислите:

а) A_n^0; б) A_n^1; в) A_n^n; г) A_5^3; д) P_5; е) C_5^3; ж) $\overline{A}_5^3$.

1.2. Вычислите:

а) C_n^0; б) C_n^1; в) C_n^n; г) C_7^2; д) A_7^2; е) $\overline{A}_7^2$; ж) P_7.

з) Докажите, что $C_n^m = C_n^{n-m}$ и вычислите C_{30}^{28}.

1.3. Комиссия состоит из председателя, его заместителя и ещё пяти человек. Сколькими способами члены комиссии могут распределить между собой обязанности?

1.4. В розыгрыше первенства по футболу принимают участие 18 команд. Сколькими способами могут распределиться золотая, серебряная и бронзовая медали, если любая команда может получить только одну медаль?

1.5. В группе из 10 человек надо выбрать трёх для уборки помещения. Сколько можно сделать различных вариантов такого выбора?

1.6. В студенческой группе 25 человек. Из них надо выбрать четверых для участия в студенческой конференции. Сколькими способами можно это сделать?

1.7. Сколькими способами можно расставить на одной книжной полке 7 книг разных авторов?

1.8. Сколькими способами можно рассадить компанию из шести человек за столом, накрытым шестью приборами?

1.9. Во взводе 3 сержанта и 30 солдат. Сколькими способами можно выделить одного сержанта и трёх солдат для патрулирования?

1.10. Хоккейная команда состоит из двух вратарей, семи защитников и десяти нападающих. Сколькими способами тренер может образовать стартовую шестёрку, состоящую из вратаря, двух защитников и трёх нападающих?

1.11. Обычно наибольшее количество очков на одной кости игры домино равно 12. Сколько костей содержала бы игра, если бы это число равнялось 18?

1.12. Сколько костей содержала бы игра домино, если бы наибольшее количество очков на одной кости равнялось 20?

1.13. Сколько различных десятизначных чисел можно написать, используя цифры 1 и 2?

1.14. Сколько различных восьмизначных чисел можно написать, используя цифры 0,1,2?

1.15. На пять сотрудников выделены три премии. Сколькими способами их можно распределить, если:

а) размер премий различен?

б) все премии одинаковые?

1.16. В классе 30 учащихся. Сколькими способами из них можно выделить двух человек для дежурства по школе, если:

а) один из них должен быть старшим?

б) старшего быть не должно?

1.17. Сколько диагоналей имеет выпуклый 12-угольник?

1.18. Сколько диагоналей имеет выпуклый 17-угольник?

1.19. Сколько существует двузначных чисел, записанных различными нечётными цифрами?

1.20. Сколько существует трёхзначных чисел, записанных различными нечётными цифрами?

1.21. Сколькими способами можно разложить пять различных писем по пяти различным конвертам, если в каждый конверт кладётся только одно письмо?

1.22. В розыгрыше первенства по футболу было сыграно 153 матча. Каждые две команды встречались между собой один раз. Сколько команд участвовало в розыгрыше первенства?

1.23. Из двух математиков и восьми экономистов надо составить комиссию из восьми человек. Сколькими способами может быть составлена комиссия, если в неё должен входить хотя бы один математик?

1.24. Из группы в 15 человек должны быть выделены бригадир и 4 члена бригады. Сколькими способами это можно сделать?

1.25. Сколько существует пятизначных чисел, которые одинаково читаются слева направо и справа налево?

1.26. Буквы азбуки Морзе представляют собой набор точек и тире. Сколько букв может быть в азбуке Морзе, если буква не должна содержать более четырёх знаков?

1.27. Сколько различных двузначных чисел можно образовать из цифр 1,2,3,4, если:

а) в каждом числе цифры не повторяются?

б) цифры в числе могут повторяться?

1.28. Сейф запирается на замок, состоящий из пяти дисков, на каждом из которых изображены цифры 0,1,2,3,4,5,6,7,8,9. Замок открывается, если на дисках набрана определённая комбинация цифр. Хватит ли десяти дней на открытие сейфа, если "рабочий день" продолжается 13 часов, а на набор одной комбинации цифр уходит 5 секунд?

Ответы

1.1. а) 1; б) n; в) $n!$; г) 60; д) 120; е) 10; ж) 125; **1.2.** а) 1; б) n; в) 1; г) 21; д) 42; е) 49; ж) 5040; з) 435; **1.3.** 42; **1.4.** 4896; **1.5.** 120; **1.6.** 12650; **1.7.** 5040; **1.8.** 720; **1.9.** 12180; **1.10.** 5040; **1.11.** 55; **1.12.** 66; **1.13.** 1024; **1.14.** 4374; **1.15.** а) 60; б) 10.**1.16.** а) 870; б) 435; **1.17.** 54; **1.18.** 119; **1.19.** 20; **1.20.** 60; **1.21.** 120; **1.22.** 18; **1.23.** 44; **1.24.** 15015; **1.25.** 900; **1.26.** 30; **1.27.** а)12; б)16; **1.28.** Может не хватить времени, так как всего возможных комбинаций 100000, а за 10 дней работы по 13 ч в день можно набрать только 93600 комбинаций.

§2. Случайные события и их классификация. Алгебра событий. Вероятность события. Теоремы сложения и умножения вероятностей

Случайные события

Определение 1. *Испытанием* (или *опытом*) называется осуществление некоторой совокупности определённых условий.

Определение 2. *Событием* называется любой результат испытания.

Определение 3. Событие называется *случайным* (обозначается прописными латинскими буквами: $A, B, C, \ldots, A_1, A_2, A_3\ldots$), если в данном испытании оно может или произойти, или не произойти.

Определение 4. Событие называется *достоверным* (обозначается E), если в данном испытании оно обязательно произойдёт.

Определение 5. Событие называется *невозможным* (обозначается $\overline{E}$), если в данном испытании оно никогда не произойдёт.

Определение 6. События называются *несовместными* в данном испытании, если они не могут наступить одновременно. В противном случае события называются *совместными*.

Определение 7. Событие B называется *независимым* от события A, если наступление события A не влияет на наступление или ненаступление события B. В противном случае события A и B называются зависимыми.

Определение 8. События называются *равновозможными*, если есть основание считать, что ни одно из них не является более возможным, чем остальные.

Определение 9. Говорят, что события $A_1, A_2, \ldots, A_n$ образуют *полную группу событий*, если в результате испытания обязательно наступает хотя бы одно из них.

Определение 10. Говорят, что события $A_1, A_2, \ldots, A_n$ образуют *полную группу равновозможных попарно несовместных событий*, если в данном испытании события $A_1, A_2, \ldots, A_n$ являются равновозможными и любые два из них – несовместные. Такие события будут называться *элементарными событиями* (или *случаями, исходами*).

Определение 11. Элементарное событие $A_i (i = \overline{1; n})$ называется благоприятствующим событию A , если его наступление влечёт за собой наступление события A.

Определение 12. Событие, обозначаемое $\bar{A}$, называется *противоположным событием* по отношению к событию A, если наступление одного из них в результате данного испытания исключает наступление другого.

Алгебра событий

Определение 13. *Суммой* (или *объединением*) событий A и B называется такое событие, обозначаемое $A+B$, которое состоит в наступлении хотя бы одного из событий A или B.

Определение 14. *Произведением* (или *совмещением*) событий A и B называется такое событие, обозначаемое $A \cdot B$, которое состоит в одновременном наступлении и события A, и события B.

Замечание 1. Если события $A_1, A_2, \ldots, A_n$ образуют полную группу попарно несовместных событий, то справедливы равенства:

$$A_1 + A_2 + \ldots + A_n = E,$$

$$A_i \cdot A_j = \overline{E} \ (i \neq j).$$

Замечание 2. Поскольку события A и $\bar{A}$ образуют полную группу и несовместны, для них справедливы равенства:

$$A + \bar{A} = E,$$
$$A \cdot \bar{A} = \overline{E}.$$

Вероятность события

Пусть для данного испытания события $A_1, A_2, \dots, A_n$ образуют полную группу равновозможных попарно несовместных событий (являются элементарными событиями).

Определение 15. *Вероятностью случайного события A* в данном испытании называется число, обозначаемое $P(A)$ и вычисляемое по формуле:

$$P(A) = \frac{m}{n}, \tag{1}$$

где n – число всех возможных элементарных событий рассматриваемого испытания; m – число благоприятствующих событию A.

Замечание 3. Ситуация, когда полную группу составляют равновозможные события, называется *классической*. Поэтому определение вероятности (1), опирающееся на такое условие, называется *классическим определением вероятности.*

Замечание 4. Нетрудно видеть, что в формуле (1) числа m, n связаны неравенствами:

$$0 \le m \le n.$$

Поэтому вероятность любого события A удовлетворяет неравенству:

$$0 \le P(A) \le 1.$$

Причём, если $A = E$ – достоверное событие, то $m = n$ и $P(E) = 1$; если $A = \overline{E}$ – невозможное событие, то $m = 0$ и $P(\overline{E}) = 0$.

Теоремы сложения и умножения вероятностей

Теорема 1 (теорема суммы несовместных событий). Вероятность наступления одного из двух несовместных событий (безразлично какого) равна сумме вероятностей этих событий:

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

Следствие 1. Вероятность наступления одного из нескольких попарно несовместных событий (безразлично какого) равна сумме вероятностей этих событий:

$$P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n).$$

Следствие 2. Вероятность события A равна единице минус вероятность его противоположного события $\bar{A}$:

$$P(A) = 1 - P(\bar{A}).$$

Теорема 2 (вероятность суммы совместных событий). Вероятность наступления хотя бы одного из двух совместных событий равна сумме вероятностей этих событий минус вероятность их совместного наступления:

$$P(A + B) = P(A) + P(B) - P(A \cdot B).$$

Определение 1. *Условной вероятностью* $P_A(B)$ называется вероятность события B при условии, что событие A уже наступило.

Теорема 3 (вероятность произведения двух независимых событий). Вероятность совместного наступления двух независимых событий равна произведению вероятностей этих событий:

$$P(A \cdot B) = P(A) \cdot P(B).$$

Следствие 3. Вероятность совместного наступления нескольких независимых событий равна произведению вероятностей этих событий:

$$P\,(A_1 \cdot A_2 \cdot \ldots \cdot A_n) = P\,(A_1) \cdot P\,(A_2) \cdot \ldots \cdot P\,(A_n).$$

Теорема 4 (вероятность произведения двух зависимых событий). Вероятность совместного наступления двух зависимых событий равна произведению вероятности одного из них на условную вероятность другого, вычисленную в предположении, что первое событие наступило:

$$P(A \cdot B) = P(A) \cdot P_A(B).$$

Следствие 4. Вероятность совместного наступления нескольких зависимых событий равна произведению вероятности одного из них на условные вероятности всех остальных, причём условная вероятность каждого последующего события вычисляется в предположении, что все предыдущие события уже наступили:

$$P(A_1 \cdot A_2 \cdot \ldots \cdot A_n) = P(A_1) \cdot P_{A_1}(A_2) \cdot P_{A_1 \cdot A_2}(A_3) \cdot \ldots \cdot P_{A_1 \cdot A_2 \cdot \ldots \cdot A_{n-1}}(A_n).$$

Задачи с решениями

Задача 2.1. Монета подбрасывается два раза.

а) Опишите полную группу возможных элементарных событий.

б) Если событие А – выпало не менее одного "орла", В – выпало не менее одной "решки", укажите, что собой представляют события:

$$\bar{A}, \bar{B}, A + B, \ A \cdot B?$$

Решение. В данной задаче испытанием является подбрасывание монеты дважды.

а) Обозначим события:

C_1 – при первом подбрасывании выпал "орёл", при втором – "решка",

C_2 – при первом подбрасывании выпала "решка", при втором – "орёл",

C_3 – оба раза выпал "орёл",

C_4 – оба раза выпала "решка".

Тогда перечисленные события C_1, C_2, C_3,C_4 образуют полную группу, так как при двух подбрасываниях монеты обязательно произойдёт одно из них. Значит, справедливо равенство:

$$C_1 + C_2 + C_3 + C_4 = E.$$

Кроме того, никакие два из указанных событий не могут наступить одновременно. Следовательно, имеет место равенство:

$$C_i \cdot C_j = \bar{E} \text{ при } i \neq j.$$

Таким образом, указанные события попарно несовместны. Причём наступление любого из событий C_1, C_2, C_3,C_4 не имеет преимущества перед остальными, а значит, эти события являются равновозможными.

Таким образом, события C_1, C_2, C_3,C_4 образуют полную группу равновозможных попарно несовместных событий. Следовательно, они являются в данном испытании полной группой элементарных событий.

б) Так как $\bar{A}$ – не выпало ни одного "орла", то $\bar{A} = C_4$ – оба раза выпала "решка". Аналогично, $\overline{B}$ – не выпало ни одной "решки", следовательно, $\overline{B} = C_3$ – оба раза выпал "орёл". А так как A означает, что выпадает не менее одного раза "орёл", то $A = C_1 + C_2 + C_3$. Аналогично заключаем:
$B = C_1 + C_2 + C_4$. Следовательно, по определению суммы и произведения событий получаем:

$$A + B = (C_1 + C_2 + C_3) + (C_1 + C_2 + C_4) = C_1 + C_2 + C_3 + C_4 = E ,$$

$$A \cdot B = (C_1 + C_2 + C_3) \cdot (C_1 + C_2 + C_4) = C_1 + C_2 .$$

Ответ: а) C_1, C_2, C_3, C_4; б) $\overline{A} = C_4$, $\overline{B} = C_3$, $A + B = E, A \cdot \text{B} = C_1 + C_2$.
.

Задача 2.2. В ящике находится 10 шаров: 3 белых и 7 чёрных. Из ящика наугад выбирается один шар. Какова вероятность того, что этот шар:

а) белый,

б) чёрный?

Решение. В данной задаче полную группу элементарных событий составляют 10 событий, так как выбор любого одного шара можно осуществить 10 способами. Из этих событий только 3 благоприятствуют выбору белого шара и 7 – выбору чёрного. Поэтому, если A – выбор белого шара, то $P(A) = \frac{3}{10} = 0{,}3$; если B – выбор чёрного шара, то $P(B) = \frac{7}{10} = 0{,}7$.

Ответ: а) 0,3 ; б) 0,7.

Задача 2.3. Из пяти карточек с буквами А, Б, В, Г, Д наугад выбираются одна за другой три карточка и располагаются в ряд (в порядке появления) слева направо. Какова вероятность, что получится слово "ДВА"?

Решение. Выбор трёх карточек из имеющихся пяти можно осуществить A_5^3 способами, так как порядок карточек имеет значение в данной задаче. Вычисляем: $\text{A}_5^3 = 5{\cdot}4{\cdot}3 = 60$.

Значит, число всех возможных элементарных событий $n = 60$. Из этих событий только одно благоприятствует событию – получению слова "ДВА", следовательно, $m = 1$. Итак, $P = \frac{1}{60}$.

Ответ: $\frac{1}{60}$.

Задача 2.4. В ящике 10 шаров: 6 белых и 4 чёрных. Из ящика наугад вынимают два шара. Какова вероятность того, что

а) оба шара белые?

б) оба шара чёрные?

в) один шар белый, другой чёрный?

Решение. Число выбора двух шаров из десяти имеющихся определяется числом всевозможных сочетаний из 10 по 2: $C_{10}^2 = \frac{10 \cdot 9}{1 \cdot 2} = 45$.

Значит, полную группу элементарных событий рассматриваемого испытания (выбор двух шаров из десяти, находящихся в ящике) составляют 45 событий. Следовательно, $n = 45$.

а) Если из элементарных событий рассматривать только те, которые состоят в выборе двух белых шаров, то находим $m = C_6^2 = \frac{6 \cdot 5}{1 \cdot 2} = 15$.

Следовательно, вероятность того, что оба шара будут белыми, вычисляется по формуле:

$$p_1 = \frac{m}{n} = \frac{15}{45} = \frac{1}{3}.$$

б) Если рассматривать событие – выбор двух чёрных шаров, то число благоприятствующих ему элементарных событий равно:

$$m = C_4^2 = \frac{4 \cdot 3}{1 \cdot 2} = 6.$$

Значит, вероятность выбора двух чёрных шаров вычисляется по формуле:

$$p_2 = \frac{m}{n} = \frac{6}{45} = \frac{2}{15}.$$

в) Если рассматривать событие – выбор одного белого и одного чёрного шаров, то для него число благоприятствующих элементарных событий равно:

$$m = C_6^1 \cdot C_4^1 = 6 \cdot 4 = 24.$$

Значит, вероятность выбора одного белого и одного чёрного шаров вычисляется по формуле:

$$p_3 = \frac{m}{n} = \frac{24}{45} = \frac{8}{15}.$$

Ответ: а) $\frac{1}{3}$, б) $\frac{2}{15}$, в) $\frac{8}{15}$.

Задача 2.5. Стрелок стреляет по мишени, разделённой на четыре области. Вероятность попадания в первую область 0,4, во вторую – 0,3. Найдите вероятность того, что стрелок при одном выстреле попадёт либо в первую область, либо во вторую.

Решение. Обозначим события:

A – стрелок попадает в первую область, B – стрелок попадает во вторую область. Эти события несовместны, так как они не могут наступить одновременно (попадание пули в одну область мишени исключает её попадание в другую область). Поэтому воспользуемся теоремой 1 (вероятность суммы несовместных событий), откуда находим:

$$P(A+B) = P(A) + P(B) = 0{,}4 + 0{,}3 = 0{,}7.$$

Ответ: 0,7.

Задача 2.6. Какова вероятность извлечь из колоды в 52 карты фигуру (валет, дама, король, туз) любой масти или карту трефовой масти?

Решение. Обозначим события:

A – извлечение из колоды карты – фигуры, B – извлечение из колоды карты трефовой масти.

Необходимо найти вероятность суммы этих событий. События A и B совместны, так как они могут наступить одновременно, если будет извлечена карта – фигура трефовой масти. Поэтому для подсчёта вероятности суммы этих событий используем теорему 2 (вероятность суммы совместных событий):

$$P(A+B) = P(A) + P(B) – P(A \cdot B).$$

В рассматриваемой задаче $P(A) = \frac{16}{52} = \frac{4}{13}$, так как всего элементарных исходов 52, что равно числу карт в колоде, из них 16 благоприятствуют событию A, что равно числу карт – фигур в колоде. Аналогично вычисляем: $P(B) = \frac{13}{52}$ (в колоде 13 карт трефовой масти). $P(A \cdot B) = \frac{4}{52} = \frac{1}{13}$ (в колоде 4 карты – фигуры трефовой масти).

Итак, находим: $P(A+B) = \frac{4}{13} + \frac{13}{52} - \frac{1}{13} = \frac{25}{52}$.

Ответ: $\frac{25}{52}$.

Задача 2.7. Два орудия стреляют по одной цели. Вероятность попадания для первого орудия равна 0,6, для второго вероятность попадания равна 0,5. Какова вероятность того, что в цель попадут оба орудия?

Решение. Обозначим события:

A – попадание в цель первого орудия, B – попадание в цель второго орудия.

Отметим, что A и B – события независимые, то есть наступление одного из них не влияет на наступление или ненаступление другого. Поэтому воспользуемся теоремой 3 (вероятность произведения независимых событий):

$$P(A \cdot B) = P(A) \cdot P(B) = 0{,}6 \cdot 0{,}5 = 0{,}3.$$

Ответ: 0,3.

Задача 2.8. Для Московской области среднее число дождливых дней в августе равно 15. Какова вероятность, что первые два дня августа не будут дождливыми?

Решение. Обозначим события:

A – 1 августа не будет дождя, B – 2 августа не будет дождя.

Необходимо рассмотреть событие $A \cdot B$ – 1 и 2 августа не будет дождя. В данной задаче $P(A) = \frac{16}{31}$, так как в августе 31 день, а не дождливых дней из них 31 – 15 = 16.

При вычислении $P(B)$ результат зависит от того, будет ли дождь 1-го августа. Следовательно, необходимо найти условную вероятность $P_A(B)$ – вероятность того, что 2-го августа не будет дождя в предположении, что 1 августа –день без дождя. Тогда получаем: $P_A(B) = \frac{15}{30} = \frac{1}{2}$,

Так как в августе осталось 30 дней, начиная со 2 августа, из них не дождливых дней осталось 15 (ведь один не дождливый день пришёлся по предположению на 1 августа). Итак, по теореме 4 (вероятность произведения зависимых событий) получаем:

$$P(A \cdot B) = P(A) \cdot P_A(B) = \frac{16}{31} \cdot \frac{1}{2} = \frac{8}{31}.$$

Ответ: $\frac{8}{31}$.

Задача 2.9. В ящике 10 деталей, среди которых 6 стандартных. Какова вероятность того, что среди трёх наугад взятых деталей окажется хотя бы одна стандартная?

Решение. События "среди взятых деталей окажется хотя бы одна стандартная" и "среди взятых деталей нет ни одной стандартной" – противоположные события, так как наступление одного из этих событий исключает наступление другого.

Обозначим:

A – среди трёх взятых деталей есть хотя бы одна стандартная, $\overline{A}$ – среди трёх взятых деталей нет ни одной стандартной.

По следствию 2 из теоремы 1 известно, что $P(A) = 1 - P(\overline{A})$. Найдём $P(\overline{A})$. Общее число элементарных событий в этой задаче – это число способов выбора трёх деталей из десяти, находящихся в ящике:

$$n = C_{10}^3 = \frac{10 \cdot 9 \cdot 8}{1 \cdot 2 \cdot 3} = 120.$$

Число нестандартных деталей равно 10 – 6 = 4. Число элементарных исходов, благоприятствующих событию $\overline{A}$: $m = C_4^3 = C_4^{4-3} = C_4^1 = 4$.

Тогда $P(\overline{A}) = \frac{m}{n} = \frac{4}{120} = \frac{1}{30}$.

Следовательно, $P(A) = 1 - P(\overline{A}) = 1 - \frac{1}{30} = \frac{29}{30}$.

Ответ: $\frac{29}{30}$.

Задачи

2.1. Два шахматиста играют одну партию. Событие A – выиграет первый игрок, B – выиграет второй игрок. Какое событие следует добавить к указанной совокупности, чтобы получилась полная группа событий?

2.2. Подбрасывают две монеты. Событие A – выпадут "орёл" и "решка", B – выпадут две "решки". Какое событие следует добавить к указанной совокупности, чтобы получилась полная группа событий?

2.3. События: A – хотя бы один из проверяемых приборов бракованный, B – все проверяемые приборы доброкачественные. Что означают события: а) $A+B$; б) $A \cdot B$; в) $\overline{A}$; г) $\overline{B}$?

2.4. Мишень состоит из трёх кругов, ограниченных концентрическими окружностями с радиусами r_1, r_2, r_3, причём $r_1<r_2<r_3$. Событие A_1 – попадание в круг радиуса r_1, A_2 – попадание в круг радиуса r_2, A_3 – попадание в круг радиуса r_3. Что означают события: а) A_1+A_2+ $+A_3$, б) $A_1 \cdot A_2 \cdot A_3$, в) $\overline{A_1} \cdot A_3$?

2.5. В ящике лежат 15 шаров: 5 чёрных и 10 белых. Наугад вынимают один шар. Какова вероятность того, что этот шар а) белый? б) чёрный? в) красный?

2.6. Из колоды (36 карт) извлекают одну карту. Какова вероятность, что извлечённая карта окажется а) тузом; б) картой – фигурой (валет, дама, король, туз)?

2.7. В ящике лежат 15 шаров: 5 чёрных и 10 белых. Наугад вынимают два шара. Какова вероятность того, что а) оба белые? б) оба чёрные? в) вынуты шары разных цветов?

2.8. В ящике 5 белых и 4 чёрных шара. Наугад вынимают два шара. Какое событие более вероятно: A – вынуты шары одного цвета, B – вынуты шары разных цветов?

2.9. В мастерскую для ремонта поступило 10 часов марки "Слава". Известно, что 6 штук из них нуждаются в общей чистке механизма. Мастер берёт первые попавшиеся 5 часов. Какова вероятность, что двое часов из взятых мастером нуждаются в общей чистке механизма?

2.10. В автобусе было 4 девушки и 5 юношей. На остановке вышли 6 человек. Какова вероятность того, что среди них 3 девушки и 3 юноши?

2.11. Брошены две игральные кости. Какова вероятность того, что сумма очков на выпавших гранях является чётным числом, не превышающим шести?

2.12. Брошены две игральные кости. Какова вероятность того, что сумма выпавших очков окажется равна шести?

2.13. На пяти одинаковых карточках написаны буквы Т,У,К,Б,Е. Карточки тщательно перемешаны. Извлекаются наудачу поочерёдно по одной карточке и укладываются слева направо. Какова вероятность того, что получится слово "БУКЕТ"?

2.14. Ребёнок, не умеющий читать, рассыпал слово "КОЛОБОК", составленное из букв разрезной азбуки, и собрал вновь. Какова вероятность того, что ребёнок собрал слово верно?

2.15. В группе 25 студентов. Из них по математике отлично успевают 5 человек, хорошо – 12, удовлетворительно – 6 и слабо – 2. Преподаватель, не знакомый с группой, вызывает одного из студентов. Какова вероятность того, что вызванный студент окажется отличником или хорошо успевающим?

2.16. В корзине 5 красных яблок, 6 жёлтых и 4 зелёных. Наугад вынимают одно яблоко. Какова вероятность того, что оно окажется жёлтым или зелёным?

2.17. Из колоды (36 карт) извлекают карту. Какова вероятность того, что это будет карта бубновой масти или туз?

2.18. Подбрасывают игральную кость. Какова вероятность того, что количество выпавших очков будет кратно двум или трём?

2.19. Буквы, составляющие слово "ОДЕССА", написаны на шести карточках. Карточки перемешаны и положены в пакет. Чему равна вероятность того, что, вынимая карточки по одной и записывая соответствующие буквы в ряд слева направо, мы получим слово "САД", если

а) после извлечения карточка снова возвращается в пакет и все они снова перемешиваются?

б) извлечённая карточка в пакет не возвращается?

2.20. Буквы, составляющие слово "КОЛОБОК", написаны на карточках, которые сложены в пакет и перемешаны. Чему равна вероятность того, что, вынимая карточки по одной и записывая соответствующие буквы в ряд слева направо, мы получим слово "БЛОК", если

а) после извлечения карточка снова возвращается в пакет и все они снова перемешиваются?

б) извлечённая карточка в пакет не возвращается?

2.21. Рабочий обслуживает три станка, работающие независимо друг от друга. Вероятность того, что в течение часа первый станок не потребует внимания рабочего, равна 0,9, для второго станка такая вероятность равна 0,8, для третьего станка – 0,85. Какова вероятность того, что в течение часа

а) ни один из станков не потребует внимания рабочего?

б) какой-нибудь один станок потребует внимания рабочего?

в) хотя бы один станок потребует внимания рабочего?

2.22. Три стрелка стреляют по одной мишени с вероятностями попадания $p_1 = 0{,}6$, $p_2 = 0{,}8$, $p_3 = 0{,}7$. Какова вероятность того, что

а) в мишень попадёт только один из них?

б) в мишень попадут двое?

в) мишень будет поражена?

2.23. В электрическую цепь последовательно включены три элемента, работающие независимо один от другого. Вероятности отказов первого, второго и третьего элементов, соответственно, равны 0,1, 0,15 и 0,2. Какова вероятность того, что тока в цепи не будет (т. е. откажет хотя бы один элемент)?

2.24. Экспедиция издательства отправляет газеты в два почтовых отделения. Вероятность своевременной доставки газет в каждое почтовое отделение равна 0,9. Какова вероятность того, что оба почтовых отделения получат газеты

а) с опозданием?

б) вовремя?

2.25. Среди 10 дружинников 3 девушки и 7 юношей. Требуется путём жеребьёвки выбрать на дежурство трёх дружинников. Чему равна вероятность того, что при извлечении одного за другим трёх жребиев окажутся выбранными 3 юноши?

2.26. Два игрока поочерёдно извлекают шары (без возвращения) из урны, содержащей 1 белый и 4 чёрных шара. Выигрывает тот, кто первым вытащит белый шар. Найдите вероятности выигрыша для каждого игрока.

2.27. Чему равна сумма вероятностей противоположных событий? Какое наибольшее и какое наименьшее значения может принимать произведение вероятностей противоположных событий?

2.28. Многолетними наблюдениями установлено, что в данном районе в сентябре 10 дней бывают дождливыми. Совхоз должен в течение первых трёх дней сентября выполнить определённую работу. Какова вероятность того, что ни один из трёх этих дней не будет дождливым?

2.29. В конверте среди 20 фотографий находится одна разыскиваемая. Из конверта наудачу взяты 4 фотографии. Какова вероятность того, что среди них окажется разыскиваемая?

2.30. В записанном номере телефона стёрлись последние три цифры. Найдите вероятности событий:

а) стёрлись различные цифры, отличные от 1, 3 и 5;

б) стёрлись одинаковые цифры;

в) две из стёршихся цифр совпадают.

Ответы

2.1. С – ничейный исход; **2.2.** С – выпало два "орла"; **2.3.** а) $A + B = E$, б) $A \cdot B = \overline{E}$, в) $\overline{A} = B$, г) $\overline{B} = A$; **2.4.** а) A_3; б) A_1; в) попадание в кольцо, ограниченное окружностями с радиусами r_1 и r_3; **2.5.** а) $\frac{2}{3}$; б) $\frac{1}{3}$; в) 0; **2.6.** а) $\frac{1}{9}$; б) $\frac{4}{9}$; **2.7.** а) $\frac{3}{7}$; б) $\frac{2}{21}$; в) $\frac{10}{21}$; **2.8.** $P(A)<P(B)$, так как $P(A)=\frac{4}{9}$, $P(B)=\frac{5}{9}$; **2.9.** $\frac{5}{11}$; **2.10.** $\frac{10}{21}$; **2.11.** $\frac{1}{4}$; **2.12.** $\frac{5}{36}$; **2.13.** $\frac{1}{120}$; **2.14.** $\frac{1}{420}$; **2.15.** $\frac{17}{25}$; **2.16.** $\frac{2}{3}$; **2.17.** $\frac{1}{3}$; **2.18.** $\frac{2}{3}$; **2.19.** а) $\frac{1}{208}$; б) $\frac{1}{60}$; **2.20.** а) $\frac{6}{2401}$; б) $\frac{1}{140}$; **2.21.** а) 0,612; б) 0,329; в) 0,388; **2.22.**

а) 0,188; б) 0,451; в) 0,976; **2.23.** 0,388; **2.24.** а) 0,01; б) 0,81; **2.25.** $\frac{7}{24}$; **2.26.** $\frac{3}{5}$; $\frac{2}{5}$; **2.27.** 1; $\frac{1}{4}$ – наибольшее; **2.28.** $\frac{57}{203}$; **2.29.** 0,2; **2.30.** а) 0,21; б) 0,01; в) 0,27.

§3. Формула полной вероятности. Повторные испытания. Формула Бернулли

Формула полной вероятности

Пусть событие A может наступить при условии наступления одного из несовместных событий H_1, H_2, … H_n, которые образуют полную группу. Поскольку заранее неизвестно, какое из этих событий наступит, их называют гипотезами. Пусть известны вероятности этих событий (гипотез) и условные вероятности события A при условии наступления каждого из них. Как найти вероятность события A? Ответ на этот вопрос даёт следующая теорема.

Теорема 1 (формула полной вероятности). Вероятность события A, которое может наступить лишь при условии наступления одного из попарно несовместных событий H_1, H_2, … H_n, образующих полную группу, равна сумме произведений вероятностей каждого из этих событий на соответствующую условную вероятность события A:

$$P(A) = P(H_1) \cdot P_{H_1}(A) + P(H_2) \cdot P_{H_2}(A) + \ldots + P(H_n) \cdot P_{H_n}(A).$$

Замечание 1. Следствием формулы полной вероятности является формула Байеса (по имени английского математика, который её вывел; опубликована в 1764 г.). Она позволяет переоценить вероятность гипотезы H_i, принятую до опыта, по результатам уже проведённого опыта, т.е. вычислить условную вероятность гипотезы H_i при условии наступления события A.

Теорема 2 (формула Байеса или теорема гипотез). Пусть попарно несовместные события H_1, H_2, … H_nобразуют полную группу. Тогда условная вероятность события $H_i (i = \overline{1,n})$ при условии, что событие A наступило, задаётся формулой:

$$P_A(H_i) = \frac{P(H_i) \cdot \mathrm{P}_{H_i}(A)}{P(A)} = \frac{P(H_i) \cdot \mathrm{P}_{H_i}(A)}{\mathrm{P}(\mathrm{H}_1) \cdot P_{\mathrm{H}_1}(A) + \mathrm{P}(\mathrm{H}_2) \cdot P_{\mathrm{H}_2}(A) + \ldots + \mathrm{P}(\mathrm{H}_n) \cdot P_{\mathrm{H}_n}(A)}.$$

Повторные испытания. Формула Бернулли

Если производится несколько испытаний, причём вероятность наступления события A в каждом испытании не зависит от исходов других испытаний, то такие испытания называют независимыми относительно события A.

В разных независимых испытаниях событие A может иметь либо различные вероятности, либо одну и ту же вероятность. Мы будем рассматривать такие независимые испытания, в которых события A имеют одну и ту же вероятность.

Теорема 3 (формула Бернулли). Пусть в серии из n одинаковых независимых испытаний в каждом испытании может наступить либо событие A с вероятностью p, либо событие $\bar{A}$ с вероятностью $q = 1 - p$. Тогда вероятность $P_n(m)$ того, что в этой серии испытаний событие A наступит ровно m раз ($m \leq n$), вычисляется по формуле Бернулли:

$$P_n(m) = C_n^m \cdot p^m \cdot q^{n-m}\text{ , где } C_n^m = \frac{n!}{m!(n-m)!}$$

Формулы Лапласа

Эти формулы дают приближенное значение вероятности наступления события A определённое число раз в серии из n независимых испытаний, если число n достаточно велико. Пусть p ($0 < p < 1$) – вероятность события A в каждом испытании, $q = 1 - p$ – вероятность события $\bar{A}$.

Теорема 4 (локальная формула Лапласа). Вероятность $P_n(m)$ наступления события A ровно m раз в серии из n одинаковых независимых испытаний приближённо вычисляется по формуле Лапласа:

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \cdot \varphi\left(\frac{m-np}{\sqrt{npq}}\right),$$

где $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$.

Замечание 2. Имеются таблицы, в которых помещены значения функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$ (в приложении 2 табл. П 2.1), соответствующие положительным значениям аргумента x. Для отрицательных значений аргумента пользуются теми же таблицами, так как функция $\varphi(x)$ является чётной, т.е. $\varphi(-x) = \varphi(x)$.

Замечание 3. Формула Лапласа тем точнее приближает формулу Бернулли, чем больше число n (более нескольких десятков) и $n \cdot p > 10$.

Теорема 5 (интегральная формула Лапласа). Вероятность $P_n(m_1, m_2)$ того, что событие A наступит от m_1 до m_2 раз в серии из n одинаковых независимых испытаний приближённо вычисляется по формуле Лапласа: $P_n(m_1, m_2) \approx \Phi\left(\frac{m_2 - \mathrm{np}}{\sqrt{npq}}\right) - \Phi\left(\frac{m_1 - \mathrm{np}}{\sqrt{npq}}\right)$,

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^x e^{-\frac{t^2}{2}}\,\mathrm{d}t$.

Замечание 4. Имеются таблицы, в которых помещены значения функции $\Phi(x)$ при $0 \leq x \leq 5$ (в приложении 2 табл. П 2.2). При $x < 0$ пользуются теми же таблицами, так как функция $\Phi(x)$ является нечётной, т.е. $\Phi(-x) = -\Phi(x)$. Для $x > 5$ можно считать $\Phi(x) = 0{,}5$.

Формула Пуассона

Теорема 6 (формула Пуассона). Пусть p – вероятность наступления события A в каждом испытании. Тогда вероятность $P_n(m)$ наступления события A ровно m раз в серии из n одинаковых независимых испытаний приближённо вычисляется по формуле Пуассона:

$$P_n(m) \approx \frac{\lambda^m \cdot e^{-\lambda}}{m!},$$

где $\lambda = np$.

Замечание 5. Формула Пуассона тем точнее, чем меньше p и больше число n (более нескольких сотен), причём $n \cdot p < 10$.

Задачи с решениями

Задача 3.1. В первой коробке находится 20 деталей, из них 18 стандартных, во второй коробке – 10 деталей, из них 9 стандартных. Из второй коробки наудачу взята одна деталь и переложена в первую коробку. Какова вероятность того, что деталь, наудачу извлечённая после этого из первой коробки, окажется стандартной?

Решение. Обозначим события:

A – из первой коробки извлечена стандартная деталь.

H_1– из второй коробки в первую переложена стандартная деталь.

H_2 – из второй коробки в первую переложена нестандартная деталь.

Событие A может наступить при условии наступления одного из событий H_1, H_2. Эти события несовместны и образуют полную группу, т. е. являются гипотезами в формуле полной вероятности. Вероятность того, что из второй коробки извлечена стандартная деталь, $P(H_1) = \frac{9}{10}$.

Вероятность того, что из второй коробки извлечена нестандартная деталь $P(H_2) = \frac{1}{10}$.

Условная вероятность того, что из первой коробки извлечена стандартная деталь, при условии, что из второй коробки в первую была переложена стандартная деталь, $P_{H_1}(A) = \frac{19}{21}$.

Условная вероятность того, что из первой коробки извлечена стандартная деталь, при условии, что из второй коробки в первую была переложена нестандартная деталь, $P_{H_2}(A) = \frac{18}{21}$.

Искомая вероятность того, что из первой коробки будет извлечена стандартная деталь, по формуле полной вероятности равна:

$$P(A) = P(H_1) \cdot P_{H_1}(A) + P(H_2) \cdot P_{H_2}(A) = \frac{9}{10} \cdot \frac{19}{21} + \frac{1}{10} \cdot \frac{18}{21} = 0{,}9.$$

Ответ: 0,9.

Задача 3.2. Два станка производят одинаковые детали, которые поступают на общий конвейер. Производительность первого станка в два раза больше производительности второго станка. Первый производит 60 %

деталей высшего сорта, а второй – 84 %. Наудачу взятая с конвейера деталь оказалась высшего сорта. Какова вероятность того, что эта деталь произведена на первом станке?

Решение. Обозначим события:

A – деталь, взятая с конвейера, оказалась высшего сорта.

Это событие наступит с одним из двух событий (гипотез):

H_1 – эта деталь произведена на первом станке,

H_2 – эта деталь произведена на втором станке.

Поскольку производительность первого станка в два раза больше производительности второго станка, вероятности гипотез равны: $P(H_1)=\frac{2}{3}$, $P(H_2)=\frac{1}{3}$.

Условные вероятности события A даны: $P_{H_1}(A) = 0{,}6$, $P_{H_2}(A) = 0{,}84$.

По формуле полной вероятности находим:

$$P(A) = P(H_1) \cdot P_{H_1}(A) + P(H_2) \cdot P_{H_2}(A) = \frac{2}{3} \cdot 0{,}6 + \frac{1}{3} \cdot 0{,}84 = 0{,}68.$$

По формуле Байеса найдём условную вероятность того, что взятая наудачу деталь высшего сорта произведена на первом станке:

$$P_A(H_1)= \frac{P(H_1) \cdot P_{H_1}(A)}{P(A)} = \frac{\frac{2}{3} \cdot 0{,}6}{0{,}68} = \frac{10}{17}.$$

Ответ: $\frac{10}{17}$.

Задача 3.3. В ящике 20 белых и 10 чёрных шаров. Поочерёдно извлекают 4 шара, причём каждый извлечённый шар возвращают в ящик перед извлечением следующего. Какова вероятность того, что среди четырёх извлечённых шаров окажется два белых?

Решение. Вероятность извлечения белого шара одна и та же во всех четырёх испытаниях, так как каждый извлечённый шар возвращается в ящик: $p = \frac{20}{30} = \frac{2}{3}$.

Тогда вероятность извлечения чёрного шара во всех четырёх испытаниях равна $q = 1 - p = \frac{1}{3}$.

Используя формулу Бернулли, находим вероятность того, что из четырёх извлечённых шаров два шара будут белыми:

$$P_4(2) = C_4^2 \cdot p^2 \cdot q^{4-2} = \frac{4 \cdot 3}{1 \cdot 2} \cdot \left(\frac{2}{3}\right)^2 \cdot \left(\frac{1}{3}\right)^2 = 6 \cdot \frac{4}{9} \cdot \frac{1}{9} = \frac{8}{27}.$$

Ответ: $\frac{8}{27}$.

Задача 3.4. Вероятность поражения мишени при одном выстреле равна 0,8. Какова вероятность того, что при 100 выстрелах мишень будет поражена 75 раз?

Решение. По условию задачи $n = 100$, $m = 75$, $p = 0{,}8$, $q = 1 - p = 0{,}2$.

Так как n – достаточно большое число, воспользуемся локальной формулой Лапласа:

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \cdot \varphi\left(\frac{m-np}{\sqrt{npq}}\right),$$

$$P_{100}(75) \approx \frac{1}{\sqrt{100 \cdot 0{,}8 \cdot 0{,}2}} \cdot \varphi\left(\frac{75-100 \cdot 0{,}8}{\sqrt{100 \cdot 0.8 \cdot 0{,}2}}\right) = \frac{1}{\sqrt{16}} \cdot \varphi\left(\frac{75-80}{\sqrt{16}}\right) = \frac{1}{4} \cdot \varphi\left(\frac{-5}{4}\right) = \frac{1}{4} \cdot \varphi(1{,}25).$$

В таблице значений функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$ находим $\varphi(1{,}25) = 0{,}1826$.

Следовательно, $P_{100}(75) \approx \frac{1}{4} \cdot 0{,}1826 = 0{,}04565$.

Ответ: 0,04565.

Задача 3.5. Вероятность поражения мишени при одном выстреле равна 0,8. Какова вероятность того, что при 100 выстрелах мишень будет поражена

а) не менее 75 раз и не более 90 раз?

б) не менее 75 раз?

в) не более 74 раз?

Решение. Воспользуемся интегральной формулой Лапласа:

$$P_n(m_1, m_2) \approx \Phi\left(\frac{m_2-\mathrm{np}}{\sqrt{npq}}\right) - \Phi\left(\frac{m_1-\mathrm{np}}{\sqrt{npq}}\right),$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^x e^{-\frac{t^2}{2}} \mathrm{d}t$ – функция Лапласа.

а) По условию задачи $n = 100$, $p = 0{,}8$, $q = 0{,}2$, $m_1 = 75$, $m_2 = 90$. Тогда, воспользовавшись таблицей значений функции $\Phi(x)$, получаем:

$P_{100}(75;\ 90) \approx \Phi\left(\frac{90-100 \cdot 0{,}8}{\sqrt{100 \cdot 0{,}8 \cdot 0{,}2}}\right) - \Phi\left(\frac{75-100 \cdot 0{,}8}{\sqrt{100 \cdot 0{,}8 \cdot 0{,}2}}\right) = \Phi\left(\frac{10}{4}\right) - \Phi\left(\frac{-5}{4}\right) = \Phi(2{,}5) +$
$+\Phi(1{,}25) = 0{,}4938 + 0{,}3944 = 0{,}8882$.

б) Требование того, чтобы событие наступило не менее 75 раз, означает следующее: число появлений события может быть равно либо 75, либо 76, … , либо 100.

Тогда следует принять $m_1 = 75$, $m_2 = 100$. Воспользовавшись таблицей значений функции Лапласа *Ф(х),* получаем: $P_{100}(75;\ 100) \approx \Phi\left(\frac{100-80}{4}\right) -$
$-\Phi\left(\frac{75-80}{4}\right) = \Phi(5) - \Phi(-1{,}25) = \Phi(5) + \Phi(1{,}25) = 0{,}5 + 0{,}3944 = 0{,}8944$.

в) Событие "мишень поражена не более 74 раз" и событие "мишень поражена не менее 75 раз" являются противоположными. Поэтому сумма их вероятностей равна 1. Следовательно, искомая вероятность $P_{100}(0;\ 74) =$
$= 1 - P_{100}(75;\ 100) = 1 - 0{,}8944 = 0{,}1056$.

Ответ: а) 0,8882; б) 0,8944; в) 0,1056.

Задача 3.6. Учебник издан тиражом 100000 экземпляров. Вероятность того, что один учебник сброшюрован неправильно, равна 0,0001. Какова вероятность того, что тираж содержит 5 бракованных книг?

Решение. По условию задачи $n = 100000$, $p = 0{,}0001$.

События "из n книг ровно m книг сброшюрованы неправильно", где m = 0,1,2, … ,100000, являются независимыми. Так как число n велико, а вероятность p мала, вероятность $P_n(m)$ можно вычислить по формуле Пуассона: $P_n(m) \approx \frac{\lambda^m \cdot e^{-\lambda}}{m!}$, где $\lambda = np$.

В рассматриваемой задаче $\lambda = 100000 \cdot 0{,}0001 = 10$. Поэтому искомая вероятность $P_{100000}(5)$ определяется равенством:

$$P_{100000}\,(5) \approx \frac{e^{-10} \cdot 10^5}{5!} \approx \ 10^5 \cdot \frac{0{,}000045}{120} = 0{,}0375.$$

Ответ: 0,0375.

Задачи

3.1. В первой урне 2 белых шара и 4 черных, во второй – 3 белых и 5 черных. Из первой урны во вторую перекладывают один шар. Какова вероятность того, что шар, извлеченный после этого из второй урны, окажется белым?

3.2. В первом ящике 5 стандартных деталей и 3 нестандартных, во втором – 8 стандартных и 1 нестандартная. Из первого ящика во второй переложена одна деталь. Какова вероятность того, что деталь, извлеченная после этого из второго ящика, окажется нестандартной?

3.3. В первой урне находятся 1 белый и 9 черных шаров, во второй – 1 черный и 5 белых шаров. Из каждой урны удалили по одному шару, а оставшиеся шары ссыпали в третью урну. Какова вероятность того, что шар, вынутый из третьей урны, белый?

3.4. В первом пакете находятся 2 пирожка с рисом и 6 пирожков с капустой, во втором пакете – 4 пирожка с рисом и 3 пирожка с капустой. Из каждого пакета вынули по одному пирожку, а оставшиеся пирожки переложили в третий пакет, из которого после этого вынули пирожок. Какова вероятность того, что из третьего пакета вынут пирожок с капустой?

3.5. Для участия в студенческих отборочных спортивных соревнованиях выделено из первой группы 4 участника, из второй – 6, из третьей – 5. Вероятности того, что студент из первой, второй, третьей группы попадет в сборную университета, соответственно равны 0,9 , 0,7 и 0,8. Наудачу выбранный студент в итоге соревнований попал в сборную университета. К какой из групп вероятнее всего принадлежал этот студент?

3.6. Изделия проверяются на стандартность одним из двух контролеров. К первому контролеру попадает 40 % всех изделий, ко второму – 60 %. Вероятность того, что изделие будет признано стандартным первым контролером, равна 0,8, а вторым – 0,95. Изделие при проверке было признано стандартным. Какова вероятность того, что изделие проверял первый контролер?

3.7. Баскетболист бросает мяч три раза. Вероятность попадания при каждом броске равна 0,7. Какова вероятность того, что он попадет

а) один раз?

б) хотя бы один раз?

3.8. При передаче сообщения вероятность искажения знака равна 0,1. Какова вероятность того, что сообщение из 10 знаков

а) не будет искажено?

б) будет содержать 3 искажения?

3.9. Что вероятнее: выиграть у равносильного противника

а) три партии из четырех или пять из восьми?

б) не менее трех партий из четырех или не менее пяти партий из восьми?

3.10. Какова вероятность того, что четырехзначный номер первой встретившейся автомашины

а) не содержит цифры 5?

б) содержит две пятерки?

3.11. В ящике находятся 15 теннисных мячей, из которых 9 новых. Для первой игры наугад берутся три мяча, которые после игры возвращаются в ящик. Для второй игры также наугад берутся три мяча. Какова вероятность того, что все мячи, взятые для второй игры, новые? (Новым считается мяч, ни разу не побывавший в игре).

3.12. В специализированную больницу поступают в среднем 50 % больных с заболеванием "L", 30 % – с заболеванием "M", 20 % – с заболеванием "N". Вероятность полного излечения от болезни "L" равна 0,7, от болезни "M" – 0,8, от болезни "N" – 0,9. Больной, поступивший в больницу, был выписан с диагнозом – здоровый. Какова вероятность того, что этот больной страдал заболеванием "L"?

3.13. В цехе 3 типа автоматических станков производят одни и те же детали. Производительность у станков одинаковая, но качество работы различное. Известно, что станки первого типа производят 90 % деталей высшего сорта, второго типа – 80 %, третьего типа – 85 %. Какова вероятность того, что взятая наудачу деталь высшего сорта произведена станком первого типа, если станков первого типа 10 штук, второго – 8 штук и третьего – 2 штуки?

3.14. Телеграфное сообщение состоит из сигналов "точка" и "тире". Статистические свойства помех таковы, что искажается в среднем 2/5 сигналов "точка" и 1/3 сигналов "тире". Известно, что среди передаваемых сигналов "точка" и "тире" содержатся в отношении 5:3. Какова вероятность того, что принят неискаженный сигнал?

3.15. На распределительной базе находятся электрические лампочки, произведенные двумя заводами. Среди них 70 % – первого завода и 30 % – второго завода. Известно, что из каждых 100 лампочек первого завода 90 штук удовлетворяют требованиям стандарта, а из 100 лампочек второго завода удовлетворяют стандарту 80 штук. Какова вероятность того, что взятая наудачу с базы лампочка будет удовлетворять требованиям стандарта?

3.16. На двух станках изготавливаются одинаковые детали. Известно, что вероятность изготовления детали высшего качества на первом станке равна 0,92 , на втором – 0,8. Изготовленные на обоих станках не рассортированные детали находятся на складе. Среди них деталей, изготовленных на первом станке, в 3 раза больше, чем деталей, изготовленных на втором. Какова вероятность того, что взятая наудачу деталь высшего качества?

3.17. Для контроля продукции из трех партий деталей взята для испытаний одна деталь. Как велика вероятность обнаружения бракованной продукции, если в одной партии $\frac{2}{3}$ деталей бракованные, а в двух других – все доброкачественные?

3.18. Установлено, что в среднем 0,5 % шариков, изготовленных для подшипников, оказываются бракованными. Какова вероятность того, что среди поступивших на контроль 10 000 шариков бракованными окажутся 50 штук?

3.19. При установившемся технологическом процессе происходит 10 обрывов нити на 100 веретенах в час. Какова вероятность того, что в течение часа на 80 веретенах произойдет 7 обрывов нити?

3.20. По данным технического контроля в среднем 2 % изготавливаемых на заводе часов нуждаются в дополнительной регулировке. Какова вероятность того, что из 300 изготовленных часов 290 штук не будут нуждаться в дополнительной регулировке?

3.21. Проверкой качества изготавливаемых на заводе часов установлено, что в среднем 98 % их отвечает предъявляемым требованиям, а 2 % нуждаются в дополнительной регулировке. Приемщик проверяет качество 300 штук изготовленных часов. Если при этом среди них обнаруживается 11 или более часов, нуждающихся в дополнительной регулировке, вся партия возвращается заводу для доработки. Какова вероятность того, что партия будет принята?

3.22. В результате проверки качества приготовленного для посева зерна было установлено, что 90 % зерен всхожие. Какова вероятность того, что среди отобранных и высаженных 1000 зерен прорастет

а) не менее 700 штук?

б) от 700 до 740 штук?

в) от 880 до 920 штук?

3.23. Вероятность выпуска нестандартной электролампы равна 0,1. Чему равна вероятность того, что в партии из 2000 ламп а) число стандартных будет не менее 1790 штук?

б) число нестандартных будет менее 201 штуки?

3.24. Сборщик получил 3 коробки деталей, изготовленных заводом №1, и 2 коробки деталей, изготовленных заводом №2.Вероятность того, что деталь завода №1 стандартна, равна 0,8, а завода №2 – 0,9. Сборщик наудачу извлек деталь из наудачу взятой коробки. Какова вероятность того, что извлечена стандартная деталь?

3.25. В первом ящике содержится 20 деталей, из них 15 стандартных, во второй – 30 деталей, из них 24 стандартных, в третьей – 10 деталей, из них 6 стандартных. Какова вероятность того, что наудачу извлеченная деталь из наудачу выбранного ящика – стандартная?

3.26. Из полного набора 28 костей домино наудачу извлечена кость. Какова вероятность того, что вторую извлеченную наудачу кость можно приставить к первой?

Ответы

3.1. $\frac{10}{27}$; **3.2.** $\frac{11}{80}$; **3.3.** $\frac{38}{105}$; **3.4.** $\frac{219}{364}$; **3.5.** $\frac{18}{59}$, $\frac{21}{59}$, $\frac{20}{59} \Rightarrow$ второй; **3.6.** $\frac{32}{89}$; **3.7.** а) 0,189, б) 0,973; **3.8.** а) $0{,}9^{10} \approx 0{,}35$, б) $0{,}12 \cdot 0{,}9^7 \approx 0{,}057$; **3.9.** а) $\frac{1}{4}$, $\frac{7}{32}$ ($p_1 > p_2$); б) $\frac{5}{16}$, $\frac{93}{256}$ ($p_2 > p_1$); **3.10.** а) 0,6561, б) 0,0486; **3.11.** $\frac{528}{5915} \approx 0{,}089$; **3.12.** $\frac{5}{11}$; **3.13.** $\frac{15}{29}$; **3.14.** $\frac{5}{8}$; **3.15.** 0,87; **3.16.** 0,89; **3.17.** $\frac{2}{9}$; **3.18.** 0,057; **3.19.** 0,139; **3.20.** 0,042; **3.21.** 0,9438; **3.22.** а) 1; б) 0; в) 0,9651; **3.23.** а) 0,7732; б) 0,5279; **3.24.** 0,84; **3.25.** $\frac{46}{63}$; **3.26.** $\frac{7}{18}$.

Глава 2. Случайные величины

§4. Дискретная случайная величина

Понятие случайной величины. Закон распределения дискретной случайной величины

Определение 1. *Случайной* называется величина, которая в результате испытания принимает только одно значение из возможного множества своих значений, наперед неизвестное и зависящее от случайных причин.

Различают два вида случайных величин: дискретные и непрерывные.

Определение 2. Случайная величина X называется *дискретной* (прерывной), если множество ее значений конечное или бесконечное, но счетное.

Другими словами, возможные значения дискретной случайной величину можно перенумеровать.

Описать случайную величину можно с помощью ее закона распределения.

Определение 3. *Законом распределения дискретной случайной величины* называют соответствие между возможными значениями случайной величины и их вероятностями.

Закон распределения дискретной случайной величины X может быть задан в виде таблицы, в первой строке которой указаны в порядке возрастания все возможные значения случайной величины, а во второй строке соответствующие вероятности этих значений, т.е.:

x	x_1	x_2	x_3	…	x_n
p	p_1	p_2	p_3	…	p_n

где $p_1 + p_2 + \ldots + p_n = 1$.

Такая таблица называется *рядом распределения дискретной случайной величины*.

Если множество возможных значений случайной величины бесконечно, то ряд $p_1 + p_2 + \ldots + p_n + \ldots$ сходится и его сумма равна 1.

Закон распределения дискретной случайной величины X можно изобразить графически, для чего в прямоугольной системе координат строят ломаную, соединяющую последовательно точки с координатами $(x_i;p_i)$, $i=1,2,\ldots n$. Полученную линию называют *многоугольником распределения* (рис. 4.1).

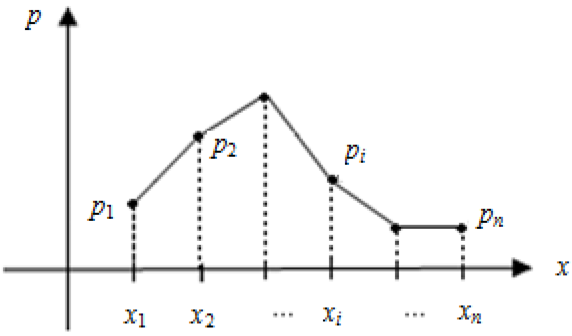


Рис. 4.1

Закон распределения дискретной случайной величины X может быть также задан аналитически (в виде формулы):

$$P(X=x_i)=\varphi(x_i), i=1,2,3\ldots n.$$

З а д а ч а 4.1. Вероятности того, что студент сдаст экзамены в сессию по математическому анализу и органической химии соответственно равны 0,7 и 0,8. Составить закон распределения случайной величины X – числа экзаменов, которые сдаст студент.

Решение. Рассматриваемая случайная величина X в результате экзамена может принять одно из следующих значений: $x_1=0$, $x_2=1$, $x_3=2$.

Найдем вероятности этих значений. Обозначим события:

A – студент сдаст экзамен по математическому анализу;

$\bar{A}$ – студент не сдаст экзамен по математическому анализу;

B – студент сдаст экзамен по органической химии;

$\bar{B}$ – студент не сдаст экзамен по органической химии.

По условию:

$P(A) = 0,7 \Rightarrow P(\bar{A}) = 1 - P(A) = 0,3;$

$P(B) = 0,8 \Rightarrow P(\bar{B}) = 1 - P(B).$

Тогда:

$P(x=0) = P(\bar{A} \cdot \bar{B}) = P(\bar{A}) \cdot P(\bar{B}) = 0,3 \cdot 0,2 = 0,06.$

$P(x=1) = P(A\bar{B} + \bar{A}B) = P(A) \cdot P(\bar{B}) + P(\bar{A}) \cdot P(B) = 0,7 \cdot 0,2 + 0,3 \cdot 0,8 = 0,38$

$P(x=1) = P(AB) = P(A) \cdot P(B) = 0,7 \cdot 0,8 = 0,56.$

Итак, закон распределения случайной величины X задается таблицей:

x	0	1	2
p	0,06	0,38	0,56

Контроль: 0,06+0,38+0,56=1.

Фунция распределения

Полное описание случайной величины дает также функция распределения.

Определение 4. *Функцией распределения дискретной случайной величины* X называется функция $F(x)$, определяющая для каждого значения x вероятность того, что случайная величина X примет значение, меньше x:

$$F(x) = \mathrm{P}(X < x)$$

Геометрически функция распределения интерпретируется как вероятность того, что случайная величина X примет значение, которое изображается на числовой прямой точкой, лежащей левее точки x.

Свойства функции распределения

1) $0 \leq F(x) \leq 1$;

2) $F(x)$ – неубывающая функция на промежутке $(-\infty;+\infty)$;

3) $F(x)$ – непрерывна слева в точках $x = x_i$ $(i=1,2,\ldots n)$ и непрерывна во всех остальных точках;

4) $F(-\infty)=P(X<-\infty)=0$ как вероятность невозможного события $X<-\infty$, $F(+\infty)=P(X<+\infty)=1$ как вероятность достоверного события $X<+\infty$.

Если закон распределения дискретной случайной величины X задан в виде таблицы:

x	x_1	x_2	x_3	…	x_n
p	p_1	p_2	p_3	…	p_n

то функция распределения $F(x)$ определяется формулой:

$$F(x) = \begin{cases} 0 \text{ при } x \le x_1, \\ p_1 \text{ при } x_1 < x \le x_{2,} \\ p_1 + p_2 \text{ при } x_2 < x \le x_3 \\ \dots \quad \dots \quad \dots \\ 1 \text{ при } x > x_n \end{cases}$$

Её график изображен на рис. 4.2:

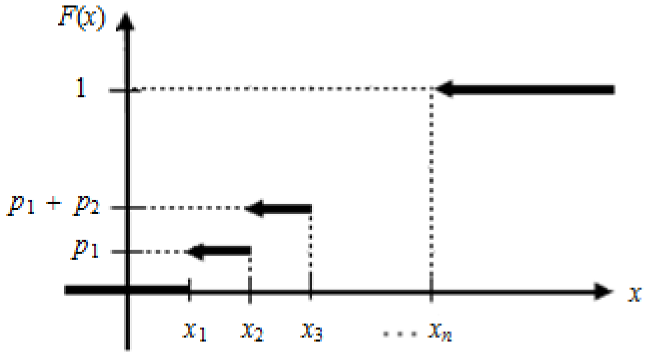


Рис. 4.2

Числовые характеристики дискретной случайной величины

Определение 5. *Математическим ожиданием* $M(X)$ дискретной случайной величины X называется сумма произведений всех ее значений на соответствующие им вероятности:

$$M(X) = \sum_{i=1}^{n} x_i p_i = x_1 p_1 + x_2 p_2 + \dots + x_n p_n$$

Математическое ожидание характеризует среднее значение случайной величины.

Свойства математического ожидания

1) $M(C)=C$, где С – постоянная величина;

2) $M(C \cdot X)=C \cdot M(X)$,

3) $M(X \pm Y)=M(X) \pm M(Y)$;

4) $M(X \cdot Y)=M(X) \cdot M(Y)$, где X,Y – независимые случайные величины;

5) $M(X \pm C)=M(X) \pm C$, где C – постоянная величина.

Для характеристики степени рассеяния возможных значений дискретной случайной величины вокруг ее математического ожидания служит дисперсия.

Определение 6. *Дисперсией* $D(X)$ случайной величины X называется математическое ожидание квадрата отклонения случайной величины от её математического ожидания:

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

Свойства дисперсии

1) $D(C)=0$, где C – постоянная величина;

2) $D(X)>0$, где X – случайная величина;

3) $D(C\cdot X)=C^2\cdot D(X)$, где C – постоянная величина;

4) $D(X+Y)=D(X)+D(Y)$, где X,Y – независимые случайные величины;

Для вычисления дисперсии часто бывает удобно пользоваться формулой:

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

$$M(X^2)=\sum_{i=1}^{n}{x_i}^2p_i=x_1^2p_1+x_2^2p_2+\ldots+x_n^2p_n$$

Дисперсия $D(X)$ имеет размерность квадрата случайной величины, что не всегда удобно. Поэтому в качестве показателя рассеяния возможных значений случайной величины используют также величину $\sqrt{D(X)}$.

Определение 7. *Средним квадратическим отклонением* $\sigma(X)$ случайной величины X называется квадратный корень из дисперсии:

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

Задача 4.2. Дискретная случайная величина X задана законом распределения:

x	–1	0	1	2	3
p	0,1	p_2	0,3	0,2	0,3

Найти p_2, функцию распределения $F(x)$ и построить её график, а также $M(X)$,$D(X)$, $\sigma(X)$.

Решение: Так как сумма вероятностей возможных значений случайной величины X равна 1, то

$P_2=1-(0{,}1+0{,}3+0{,}2+0{,}3)=0{,}1$

Найдем функцию распределения $F(x)=P(X<x)$.

Геометрически это равенство можно истолковать так: $F(x)$ есть вероятность того, что случайная величина примет значение, которое изображается на числовой оси точкой, лежащей левее точки x.

Если $x\le -1$, то $F(x)=0$, так как на промежутке $(-\infty;\,x)$ нет ни одного значения данной случайной величины;

Если $-1<x\le 0$, то $F(x)=P(X=-1)=0{,}1$, так как в промежуток $(-\infty;\,x)$ попадает только одно значение $x_1=-1$;

Если $0<x\le 1$, то $F(x)=P(X=-1)+P(X=0)=0{,}1+0{,}1=0{,}2$, так как в промежуток $(-\infty;\,x)$ попадают два значения $x_1=-1$ и $x_2=0$;

Если $1<x\le 2$, то $F(x)=P(X=-1)+P(X=0)+P(X=1)=0{,}1+0{,}1+0{,}3=0{,}5$, так как в промежуток $(-\infty;\,x)$ попадают три значения

$x_1=-1$, $x_2=0$ и $x_3=1$;

Если $2 < x \le 3$, то $F(x) = P(X = -1) + P(X = 0) + P(X = 1) + P(X = 2) = 0{,}1+0{,}1+0{,}3+0{,}2=0{,}7$, так как в промежуток $(-\infty;\ x)$ попадают четыре значения $x_1 = -1$, $x_2 = 0$, $x_3 = 1$ и $x_4 = 2$;

Если $x > 3$, то $F(x)=P(X = -1)+P(X = 0)+P(X = 1)+P(X = 2)+P(X = 3) = 0{,}1+0{,}1+0{,}3+0{,}2+0{,}3=1$, так как в промежуток $(-\infty;\ x)$ попадают пять значений $x_1= -1$, $x_2 = 0$, $x_3 = 1$, $x_4 = 2$ и $x_5 = 3$.

Итак,

$$F(x) = \begin{cases} 0 \text{ при } x \le -1, \\ 0{,}1 \text{ при } -1 < x \le 0, \\ 0{,}2 \text{ при } 0 < x \le 1, \\ 0{,}5 \text{ при } 1 < x \le 2, \\ 0{,}7 \text{ при } 2 < x \le 3, \\ 1 \text{ при } x > 3. \end{cases}$$

Изобразим функцию $F(x)$ графически (рис. 4.3):

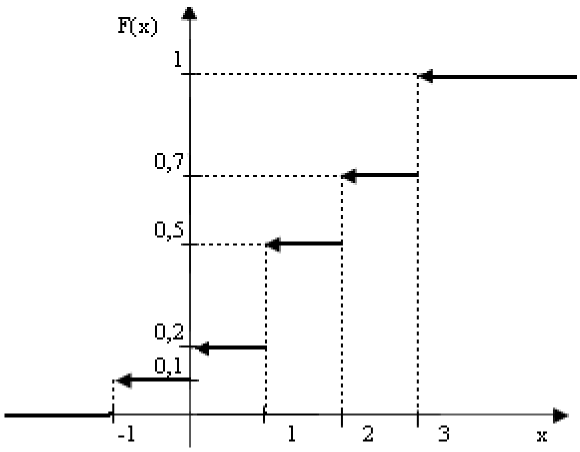


Рис. 4.3

Найдем числовые характеристики случайной величины:

$$M(X) = \sum_{k=1}^{n} x_k p_k = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n$$

$$M(X) = -1 \cdot 0{,}1 + 0 \cdot 0{,}1 + 1 \cdot 0{,}3 + 2 \cdot 0{,}2 + 3 \cdot 0{,}3 = 1{,}5$$

$$D(X) = \sum_{k=1}^{n} x_k^2 p_k - \left(M(X^2)\right) = x_1^2 p_1 + x_2^2 p_2 \ldots + x_n^2 p_n - \left(M(X^2)\right)$$

$$D(X) = (-1)^2 \cdot 0{,}1 + 0^2 \cdot 0{,}1 + 1^2 \cdot 0{,}3 + 2^2 \cdot 0{,}2 + 3^2 \cdot 0{,}3 - (1{,}5)^2 = \\ = 1{,}65$$

$$\sigma(X) = \sqrt{D(X)} = \sqrt{1{,}65} \approx 1{,}2845.$$

Некоторые законы распределения дискретной случайной величины

Определение 8. *Биномиальным* называется закон распределения дискретной случайной величины X – числа появлений события A в n независимых повторных испытаниях, в каждом из которых события A может наступить с вероятностью p или не наступить с вероятностью $q=1-p$. Тогда $P(X=m)$ – вероятность появления события A ровно m раз в n испытаниях вычисляется по формуле Бернулли:

$$P(X=m)=C^m_n p^m q^{n-m}.$$

Математическое ожидание, дисперсию и среднее квадратическое отклонение случайной величины X, распределенной по *биномиальному закону*, находят, соответственно, по формулам:

$$M(X)=np,$$
$$D(X)=npq,$$
$$\sigma(X)=\sqrt{npq}.$$

Если число испытаний n очень велико, а вероятность появления события A в каждом испытании очень мала ($p \le 0{,}1$), то для вычисления $P(X=m)$ используют формулу Пуассона:

$$P(\mathrm{X}=m)=P_n(m)=\frac{e^{-\lambda}\cdot\lambda^m}{m!},$$

где $\lambda=np$.

Тогда говорят, что случайная величина X распределена *по закону Пуассона.*

Так как вероятность p события A в каждом испытании мала, то *закон распределения Пуассона* называется законом редких явлений.

Задача 4.3. Составить закон распределения случайной величины X – числа выпадений пятерки при трех бросаниях игральной кости. Вычислить $M(X)$, $D(X)$, $\sigma(X)$ этой величины.

Решение: Испытание состоит в одном бросании игральной кости. Так как кость бросается 3 раза, то число испытаний $n=3$.

Вероятность события A – "выпадение пятёрки" в каждом испытании одна и та же и равна 1/6, т.е. $P(A)=p=1/6$, тогда $P(\bar{A})=1-p=q=5/6$,
где $\bar{A}$ – "выпадения не пятёрки".

Случайная величина X может принимать значения: 0;1;2;3.

Вероятность каждого из возможных значений X найдём по формуле Бернулли:

$P(X=0)=P_3(0)=C^0_3p^0q^3=1\cdot(1/6)^0\cdot(5/6)^3=125/216;$
$P(X=1)=P_3(1)=C^1_3p^1q^2=3\cdot(1/6)^1\cdot(5/6)^2=75/216;$
$P(X=2)=P_3(2)=C^2_3p^2q=3\cdot(1/6)^2\cdot(5/6)^1=15/216;$
$P(X=3)=P_3(3)=C^3_3p^3q^0=1\cdot(1/6)^3\cdot(5/6)^0=1/216.$

Таким образом закон распределения случайной величины X имеет вид:

x	0	1	2	3
p	125/216	75/216	15/216	1/216

Контроль: 125/216+75/216+15/216+1/216=1.

Найдем числовые характеристики случайной величины X:

$M(X) = np = 3 \cdot (1/6) = 1/2$,

$D(X) = npq = 3 \cdot (1/6) \cdot (5/6) = 5/12$,

$\sigma(X) = \sqrt{npq} = \sqrt{5/12} = \sqrt{15}/6$.

Задача 4.4. Станок-автомат штампует детали. Вероятность того, что изготовленная деталь окажется бракованной равна 0,002. Найти вероятность того, что среди 1000 отобранных деталей окажется:

а) 5 бракованных;

б) хотя бы одна бракованная.

Решение: Число n = 1000 велико, вероятность изготовления бракованной детали p = 0,002 мала, и рассматриваемые события (деталь окажется бракованной) независимы, поэтому имеет место формула Пуассона:

$$P_n(m) = \frac{e^{-\lambda} \cdot \lambda^m}{m!}$$

Найдем $\lambda = np = 1000 \cdot 0{,}002 = 2$.

а) Найдем вероятность того, что будет 5 бракованных деталей среди отобранных (m = 5):

$\mathrm{P}_{1000}(5) = \frac{e^{-2} \cdot 2^5}{5!} = \frac{32 \cdot 0{,}13534}{120} = 0{,}0361$.

б) Найдем вероятность того, что будет хотя бы одна бракованная деталь среди отобранных.

Событие A – "хотя бы одна из отобранных деталей бракованная" является противоположным событию $\bar{A}$ – "все отобранные детали не бракованные". Следовательно, $P(A) = 1 - P(\bar{A})$. Отсюда искомая вероятность равна: $P(A) = 1 - P_{1000}(0) = 1 - \frac{e^{-2} \cdot 2^0}{0!} = 1 - e^{-2} = 1 - 0{,}13534 \approx 0{,}865$

Задачи

4.1. Дискретная случайная величина X задана законом распределения:

x	–2	0	2	5
p	0,3	0,2	p_3	0,1

Найти p_3, функцию распределения $F(X)$ и построить ее график, а также $M(X)$, $D(X)$, $\sigma(X)$.

4.2. Дискретная случайная величина X задана законом распределения:

x	–1	0	1	2	3
p	0,3	0,1	0,2	p_4	0,3

Найти p_4, функцию распределения $F(X)$ и построить ее график, а также $M(X)$,$D(X)$, $\sigma(X)$.

4.3. В коробке 9 фломастеров, из которых 2 фломастера уже не пишут. Наудачу берут 3 фломастера. Случайная величина X – число пишущих фломастеров среди взятых. Составить закон распределения случайной величины.

4.4. На стеллаже библиотеки в случайном порядке расставлено 6 учебников, причем 4 из них в переплете. Библиотекарь берет наудачу 4 учебника. Случайная величина X – число учебников в переплете среди взятых. Составить закон распределения случайной величины.

4.5. В билете две задачи. Вероятность правильного решения первой задачи равна 0,9, второй – 0,7. Случайная величина X – число правильно решенных задач в билете. Составить закон распределения, вычислить математическое ожидание и дисперсию этой случайной величины, а также найти функцию распределения $F(x)$ и построить ее график.

4.6. Три стрелка стреляют по мишени. Вероятность попадания в мишень при одном выстреле для первого стрелка равна 0,5, для второго – 0,8, для третьего – 0,7. Случайная величина X – число попаданий в мишень, если стрелки делают по одному выстрелу. Найти закон распределения, $M(X)$,$D(X)$.

4.7. Баскетболист бросает мяч в корзину с вероятностью попадания при каждом броске 0,8. За каждое попадание он получает 10 очков, а в случае промаха очки ему не начисляют. Составить закон распределения случайной величины X – числа очков, полученных баскетболистом за 3 броска. Найти $M(X)$,$D(X)$, а также вероятность того, что он получит более 10 очков.

4.8. На карточках написаны буквы, всего 5 гласных и 3 согласных. Наугад выбирают 3 карточки, причем каждый раз взятую карточку возвращают назад. Случайная величина X – число гласных букв среди взятых. Составить закон распределения случайной величины X и найти $M(X)$,$D(X)$,$\sigma(X)$.

4.9. В среднем по 60 % договоров страховая компания выплачивает страховые суммы в связи с наступлением страхового случая. Составить закон распределения случайной величины X – числа договоров, по которым была выплачена страховая сумма среди наудачу отобранных четырех договоров. Найти числовые характеристики этой величины.

4.10. Радиостанция через определенные промежутки времени посылает позывные сигналы (не более четырех) до установления двусторонней связи. Вероятность получения ответа на позывной сигнал равна 0,3. Случайная величина X – число посланных позывных сигналов. Составить закон распределения и найти $F(x)$.

4.11. Имеется 3 ключа, из которых только один подходит к замку. Составить закон распределения случайной величины X – числа попыток открывания замка, если испробованный ключ в последующих попытках не участвует. Найти $M(X)$,$D(X)$.

4.12. Производятся последовательные независимые испытания трех приборов на надежность. Каждый следующий прибор испытывается только в том случае, если предыдущий оказался надежным. Вероятность выдержать испытание для каждого прибора равна 0,9. Составить закон распределения случайной величины X – числа испытанных приборов.

4.13. Дискретная случайная величина X имеет три возможные значения: $x_1=1$, x_2, x_3, причем $x_1 < x_2 < x_3$. Вероятность того, что X примет значения x_1 и x_2, соответственно равны 0,3 и 0,2. Известно, что $M(X)=2,2$, $D(X)=0,76$. Составить закон распределения случайной величины.

4.14. Блок электронного устройства содержит 100 одинаковых элементов. Вероятность отказа каждого элемента в течении времени T равна 0,002. Элементы работают независимо. Найти вероятность того, что за время T откажет не более двух элементов.

4.15. Учебник издан тиражом 50000 экземпляров. Вероятность того, что учебник сброшюрован неправильно, равна 0,0002. Найти вероятность того, что тираж содержит:

а) четыре бракованные книги,

б) менее двух бракованных книг.

4.16. Число вызовов, поступающих на АТС каждую минуту, распределено по закону Пуассона с параметром $\lambda=1,5$. Найдите вероятность того, что за минуту поступит:

а) два вызова;

б)хотя бы один вызов.

4.17. Даны законы распределения двух независимых случайных величин:

X:

x	–2	0	2
p	0,5	0,2	0,3

Y:

y	0	1	3
p	0,2	0,5	0,3

Найти $M(Z)$,$D(Z)$, если $Z=3X+Y$.

4.18. Даны законы распределения двух независимых случайных величин:

X:

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

Y:

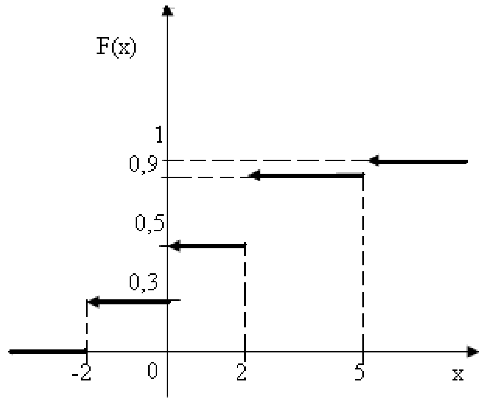
y	3	4	5
p	0,2	0,4	0,4

Найти $M(Z)$,$D(Z)$, если $Z=X+2Y$.

Ответы

4.1. p_3=0,4;

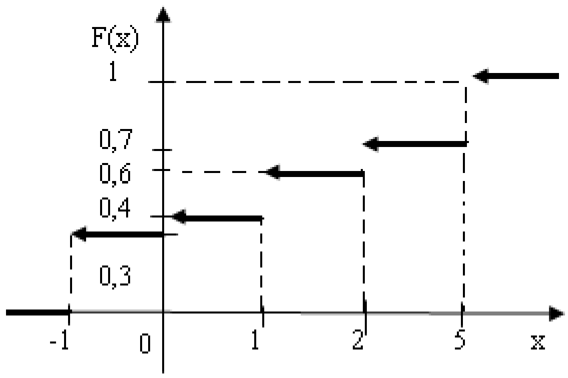
$$F(x)=\begin{cases} 0 & \text{при } x\le -2,\\ 0{,}3 & \text{при } -2<x\le 0,\\ 0{,}5 & \text{при } 0<x\le 2,\\ 0{,}9 & \text{при } 2<x\le 5,\\ 1 & \text{при } x>5 \end{cases}$$



$M(X)=0{,}7$; $D(X)=4{,}87$; $\sigma(X)\approx 2{,}193$.

4.2. p_4=0,1;

$$F(x)=\begin{cases} 0 & \text{при } x\le -1,\\ 0{,}3 & \text{при } -1<x\le 0,\\ 0{,}4 & \text{при } 0<x\le 1,\\ 0{,}6 & \text{при } 1<x\le 2,\\ 0{,}7 & \text{при } 2<x\le 3,\\ 1 & \text{при } x>3 \end{cases}$$



$M(X)=1$; $D(X)=2{,}6$; $\sigma(X)\approx 1{,}612$.

4.3.

x	1	2	3
p	7/84	1/2	35/84

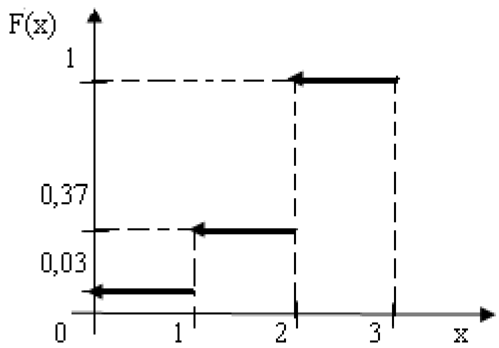
4.4.

x	2	3	4
p	2/5	8/15	1/15

4.5.

x	0	1	2
p	0,03	0,34	0,63

$$F(x)=\begin{cases} 0 & \text{при } x\le 0,\\ 0{,}03 & \text{при } 0<x\le 1,\\ 0{,}37 & \text{при } 1<x\le 2,\\ 1 & \text{при } x>2 \end{cases}$$



4.6.

x	0	1	2	3
p	0,03	0,22	0,47	0,28

$M(X)$=2; $D(X)$=0,62

4.7.

x	0	10	20	30
p	0,008	0,096	0,384	0,512

$M(X)$=2,4; $D(X)$=0,48, $P(X >10)$=0,896

4.8.

x	0	1	2	3
p	27/512	135/512	225/512	125/512

$M(X)$=15/8; $D(X)$=45/64; $\sigma(X) \approx 3\sqrt{5}/8$

4.9.

x	0	1	2	3	4
p	0,0256	0,1536	0,3456	0,3456	0,1296

$M(X)$=2,4; $D(X)$=0,96

4.10.

x	1	2	3	4
p	0,3	0,21	0,147	0,343

$$F(x) = \begin{cases} 0, & \text{при } x \le 1, \\ 0{,}3, & \text{при } 1<x\le 2, \\ 0{,}51, & \text{при } 2<x\le 3, \\ 0{,}657, & \text{при } 3<x\le 4, \\ 1, & \text{при } x>4 \end{cases}$$

4.11.

x	1	2	3
p	1/3	1/3	1/3

$M(X)$=2; $D(X)$=2/3

4.12.

x	1	2	3
p	0,9	0,09	0,01

4.13.

x	1	2	3
p	0,3	0,2	0,5

4.14. $1{,}22 \cdot e^{-0{,}2} \approx 0{,}999$
4.15. а)0,0189; б) 0,00049
4.16. а) 0,0702; б) 0,77687
4.17. 0,2; 28,6
4.18. 11,2; 4.

§5. Непрерывная случайная величина

Понятие непрерывной случайной величины. Функция распределения непрерывной случайной величины

Определение 1. *Непрерывной* называют величину, все возможные значения которой полностью заполняют конечный или бесконечный промежуток числовой оси.

Очевидно, число возможных значений непрерывной случайной величины бесконечно.

Непрерывную случайную величину можно задавать с помощью функции распределения.

Определение 2. Ф*ункцией распределения* непрерывной случайной величины X называется функция $F(x)$, определяющая для каждого значения $x \in R$ вероятность того, что случайная величина X в результате испытания примет значение, меньшее x:

$$F(x)=P(X<x),\ \text{где}\ x \in R.$$

Функцию распределения иногда называют интегральной функцией распределения.

Свойства функции распределения

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

2) У непрерывной случайной величины функция распределения непрерывна в любой точке и дифференцируема всюду, кроме, быть может, отдельных точек.

3) Вероятность попадания случайной величины X в один из промежутков $(a;b)$, $[a;b)$, $(a;b]$, $[a;b]$, равна разности значений функции $F(x)$ в точках a и b, т.е. $P(a<X<b)= F(b)-F(a)$.

4) Вероятность того, что непрерывная случайная величина X примет одно определенное значение равна 0.

5) $F(-\infty)=0$, $F(+\infty)=1$.

Задание непрерывной случайной величины с помощью функции распределения не является единственным.

Функция плотности распределения вероятностей

Определение 3. *Функцией плотности распределения вероятностей* $f(x)$ (или плотностью распределения) непрерывной случайной величины X называется производная от ее функции распределения, т.е.:

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

Плотность распределения вероятностей иногда называют дифференциальной функцией распределения или дифференциальным законом распределения.

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

Свойства функции плотности распределения вероятностей

1) $f(x)\geq 0$, при $x \in R$.

2) $F(x)=\int\limits_{-\infty}^{x} f(x)\mathrm{d}x$.

Геометрически функция распределения равна площади фигуры, ограниченной сверху кривой распределения, снизу осью Ox и лежащей левее точки x (рис. 5.1).

3) $P(a<X<b)=\int\limits_{a}^{b} f(x)\mathrm{d}x$.

Геометрически полученная вероятность равна площади криволинейной трапеции, ограниченной сверху кривой распределения, снизу осью Ox, слева и справа прямыми $x=a$, $x=b$ (рис. 5.2).

4) $\int\limits_{+\infty}^{-\infty} f(x)\mathrm{d}x=1$ – условие нормировки.

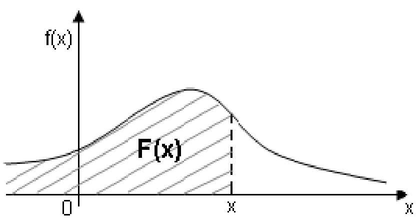


Рис. 5.1

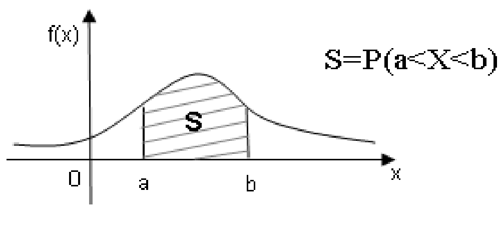


Рис. 5.2

Задача 5.1. Случайная величина X задана плотностью распределения вероятностей:

$$f(x) = \begin{cases} 0 \text{ при } x \le 2, \\ c(x-2) \text{ при } 2 < x \le 6, \\ 0 \text{ при } x > 6. \end{cases}$$

Найти: а) значение c; б) функцию распределения $F(x)$ и построить ее график; в) $P(3\le x <5)$

Решение:

а) Значение c найдем из условия нормировки: $\int\limits_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1$

Следовательно,

$$\int\limits_{-\infty}^{+\infty} f(x)\mathrm{d}x = \int\limits_{-\infty}^{2} 0\mathrm{d}x + \int\limits_{2}^{6} c(x-2)\mathrm{d}x + \int\limits_{6}^{+\infty} 0\mathrm{d}x = c\int\limits_{2}^{6}(x-2)\mathrm{d}x = c\,(x^2/2-2x)\Big|_2^6 =$$

$= c\,(36/2–12 – (4/2 – 4) = 8c$;

c =1/8.

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

Поэтому,

если $x \le 2$, то $F(x)=\int_{-\infty}^{x} 0dx = 0$;

если $2 < x \le 6$, то $F(x)=\int_{-\infty}^{2} 0dx + \int_{2}^{x}\frac{1}{8}\cdot(x-2)dx =$

$= 1/8(x^2/2 - 2x)\Big|_2^x = 1/8(x^2/2–2x– (4/2–4))=1/8(x^2/2–2x+2)=1/16(x–2)^2$;

если $x > 6$, то $F(x)=\int_{-\infty}^{2} 0dx + \int_{2}^{6} 1/8(x-2)dx + \int_{6}^{x} 0dx=$

$$= 1/8\int_{2}^{6}(x-2)dx = 1/8(x^2/\,2 - 2x)\Big|_2^6 =$$

$$= 1/8(36/2 - 12 - (4/2 + 4)) = 1/8\cdot 8 = 1.$$

Таким образом,

$$F(x) = \begin{cases} 0, \text{при } x \le 2, \\ (x-2)^2/16, \text{при } 2 < x \le 6, \\ 1, \text{при } x > 6. \end{cases}$$

График функции $F(x)$ изображен на рис. 5. 3.

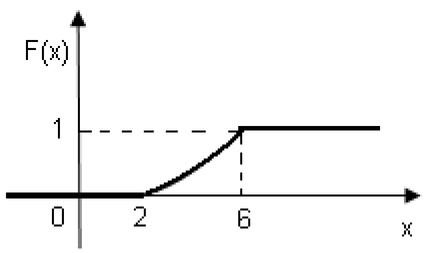


Рис. 5.3

в) $P(3\le X <5) = F(5) – F(3) = (5–2)^2/16 – (3–2)^2/16 = 9/16–1/16 =$
$=8/16=1/2$.

Задача 5.2. Случайная величина X задана функцией распределения:

$$F(x) = \begin{cases} 0 \text{ при } x \leq 0, \\ \dfrac{3 \cdot arctg\, x}{\pi} \text{ при } 0 < x \leq \sqrt{3}, \\ 1 \text{ при } x > \sqrt{3}. \end{cases}$$

Найти дифференциальную функцию распределения $f(x)$.

Решение: Так как $f(x) = F'(x)$, то

$$f(x) = \begin{cases} 0, \text{при } x \leq 0 \\ \dfrac{3}{\pi \cdot (1 + x^2)} \text{ при } 0 < x \leq \sqrt{3}, \\ 0 \text{ при } x > \sqrt{3}. \end{cases}$$

Числовые характеристики непрерывной случайной величины

Понятия математического ожидания $M(X)$ и дисперсии $D(X)$, введенные ранее для дискретной случайной величины, можно распространить на непрерывные случайные величины.

- ***Математическое ожидание M(X)*** непрерывной случайной величины X определяется равенством:

$$M(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx,$$

при условии, что этот интеграл сходится.

- ***Дисперсия D(X)*** непрерывной случайной величины X определяется равенством:

$$D(X) = \int_{-\infty}^{+\infty} (x - M(x))^2 \cdot f(x) dx \text{ или}$$

$$D(X) = \int_{-\infty}^{+\infty} x^2 \cdot f(x) dx - \left(M(x)\right)^2.$$

- ***Среднее квадратическое отклонение σ(X)*** непрерывной случайной величины определяется равенством:

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

Все свойства математического ожидания и дисперсии, рассмотренные ранее для дискретных случайных величин, справедливы и для непрерывных.

Задача 5.3. Случайная величина X задана дифференциальной функцией $f(x)$:

$$f(x) = \begin{cases} 0 \text{ при } x \le 0, \\ x/3 \text{ при } 0 < x \le 2, \\ 1/3 \text{ при } 2 < x \le 3 \\ 0 \text{ при } x > 3. \end{cases}$$

Найти *M*(*X*), *D*(*X*), σ(*X*), а также *P*(1 < *x* < 5).

Решение:

$$M(X) = \int_{-\infty}^{+\infty} x \cdot f(x)dx = \int_{-\infty}^{0} x \cdot 0dx + \int_{0}^{2} x \cdot x/3dx + \int_{2}^{3} x/3dx +$$

$$+ \int_{3}^{+\infty} 0 \cdot x \cdot dx = 1/3 \int_{0}^{2} x^2 dx + 1/3 \int_{2}^{3} xdx =$$

$$= x^3/9 \Big|_0^2 + x^2/6 \Big|_2^3 = 8/9 - 0 + 9/6 - 4/6 = 31/18,$$

$$D(X) = \int_{-\infty}^{x} x^2 \cdot f(x)dx - \left(M(x)\right)^2 = \int_{0}^{2} x^2 \cdot \frac{x}{3} \cdot dx + \int_{2}^{3} \frac{x^2}{3} dx -$$

$$-\left(\frac{31}{18}\right)^2 = x^4/12 \Big|_0^2 + x^3/9 \Big|_2^3 - \left(\frac{31}{18}\right)^2 = \frac{16}{12} - 0 + \frac{27}{9} - \frac{8}{9} - \left(\frac{31}{18}\right)^2 =$$

$$= 31/9 - \left(\frac{31}{18}\right)^2 = 31/9(1 - 31/36) = 155/324,$$

$$\sigma(X) = \sqrt{D(X)} = \sqrt{155/324} = \sqrt{155}/18.$$

$$P(1 < x < 5) = \int_{1}^{5} f(x)dx = \int_{1}^{2} \frac{xdx}{3} + \int_{2}^{3} \frac{1dx}{3} + \int_{3}^{5} 0dx =$$

$$= x^2/6 \Big|_1^2 + 1/3x \Big|_2^3 = 4/6 - 1/6 + 1 - 2/3 = 5/6.$$

Задачи

5.1. Непрерывная случайная величина *X* задана функцией распределения:

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

Найти дифференциальную функцию распределения *f* (*x*), а также *P*(–1/2 < *X* < 1/2).

5.2. Непрерывная случайная величина *X* задана функцией распределения:

$$F(x) = \begin{cases} 0 \text{ при } x \le \pi/6, \\ -\cos 3x \text{ при } \pi/6 < x \le \pi/3, \\ 1 \text{ при } x > \pi/3. \end{cases}$$

Найти дифференциальную функцию распределения *f* (*x*), а также *P*(2π /9 < *X* < π /2).

5.3. Непрерывная случайная величина *X* задана плотностью распределения:

$$f(x) = \begin{cases} 0 \text{ при } x \le 2, \\ c \cdot x \text{ при } 2 < x \le 4, \\ 0 \text{ при } x > 4. \end{cases}$$

Найти: а) число *c*; б) *M*(*X*), *D*(*X*).

5.4. Непрерывная случайная величина *X* задана плотностью распределения:

$$f(x) = \begin{cases} 0 \text{ при } x \le 0, \\ c \cdot \sqrt{x} \text{ при } 0 < x \le 1, \\ 0 \text{ при } x > 1. \end{cases}$$

Найти: а) число *c*; б) *M*(*X*), *D*(*X*).

5.5. Задана плотность распределения вероятностей непрерывной случайной величины *X*:

$$f(x) = \begin{cases} \dfrac{5-x}{2} \text{ при } x \epsilon [3; 5], \\ 0 \text{ при } x \notin [3; 5], \end{cases}$$

Найти: а) *F*(*x*) и построить ее график; б) *M*(*X*), *D*(*X*), σ(*X*); в) вероятность того, что в четырех независимых испытаниях величина *X* примет ровно 2 раза значение, принадлежащее интервалу (1;4).

5.6. Задана плотность распределения вероятностей непрерывной случайной величины *X*:

$$f(x) = \begin{cases} 2(x-2) \text{ при } x \epsilon [2; 3], \\ 0 \text{ при } x \notin [2; 3], \end{cases}$$

Найти: а) *F*(*x*) и построить ее график; б) *M*(*X*), *D*(*X*), σ(*X*); в) вероятность того, что в трех независимых испытаниях величина *X* примет ровно 2 раза значение, принадлежащее отрезку [1; 2,5].

5.7. Функция *f*(*x*) задана в виде:

$$f(x) = \begin{cases} \dfrac{c}{\sqrt{1-x^2}} \text{ при } x \epsilon [-\sqrt{3}/2; \sqrt{3}/2], \\ 0 \text{ при } x \notin [-\sqrt{3}/2; \sqrt{3}/2]. \end{cases}$$

Найти: а) значение постоянной *c*, при которой функция будет плотностью вероятности некоторой случайной величины *X*; б) функцию распределения *F*(*x*).

5.8.Функция *f*(*x*) задана в виде:

$$f(x) = \begin{cases} \dfrac{c}{\cos^2 x} \text{ при } x \epsilon [-\pi/4; \pi/4], \\ 0 \text{ при } x \notin [-\pi/4; \pi/4]. \end{cases}$$

Найти: а) значение постоянной *c*, при которой функция будет плотностью вероятности некоторой случайной величины *X*; б) функцию распределения *F*(*x*).

5.9. Случайная величина X, сосредоточенная на интервале (3;7), задана функцией распределения $F(x)= \frac{1}{16}(x^2-6x+9)$. Найти вероятность того, что случайная величина X примет значение: а) меньше 5, б) не меньше 7.

5.10. Случайная величина X, сосредоточенная на интервале (–1;4), задана функцией распределения $F(x)=\frac{x}{5}+\frac{1}{5}$. Найти вероятность того, что случайная величина X примет значение: а) меньше 2, б) меньше 4.

5.11. Случайная величина задана дифференциальной функцией распределения:

$$f(x) = \begin{cases} \dfrac{c \cdot \ln x}{x} \text{ при } x \epsilon [1; e], \\ 0 \text{ при } x \notin [1; e]. \end{cases}$$

Найти: а) число c; б) $M(X)$; в) вероятность $P(X > M(X))$.

5.12. Случайная величина задана дифференциальной функцией распределения:

$$f(x) = \begin{cases} \dfrac{1}{2}\sin x \quad \text{при } x \epsilon [0; \pi], \\ 0 \text{ при } x \notin [0; \pi], \end{cases}$$

Найти: а) $M(X)$; б) вероятность $P(X \leq M(X))$.

5.13. Распределение Ремя задается плотностью вероятности:

$$f(x) = \begin{cases} 0 \text{ при } x < 0, \\ x \cdot e^{\frac{-x^2}{2}} \text{при } x \geq 0. \end{cases}$$

Доказать, что $f(x)$ действительно является плотностью распределения вероятностей.

5.14. Задана плотность распределения вероятностей непрерывной случайной величины X:

$$f(x) = \begin{cases} 0 \text{ при } x < 0, \\ c \cdot x \cdot e^{-x} \text{при } x \geq 0. \end{cases}$$

Найти число c.

5.15. Случайная величина X распределена по закону Симпсона (равнобедренного треугольника) на отрезке [–2;2] (рис. 5.4). Найти аналитическое выражение для плотности вероятности $f(x)$ на всей числовой оси.

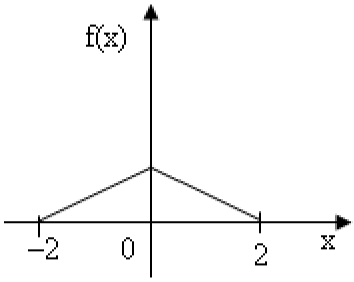


Рис. 5.4

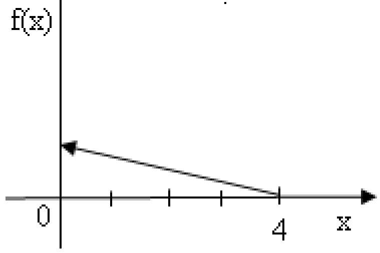


Рис. 5.5

5.16. Случайная величина *X* распределена по закону "прямоугольного треугольника" в интервале (0;4) (рис. 5.5). Найти аналитическое выражение для плотности вероятности *f*(*x*) на всей числовой оси.

Ответы

5.1. $f(x) = \left[\begin{array}{l} 0 \text{ при } x \le 0, \\ \frac{2x}{(x+1)^2} \text{ при } 0 < x \le 1, \\ 1 \text{ при } x > 1. \end{array}\right.$

P (–1/2<X<1/2)=2/3.

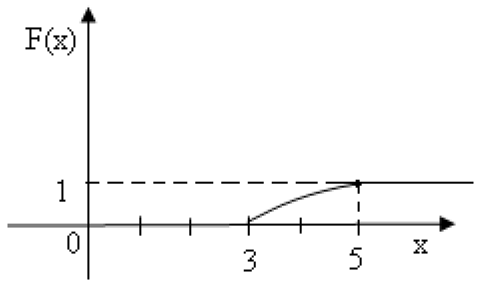
5.2. $f(x) = \left[\begin{array}{l} 0 \text{ при } x \le \pi/6, \\ 3sin3x \text{ при } \pi/6 < x \le \pi/3, \\ 0 \text{ при } x > \pi/3. \end{array}\right.$

P (2π /9<X< π /2)=1/2.

5.3. а) c =1/6, б) $M(X)=3\frac{1}{9}$, в) $D(X)$=26/81.

5.4. а) c=3/2, б) $M(X)$=3/5, в) $D(X)$=12/175.

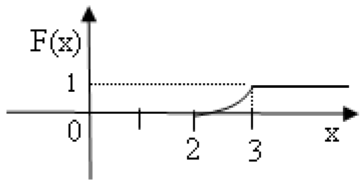
5.5. $a)F(x) = \left[\begin{array}{l} 0 \text{ при } x \le 3, \\ 1 - \frac{(5-x)^2}{4} \text{ при } 3 < x \le 5, \\ 1 \text{ при } x > 5. \end{array}\right.$



б) $M(X)=3\frac{2}{3}$, $D(X)$=2/9, σ $(X)=\sqrt{2}/3$.

в) 3/8.

5.6. $a)F(x) = \left[\begin{array}{l} 0 \text{ при } x \le 2, \\ (x-2)^2 \text{при } 2 < x \le 3, \\ 1 \text{ при } x > 3. \end{array}\right.$



б) $M(X)=2\frac{2}{3}$, $D(X)=3\frac{7}{12}$, σ $(X)=\sqrt{\frac{43}{12}} \approx 1{,}893$.

в) 9/64.

5.7. а) с $=\frac{3}{2\pi}$; б) $F(x) = \left[\begin{array}{l} 0 \text{ при } x \le\ -\sqrt{3}/2, \\ \frac{3}{2\pi}\arcsin x + \frac{1}{2} \text{при } -\sqrt{3}/2 < x \le \sqrt{3}/2 \\ 1 \text{ при } x > \sqrt{3}/2. \end{array}\right.$

5.8. а) c =1/2; б) $F(x) = \left[\begin{array}{c} 0 \text{ при } x \le -\pi/4, \\ 1/2 \cdot (\text{tg}\, x + 1) \text{ при } -\pi/4 < x \le \pi/4 \\ 1 \text{ при } x > \pi/4. \end{array}\right.$

5.9. а)1/4; б) 0.

5.10. а)3/5; б) 1.

5.11. а) c = 2; б) $M(X)$= 2; в) $1 - ln^2 2 \approx 0{,}5185$.

5.12. а) $M(X) = \pi/2$; б) 1/2

5.14. $c = 1$.

5.15. $f(x) = \begin{cases} \frac{2-|x|}{4} \text{ при } x \in [-2; 2], \\ 0 \text{ при } x \notin [-2; 2]. \end{cases}$

5.16. $f(x) = \begin{cases} \frac{4-x}{8} \text{ при } x \in [0; 4], \\ 0 \text{ при } x \notin [0; 4]. \end{cases}$

§6. Некоторые законы распределения непрерывной случайной величины

Равномерный закон распределения

Определение 1. Непрерывная случайная величина X имеет равномерный закон распределения на некотором интервале $(a;b)$, которому принадлежат все возможные значения X, если плотность распределения вероятностей $f(x)$ постоянная на этом интервале и равна 0 вне его, т.е.

$$f(x) = \begin{cases} 0 \text{ при } x \leq a, \\ \frac{1}{b-a} \text{ при } a < x \leq b, \\ 0 \text{ при } x \geq b. \end{cases}$$

График функции $f(x)$ изображен на рис. 6.1.

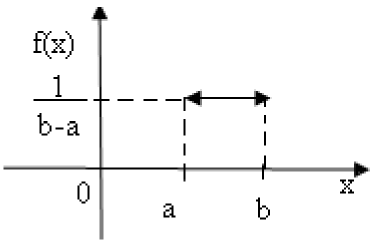


Рис. 6.1

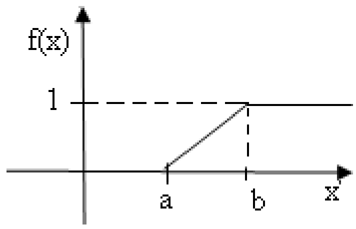


Рис. 6.2

Функция распределения случайной величины X, распределенной по равномерному закону, задается формулой:

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

Ее график изображен на рис. 6.2.

Числовые характеристики случайной величины, равномерно распределенной на интервале $(a;b)$, вычисляются по формулам:

$$M(X) = \frac{a+b}{2},\ D(X) = \frac{(b-a)^2}{12},\ \sigma(X) = \frac{b-a}{2\sqrt{3}}.$$

Задача 6.1. Случайная величина X равномерно распределена на отрезке [3;7]. Найти:

а) плотность распределения вероятностей $f(x)$ и построить ее график;

б) функцию распределения $F(x)$ и построить ее график;

в) $M(X)$, $D(X)$, $\sigma(X)$.

Решение: Воспользовавшись формулами, рассмотренными выше, при $a = 3$, $b = 7$, находим:

$$\text{а) } f(x) = \begin{cases} 0 \text{ при } x < 3, \\ \frac{1}{4} \text{ при } 3 \le x \le 7, \\ 0 \text{ при } x > 7. \end{cases}$$

Построим ее график (рис. 6.3):

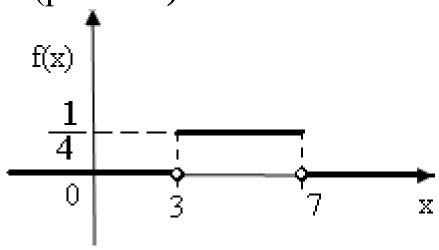


Рис. 6.3

$$\text{б) } F(x) = \begin{cases} 0 \text{ при } x \le 3, \\ \frac{x-3}{4} \text{ при } 3 < x \le 7, \\ 1 \text{ при } x > 7. \end{cases}$$

Построим ее график (рис. 6.4):

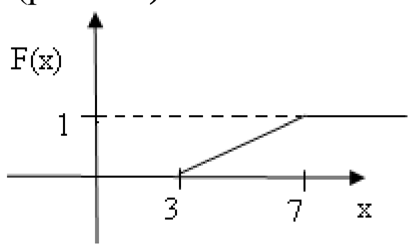


Рис. 6.4

в) $M(X) = \frac{a+b}{2} = \frac{3+7}{2} = 5$, $D(X) = \frac{(b-a)^2}{12} = \frac{(7-3)^2}{12} = \frac{4}{3}$, $\sigma(X) = \frac{b-a}{2\sqrt{3}} = \frac{2}{\sqrt{3}} = \frac{2\sqrt{3}}{3}$.

Показательный (экспоненциальный) закон распределения

Определение 2. Непрерывная случайная величина X имеет *показательный (экспоненциальный) закон распределения* с параметром $\lambda > 0$, если функция плотности распределения вероятностей имеет вид:

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

Функция распределения случайной величины X, распределенной по показательному закону, задается формулой:

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

Кривая распределения $f(x)$ и график функции распределения $F(x)$ случайной величины X приведены на рис. 6.5 и рис. 6.6.

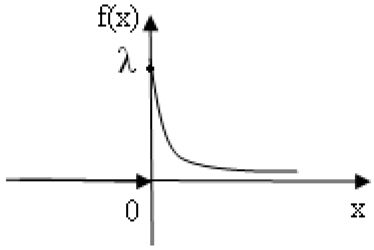


Рис. 6.5

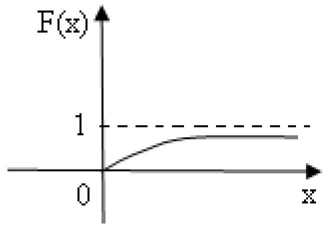


Рис. 6.6

Математическое ожидание, дисперсия и среднее квадратическое отклонение показательного распределения соответственно равны:

$$M(X) = \frac{1}{\lambda},\ D(X) = \frac{1}{\lambda^2},\ \sigma(X) = \frac{1}{\lambda}.$$

Таким образом, математическое ожидание и среднее квадратическое отклонение показательного распределения равны между собой.

Вероятность попадания X в интервал $(a;b)$ вычисляется по формуле:

$$P(a < X < b) = \mathrm{e}^{-\lambda a} - \mathrm{e}^{-\lambda b}, \text{ если } (a;b) \in [0; +\infty)$$

Задача 6.2. Среднее время безотказной работы прибора равно 100 ч. Полагая, что время безотказной работы прибора имеет показательный закон распределения, найти:

а) плотность распределения вероятностей;

б) функцию распределения;

в) вероятность того, что время безотказной работы прибора превысит 120 ч.

Решение. По условию математическое ожидание $M(X) = \frac{1}{\lambda} = 100$, откуда $\lambda = 1/100 = 0{,}01$.

Следовательно,

а) $f(x) = \begin{cases} 0 \text{ при } x < 0, \\ 0{,}01\mathrm{e}^{-0{,}01x} \text{ при } x \geq 0. \end{cases}$

б) $F(x) = \begin{cases} 0 \text{ при } x < 0, \\ 1 - \mathrm{e}^{-0{,}01x} \text{ при } x \geq 0. \end{cases}$

в) Искомую вероятность найдем, используя функцию распределения:

$$P(X > 120) = 1 - F(120) = 1 - (1 - e^{-1{,}2}) = e^{-1{,}2} \approx 0{,}3.$$

Нормальный закон распределения

Определение 3. Непрерывная случайная величина X имеет *нормальный закон распределения (закон Гаусса),* если ее плотность распределения имеет вид:

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

где $m = M(X)$, $\sigma^2 = D(X)$, $\sigma > 0$.

Кривую нормального закона распределения называют *нормальной или гауссовой кривой* (рис. 6.7).

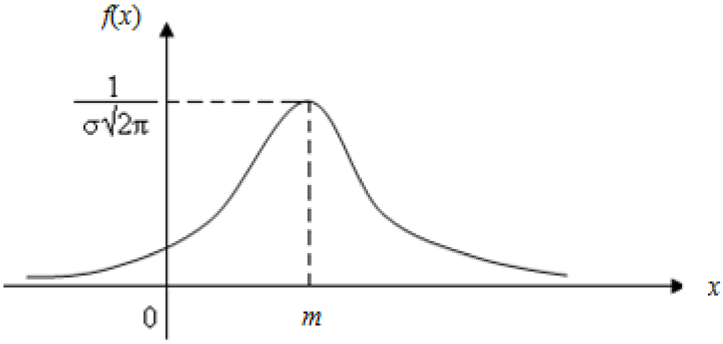


Рис. 6.7

Нормальная кривая симметрична относительно прямой $x = m$, имеет максимум в точке $x = m$, равный $\frac{1}{\sigma\sqrt{2\pi}}$.

Функция распределения случайной величины X, распределенной по нормальному закону, выражается через функцию Лапласа $\Phi(x)$ по формуле:

$$F(x) = \Phi\left(\frac{x-m}{\sigma}\right) + \frac{1}{2},$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{\frac{-t^2}{2}} dt$$ – функция Лапласа.

Замечание. Функция $\Phi(x)$ является нечетной ($\Phi(-x) = -\Phi(x)$), кроме того, при $x > 5$ можно считать $\Phi(x) \approx 1/2$.

Таблица значений функции $\Phi(x)$ приведена в приложении (табл. П 2.2).

График функции распределения $F(x)$ изображен на рис. 6.8.

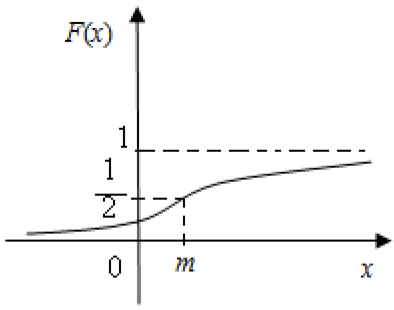


Рис. 6.8

Вероятность того, что случайная величина X примет значения, принадлежащие интервалу $(a;b)$ вычисляются по формуле:

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

Вероятность того, что абсолютная величина отклонения случайной величины от ее математического ожидания меньше положительного числа δ вычисляется по формуле:

$$P(|X - m| < \delta) = 2 \cdot \Phi\left(\frac{\delta}{\sigma}\right).$$

В частности, при m=0 справедливо равенство:

$$P(|X| < \delta) = 2 \cdot \Phi\left(\frac{\delta}{\sigma}\right).$$

"Правило трех сигм"

Если случайная величина X имеет нормальный закон распределения с параметрами m и σ, то практически достоверно, что ее значения заключены в интервале $(m - 3\sigma; m + 3\sigma)$, так как $P(|X - m| < 3\delta) = 0{,}9973$.

Задача 6.3. Случайная величина X распределена нормально с математическим ожиданием 32 и дисперсией 16. Найти: а) плотность распределения вероятностей $f(x)$; б) вероятность того, что в результате испытания X примет значение из интервала (28;38).

Решение: По условию $m = 32$, $\sigma^2 = 16$, следовательно, $\sigma = 4$, тогда

а) $f(x) = \frac{1}{4\sqrt{2\pi}} \cdot e^{\frac{-(x-32)^2}{32}}$

б) Воспользуемся формулой:

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

Подставив $a = 28$, $b = 38$, $m = 32$, $\sigma = 4$, получим

$P\,(28<X<38)=\Phi\left(\frac{38-32}{4}\right) - \Phi(\frac{28-32}{4}) = \Phi(1{,}5)+\Phi(1)$

По таблице значений функции $\Phi(x)$ находим $\Phi(1{,}5) = 0{,}4332$, $\Phi(1) = 0{,}3413$.

Итак, искомая вероятность:

$P(28<X<38) = 0{,}4332+0{,}3413=0{,}7745$.

Задачи

6.1. Случайная величина X равномерно распределена в интервале (–3;5). Найдите:

а) плотность распределения $f(x)$;

б) функции распределения $F(x)$;

в) числовые характеристики;

г) вероятность $P(4<x<6)$.

6.2. Случайная величина X равномерно распределена на отрезке [2;7]. Найдите:

а) плотность распределения $f(x)$;

б) функцию распределения $F(x)$;

в) числовые характеристики;

г) вероятность $P(3 \le x \le 6)$.

6.3. На шоссе установлен автоматический светофор, в котором 2 минуты для транспорта горит зеленый свет, 3 секунды – желтый и 30 секунд – красный и т.д. Машина проезжает по шоссе в случайный момент времени. Найти вероятность того, что машина проедет мимо светофора, не останавливаясь.

6.4. Поезда метрополитена идут регулярно с интервалом 2 минуты. Пассажир выходит на платформу в случайный момент времени. Какова вероятность того, что ждать поезд пассажиру придется больше 50 секунд. Найти математическое ожидание случайной величины *X* – время ожидания поезда.

6.5. Найти дисперсию и среднее квадратическое отклонение показательного распределения, заданного функцией распределения:

$$F(x) = \begin{cases} 0 \text{ при } x < 0, \\ 1 - e^{-8x} \text{при } x \geq 0. \end{cases}$$

6.6. Непрерывная случайная величина *X* задана плотностью распределения вероятностей:

$$f(x) = \begin{cases} 0 \text{ при } x < 0, \\ 0{,}7 \cdot e^{-0{,}7x} \text{при } x \geq 0. \end{cases}$$

а) Назовите закон распределения рассматриваемой случайной величины.

б) Найдите функцию распределения $F(x)$ и числовые характеристики случайной величины *X*.

6.7. Случайная величина *X* распределена по показательному закону, заданному плотностью распределения вероятностей:

$$f(x) = \begin{cases} 0 \text{ при } x < 0, \\ 0{,}4 \cdot e^{-0{,}4x} \text{при } x \geq 0. \end{cases}$$

Найти вероятность того, что в результате испытания *X* примет значение из интервала (2,5;5).

6.8. Непрерывная случайная величина *X* распределена по показательному закону, заданному функцией распределения:

$$F(x) = \begin{cases} 0 \text{ при } x < 0, \\ 1 - e^{-0{,}6x} \text{при } x \geq 0. \end{cases}$$

Найти вероятность того, что в результате испытания *X* примет значение из отрезка [2;5].

6.9. Математическое ожидание и среднее квадратическое отклонение нормально распределенной случайной величины соответственно равны 8 и 2. Найдите:

а) плотность *распределения* $f(x)$;

б) вероятность того, что в результате испытания *X* примет значение из интервала (10;14).

6.10. Случайная величина *X* распределена нормально с математическим ожиданием 3,5 и дисперсией 0,04. Найдите:

а) плотность распределения $f(x)$;

б) вероятность того, что в результате испытания *X* примет значение из отрезка [3,1; 3,7].

6.11. Случайная величина X распределена нормально с $M(X) = 0$ и $D(X)=1$. Какое из событий: $|X|\leq 0,6$ или $|X|\geq 0,6$ имеет большую вероятность?

6.12. Случайная величина X распределена нормально с $M(X) = 0$ и $D(X)=1$.Из какого интервала (–0,5; –0,1) или (1; 2) при одном испытании она примет значение с большей вероятностью?

6.13. Текущая цена за одну акцию может быть смоделирована с помощью нормального закона распределения с $M(X)=10$ ден. ед. и $\sigma(X) = 0,3$ ден. ед. Найти:

а) вероятность того, что текущая цена акции будет от 9,8 ден. ед. до 10,4 ден. ед.;

б) с помощью "правила трех сигм" найти границы, в которых будет находиться текущая цена акции.

6.14. Производится взвешивание вещества без систематических ошибок. Случайные ошибки взвешивания подчинены нормальному закону со средним квадратическим отклонением σ = 5г. Найти вероятность того, что в четырех независимых опытах ошибка при трех взвешиваниях не превзойдет по абсолютной величине 3 г.

6.15. Случайная величина X распределена нормально с $M(X)=12,6$. Вероятность попадания случайной величины в интервал (11,4; 13,8) равна 0,6826. Найдите среднее квадратическое отклонение σ.

6.16. Случайная величина X распределена нормально с $M(X) = 12$ и $D(X) = 36$. Найти интервал, в который с вероятностью 0,9973 попадет в результате испытания случайная величина X.

6.17. Деталь, изготовленная автоматом, считается бракованной, если отклонение X ее контролируемого параметра от номинала превышает по модулю 2 единицы измерения. Предполагается, что случайная величина X распределена нормально с $M(X) = 0$ и $\sigma(X) = 0,7$. Сколько процентов бракованных деталей выдает автомат?

3.18. Параметр X детали распределен нормально с математическим ожиданием 2, равным номиналу, и средним квадратическим отклонением 0,014. Найти вероятность того, что отклонение X от номинала по модулю не превысит 1 % номинала.

Ответы

6.1.

а)$f(x) = \begin{cases} 0 \text{ при } x \leq \text{-}3, \\ \frac{1}{8} \text{при} -3 < x < 5, \\ 0 \text{ при } x \geq 5. \end{cases}$ б)$F(x) = \begin{cases} 0 \text{ при } x \leq -3, \\ \frac{x+3}{8} \text{ при} -3 < x < 5, \\ 1 \text{ при } x \geq 5. \end{cases}$

в) $M(X)=1$, $D(X)=16/3$, $\sigma(X)= 4/\sqrt{3}$, г)1/8.

6.2.

а) $f(x) = \begin{cases} 0 \text{ при } x < 2, \\ \frac{1}{5} \text{при } 2 \le x \le 7, \\ 1 \text{ при } x > 7. \end{cases}$ б) $F(x) = \begin{cases} 0 \text{ при } x \le 2, \\ \frac{x-2}{7} \text{ при } 2 < x \le 7, \\ 1 \text{ при } x > 7. \end{cases}$

в) $M(X)$=4,5, $D(X)$ =$2\frac{1}{2}$, σ (X)= $\frac{5}{2\sqrt{3}}$, г)3/5.

6.3. 40/51.

6.4. 7/12, $M(X)$=1.

6.5. $D(X)$ = 1/64, σ (X)=1/8

6.6. $F(x) = \begin{cases} 0 \text{ при } x < 0, \\ 1 - e^{-0,7x} \text{ при } x \ge 0, \end{cases}$ $M(X)$=$1\frac{3}{7}$, $D(X)$ =$2\frac{2}{49}$, σ (X)= $1\frac{3}{7}$.

6.7. P(2,5<X<5)=$e^{-1}-e^{-2}$≈0,2325

6.8. P(2≤X≤5)=0,252.

6.9.

а) $f(x) = \frac{1}{2\sqrt{2\pi}} \cdot e^{\frac{-(x-8)^2}{8}}$,

б) $P(10 < X < 14) \approx 0{,}1574$.

а) $f(x) = \frac{5}{\sqrt{2\pi}} \cdot e^{\frac{25(x-3,5)^2}{2}}$,

б) $P(3{,}1 \le X \le 3{,}7) \approx 0{,}8185$.

6.10.

6.11. $|x|$≥0,6.

6.12. (–0,5; –0,1).

6.13. а) P(9,8 ≤ X ≤ 10,4) ≈ 0,6562

б) (9,1; 10,9).

6.14. 0,111.

6.15. σ = 1,2.

6.16. (–6; 30).

6.17. 0,4 %.

6.18. 0,8472.

Глава 3. Элементы математической статистики

§7. Статистическое распределение выборки

Задачи математической статистики

Математическая статистика разрабатывает методы планирования и анализы эксперимента.

К типичным задачам математической статистики относятся:

- задача определения закона распределения случайной величины по статистическим данным;
- задача нахождения неизвестных параметров распределения случайной величины;
- задача проверки правдоподобия выдвигаемых по статистическим данным гипотез о законе распределения случайной величины, о её параметрах.

Генеральная и выборочная совокупности

Определение 1. *Выборочной совокупностью* (или *выборкой*) называется совокупность случайно отобранных объектов.

Определение 2. Генеральной совокупностью называется совокупность всех однородных объектов, из которых производится выборка.

Определение 3. Объемом совокупности (генеральной или выборочной) называется число объектов этой совокупности.

Определение 4. Выборка называется представительной (или репрезентативной), если она осуществлена случайным образом, когда все объекты генеральной совокупности имели равные вероятности попасть в выборку.

Определение 5. Статистическим рядом, соответствующим полученной случайной выборке, называется набор значений (вариант) качественного или количественного признака объектов выборки, которые располагают в порядке возрастания.

Определение 6. Интервальным статистическим рядом, соответствующим полученной случайной выборке, называется упорядоченная последовательность интервалов $[a_i; a_{i+1})$, $i = 1,2,\ldots$ с указанием количества m_i значений x_i, попавших в них (табл. 7.1)

Таблица 7.1

Интервальный статистический ряд

№	$[a_i ; a_{i+1})$	m_i
1	$[a_1 ; a_2)$	m_1
2	$[a_2 ; a_3)$	m_2
…	… … …	…
k	$[a_k ; a_{k+1}]$	m_k

Причем, если n – объем выборки, то $n = m_1 + m_2 + \ldots + m_k$. Интервалы $[a_1 ; a_2)$, $[a_2 ; a_3)$, …, $[a_k ; a_{k+1}]$ имеют не обязательно равные длины. Число k не должно быть большим, но и не малым. Обычно берут $7 \le k \le 20$.

Замечание 1. Иногда для упрощения исследования интервальный статистический ряд заменяют дискретным рядом, где в качестве значений исследуемого признака берут середины или одну из границ соответствующих интервалов.

Статистическое распределение выборки. Полигон и гистограмма

Определение 7. Статистической функцией распределения (или функцией распределения выборки) называется функция $\tilde{F}(x)$, задающая для каждого значения x статистического ряда относительную частоту события $X < x$,

т.е. $\tilde{F}(x)= \frac{m_x}{n}$,

где n – объем выборки; m_x – число выборочных значений, меньших x.

Свойства функции $\tilde{F}(x)$

1. $0 \le \tilde{F}(x) \le 1$;
2. $\tilde{F}(x)$ – неубывающая функция;
3. $\tilde{F}(-\infty) = 0$, $\tilde{F}(+\infty)=1$.

Замечание 2. В дальнейшем интегральную функцию распределения $F(x)$ генеральной совокупности X будем называть *теоретической,* а функцию $\tilde{F}(x)$ – *эмпирической* функцией распределения. Отличие между ними состоит в том, что $F(x)$ – вероятность события $X < x$, а $\tilde{F}(x)$ – его суммарная относительная частота в n опытах. Однако функции $\tilde{F}(x)$ и $F(x)$ обладают одинаковыми свойствами.

Определение 8. *Полигоном* (или *многоугольником*) статистического распределения называется ломаная линия на плоскости Oxy, соединяющая точки $(x_i\,;\, m_i/\,n)$, $i = 1, \dots k$,

где n – объем выборки; x_i – значения статистического ряда; m_i – число значений x_i в этом ряде (рис. 7.1).

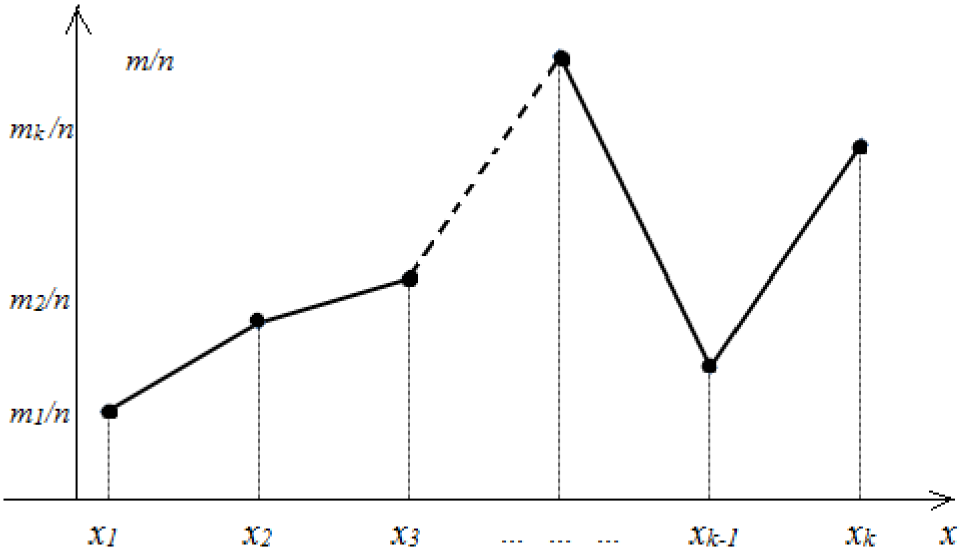


Рис. 7.1

Определение 9. *Гистограммой* интервального статистического ряда называется ступенчатая фигура, построенная по правилу: на плоскости Oxy на отрезках, изображающих интервалы статистического ряда, как на основаниях, строят прямоугольники с высотами, равными относительным частотам соответствующих интервалов (рис. 7.2).

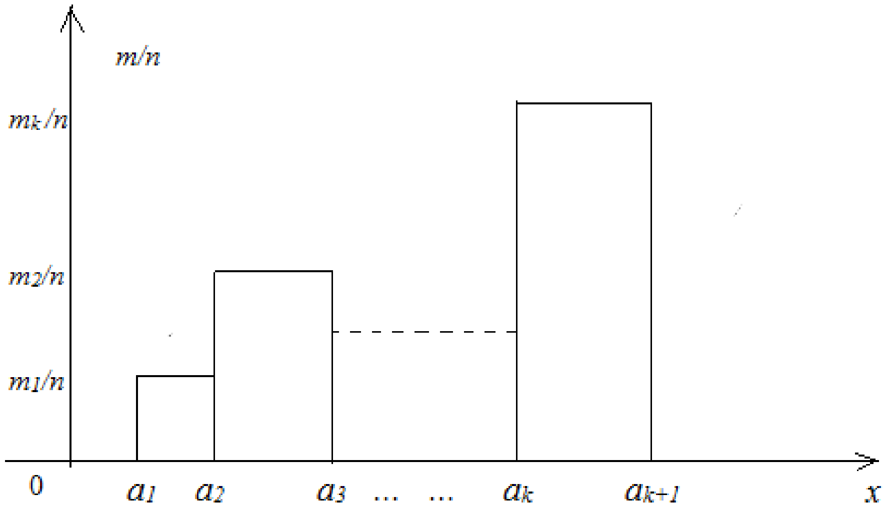


Рис. 7.2

Замечание 3. Полигон и гистограмма являются графическими приближениями дифференциальной функции распределения исследуемой случайной величины.

Задачи с решениями

Задача 7.1. В результате эксперимента получены следующие значения случайной величины X:

3; 6; 8; 11; 6; 10; 7; 9; 7; 3; 4; 8;
7; 9; 4; 9; 11; 7; 8; 4; 10; 5; 6; 7; 2.

Требуется: а) составить статистический ряд;
б) построить статистическое распределение;
в) изобразить полигон распределения.

Решение. а) Объем выборки n = 25.

Построим статистический ряд данной выборки: в первой строке таблицы укажем все различные значения, принимаемые случайной величиной X; во второй строке укажем, сколько раз она приняла эти значения.

Таблица 7.2

x_i	2	3	4	5	6	7	8	9	10	11
m_i	1	2	3	1	3	5	3	3	2	2

б) Найдем статистическое распределение случайной величины X, для чего в табл. 7.2 заменим вторую строку строкой, содержащей относительные частоты $\frac{m_i}{n}$.

Таблица 7.3

x_i	2	3	4	5	6	7	8	9	10	11
$\frac{m_i}{n}$	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{3}{25}$	$\frac{1}{25}$	$\frac{3}{25}$	$\frac{1}{5}$	$\frac{3}{25}$	$\frac{3}{25}$	$\frac{2}{25}$	$\frac{2}{25}$

Контроль:

$$\frac{1}{25}+\frac{2}{25}+\frac{3}{25}+\frac{1}{25}+\frac{3}{25}+\frac{1}{5}+\frac{3}{25}+\frac{3}{25}+\frac{2}{25}+\frac{2}{25}=1.$$

в) На плоскости Oxy построим точки:

$$\left(2;\ \frac{1}{25}\right),\left(3;\ \frac{2}{25}\right),\left(4;\ \frac{3}{25}\right),\left(5;\ \frac{1}{25}\right),\left(6;\ \frac{3}{25}\right),\left(7;\ \frac{1}{5}\right),$$
$$\left(8;\ \frac{3}{25}\right),\left(9;\ \frac{3}{25}\right),\left(10;\ \frac{2}{25}\right),\left(11;\ \frac{2}{25}\right).$$

Соединим их (рис. 7.3). Полученная ломаная – полигон данного распределения.

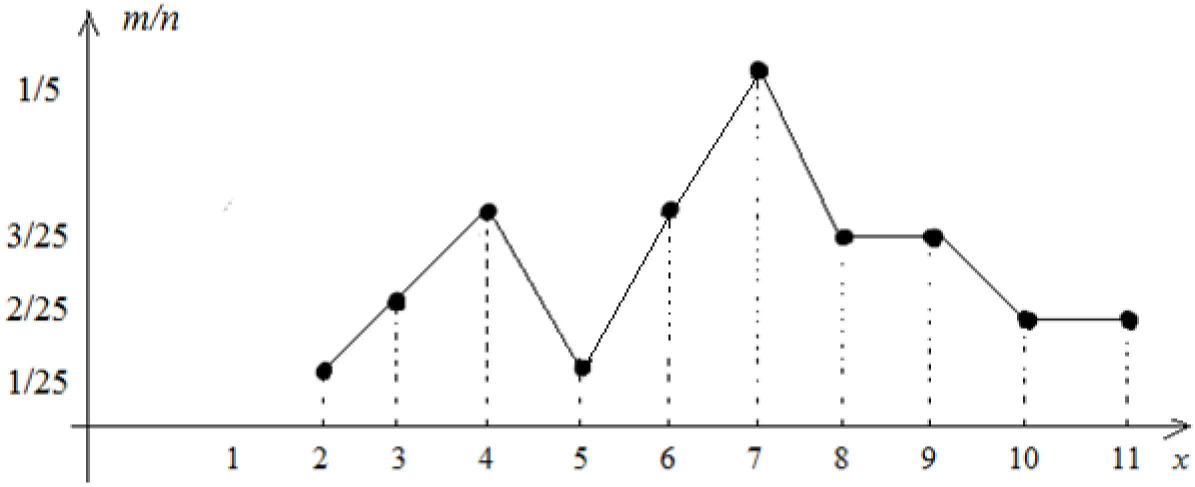


Рис. 7.3

Ответ: а) табл. 7.2, б) табл. 7.3, в) рис. 7.3.

Задача 7.2. В результате эксперимента получены следующие значения случайной величины X:

16; 17; 9; 13; 21; 11; 7; 7; 19; 5; 17; 5; 20;

18; 11; 4; 6; 22; 21; 15; 15; 23; 19; 25; 1.

Требуется: а) построить интервальный статистический ряд, разбив промежуток [0; 25] на 5 промежутков равной длины;

б) построить гистограмму относительных частот.

Решение. а) Объем выборки $n = 25$. По экспериментальным данным составим таблицу (табл. 7.4). В её первой строке укажем промежутки разбиения:

[0; 5), [5; 10), [10; 15), [15; 20) [20; 25].

Во второй строке укажем соответствующие числа m_i – сколько раз случайная величина X приняла значение из этого промежутка.

Таблица 7.4

$[a_i; a_{i+1})$	[0; 5)	[5; 10)	[10; 15)	[15; 20)	[20; 25]
m_i	2	6	3	8	6

Контроль: 2 + 6 + 3 + 8 + 6 = 25.

По табл. 7.4 составим интервальный статистический ряд, где во второй строке указаны относительные частоты (табл. 7.5).

Таблица 7.5

$[a_i; a_{i+1})$	[0; 5)	[5; 10)	[10; 15)	[15; 20)	[20; 25]
$\frac{m_i}{n}$	$\frac{2}{25}$	$\frac{6}{25}$	$\frac{3}{25}$	$\frac{8}{25}$	$\frac{6}{25}$

б) На оси *Ox* отложим промежутки:

[0; 5), [5; 10), [10; 15), [15; 20) [20; 25]

интервального статистического ряда, а на оси *Oy* – относительные частоты. Построив по этим данным прямоугольники с основаниями $[a_i; a_{i+1})$ и высотами $\frac{m_i}{n}$, получим ступенчатую фигуру – гистограмму (рис. 7.4).

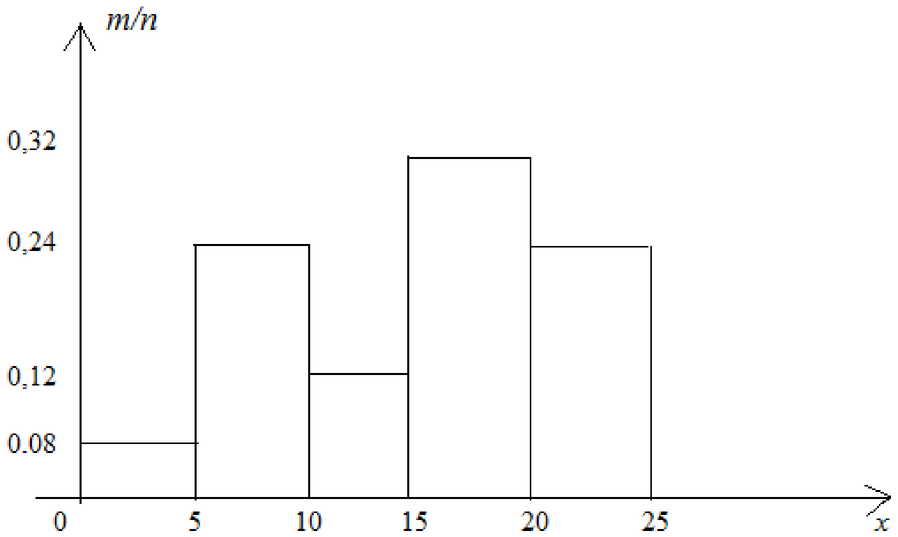


Рис. 7.4

Ответ: а) табл. 7.4; б) рис. 7.5.

Задача 7.3. Дан статистический ряд

x_i	15	16	17	18
$\frac{m_i}{n}$	0,4	0,1	0,3	0,2

Найти статистическую функцию распределения и построить её график.

Решение. Воспользовавшись формулой

$$\tilde{F}(x) = \frac{m_x}{n},$$

где n – объем выборки; m_x – число выборочных значений, меньших x, вычисляем:

$$\tilde{F}(x) = \begin{cases} 0 \text{ при } x \leq 15, \\ 0{,}4 \text{ при } 15 < x \leq 16, \\ 0{,}5 \text{ при } 16 < x \leq 17, \\ 0{,}8 \text{ при } 17 < x \leq 18, \\ 1 \text{ при } x \geq 18. \end{cases} \qquad (1)$$

Построим график функции $\tilde{F}(x)$.

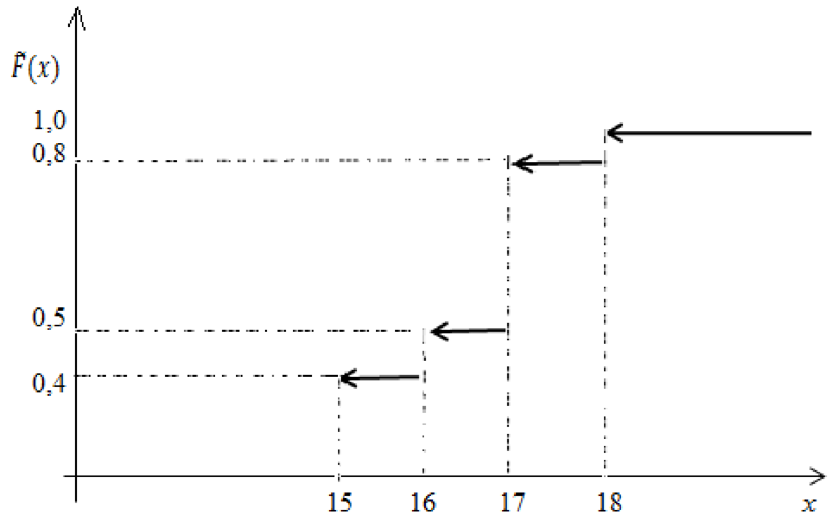


Рис. 7.5

Ответ: а) формула (1); б) рис. 7.5.

Задачи

7.1. В результате испытаний получены следующие статистические значения случайной величины *X*:

–0,9; 0,1; –2,9; 1,1; 5,1; 0,1; –6,9; 1,1; –3,9; –0,9;

5,1; –8,9; –2,9; 0,1; 1,1; 6,1; 3,1; 0,1; 1,1; –2,9;

–2,9; 0,1; –3,9; –0,9; 0,1; 3,1; 0,1; 1,1; 3,1; 0,1.

Требуется: а) составить статистический ряд;

б) найти статистическую функцию распределения $\tilde{F}(x)$;

в) изобразить полигон относительных частот.

7.2. В результате испытаний получены следующие статистические значения случайной величины *X*:

1,36; 1,37; 1,35; 1,31; 1,34; 1,36; 1,38; 1,35; 1,39; 1,40;

1,33; 1,34; 1,36; 1,35; 1,37; 1,41; 1,36; 1,34; 1,39; 1,36;

1,35; 1,37; 1,38; 1,40; 1,37; 1,36; 1,35; 1,34; 1,37; 1,38.

Требуется: а) составить статистический ряд;

б) найти статистическую функцию распределения $\tilde{F}(x)$;

в) изобразить полигон относительных частот.

7.3. В результате испытаний получены следующие статистические значения случайной величины X:

3,45; 3,47; 3,47; 3,43; 3, 46; 3,44; 3,40; 3,45; 3,41; 3,42;

3,47; 3,49; 3,41; 3,48; 3,43; 3,40; 3,43; 3,47; 3,45; 3,44;

3,41; 3,40; 3,48; 3,46; 3,51; 3,39; 3,50; 3,50; 3,47; 3,38;

3,44; 3,40; 3,40; 3,44; 3,47; 3,53; 3,46; 3,46; 3,52; 3,47;

3,41; 3,44; 3,47; 3,45; 3,44; 3,45; 3,47; 3,42; 3,44; 3,50;

3,45; 3,50; 3,42; 3,48; 3,40; 3,45; 3,48; 3,48; 3,46; 3,47;

3,44; 3,44; 3,47; 3,43; 3,44; 3,47; 3,44; 3,45; 3,44; 3,46;

3,46; 3,44; 3,44; 3,44; 3,44; 3,46; 3,44; 3,42; 3,50; 3,46;

3,48; 3,43; 3,40; 3,46; 3,46; 3,47; 3,45; 3,48; 3,42; 3,46;

3,48; 3,38; 3,45; 3,43; 3,52; 3,43; 3,50; 3,51; 3,41; 3,52.

Построить: а) интервальный статистический ряд;

б) статистический ряд, рассматривая в качестве значений середины интервалов;

в) статистическую функцию распределения $\tilde{F}(x)$;

г) гистограмму относительных частот.

7.4. При измерении диаметров ста подшипниковых шариков, выбранных из большой партии шариков для определения стандартности, получены следующие результаты:

8,31; 8,42; 8,37; 8,40; 8,40; 8,30; 8,30; 8,42; 8,32; 8,29;

8,33; 8,36; 8,34; 8,37; 8,32; 8,36; 8,38,8,38; 8,33; 8,36;

8,40; 8,36; 8,32; 8,36; 8,36; 8,30; 8,30; 8,33; 8,35; 8,37;

8,37; 8,30; 8,41; 8,34; 8,33; 8,37; 8,34; 8,38; 8,29; 8, 34;

8,31; 8,36; 8,37; 8,30; 8,41; 8,34; 8,34; 8,37; 8,354 8,40;

8,34; 8,36; 8,37; 8,37; 8,41; 8,35; 8,38; 8,33; 8,36; 8,36;

8,36; 8,37; 8,36; 8,40; 8,37; 8,34; 8,37; 8,32; 8,35; 8,36;

8,37; 8,41; 8,36; 8,36; 8,36; 8,40; 8,34; 8,40; 8,34; 8,33;

8,35; 8,37; 8,34; 8,36; 8,37; 8,37; 8,35; 8,36; 8,34; 8,42;

8,36; 8,33; 8,34; 8,35; 8,36;8,32; 8,38; 8,32; 8,36; 8,37;

Построить: а) интервальный статистический ряд;

б) статистический ряд, рассматривая в качестве значений середины интервалов;

в) статистическую функцию распределения $\tilde{F}(x)$;

г) гистограмму относительных частот.

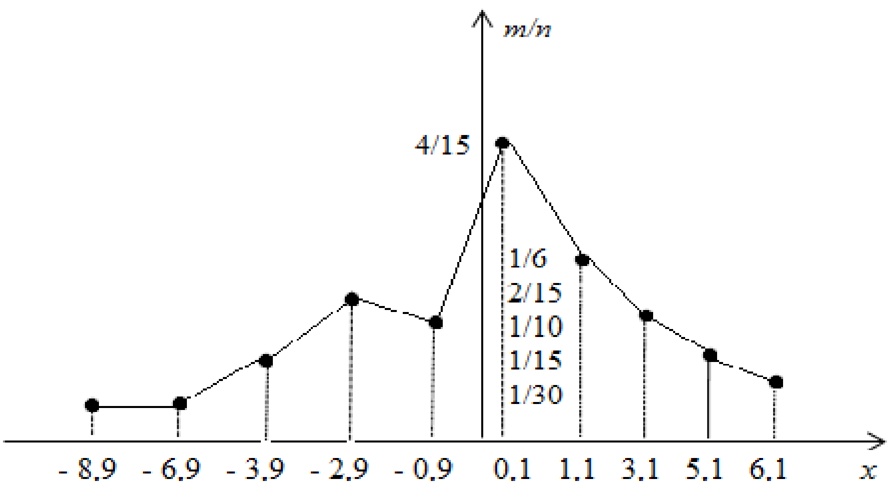
Ответы

7.1. а)

x_i	–8,9	–6,9	–3,9	–2,9	–0,9	0,1	1,1	3,1	5,1	6,1
m_i	1	1	2	4	3	8	5	3	2	1
$\frac{m_i}{n}$	$\frac{1}{30}$	$\frac{1}{30}$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{1}{10}$	$\frac{4}{15}$	$\frac{1}{6}$	$\frac{1}{10}$	$\frac{1}{15}$	$\frac{1}{30}$

б) $$\tilde{F}(x) = \begin{cases} 0 & \text{при} \quad x \le -8{,}9, \\ 1/30 & \text{при } -8{,}9 < x \le -6{,}9, \\ 1/15 & \text{при } -6{,}9 < x \le -3{,}9, \\ 2/15 & \text{при } -3{,}9 < x \le -2{,}9, \\ 4/15 & \text{при } -2{,}9 < x \le -0{,}9, \\ 11/30 & \text{при } -0{,}9 < x \le 0{,}1, \\ 19/30 & \text{при } 0{,}1 < x \le 1{,}1, \\ 4/5 & \text{при } 1{,}1 < x \le 3{,}1, \\ 9/10 & \text{при } 3{,}1 < x \le 5{,}1, \\ 29/30 & \text{при } 5{,}1 < x \le 6{,}1, \\ 1 & \text{при} \quad x > 6{,}1. \end{cases}$$

в)



7.2. а)

x_i	1,31	1,33	1,34	1,35	1,36	1,37	1,38	1,39	1,40	1,41
m_i	1	1	4	5	6	5	3	2	2	1
$\frac{m_i}{n}$	$\frac{1}{30}$	$\frac{1}{30}$	$\frac{2}{15}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{10}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{30}$

б) $$\tilde{F}(x) = \begin{cases} 0 & \text{при } x \le 1{,}31, \\ 1/30 & \text{при } 1{,}31 < x \le 1{,}33, \\ 1/15 & \text{при } 1{,}33 < x \le 1{,}34, \\ 1/5 & \text{при } 1{,}34 < x \le 1{,}35, \\ 11/30 & \text{при } 1{,}35 < x \le 1{,}36, \\ 17/30 & \text{при } 1{,}36 < x \le 1{,}37, \\ 11/15 & \text{при } 1{,}37 < x \le 1{,}38, \\ 5/6 & \text{при } 1{,}38 < x \le 1{,}39, \\ 9/10 & \text{при } 1{,}39 < x \le 1{,}40, \\ 29/30 & \text{при } 1{,}40 < x \le 1{,}41, \\ 1 & \text{при } x > 1{,}41. \end{cases}$$

7.3. а)

$[x_i;\ x_{i+1})$	m_i	$\frac{m_i}{n}$
[3,38; 3,40)	3	0,03
[3,40; 3,42)	12	0,12
[3,42; 3,44)	12	0,12
[3,44; 3,46)	27	0,27
[3,46; 3,48)	25	0,25
[3,48; 3,50)	9	0,09
[3,50; 3,52)	8	0,08
[3,52; 3,54]	4	0,04

Указания:

1) Из заданной выборки найти:
   $x_{наиб} = 3{,}53$, $x_{наим} = 3{,}38$, $x_{наиб} - x_{наим} = 0{,}15$.
2) Число интервалов определить по формуле:
   $k = 1+3{,}322 \cdot \lg n = 1 + 3{,}322 \cdot \lg 100 = 1+6{,}644 = 7{,}644 \approx 8$.
3) Взять в качестве шага, то есть длины интервалов, число:
   $0{,}15 : 8 = 0{,}019 \approx 0{,}02$.
4) Из данной выборки найти m_i – число значений, попавших в промежуток $[x_i;\ x_{i+1})$.

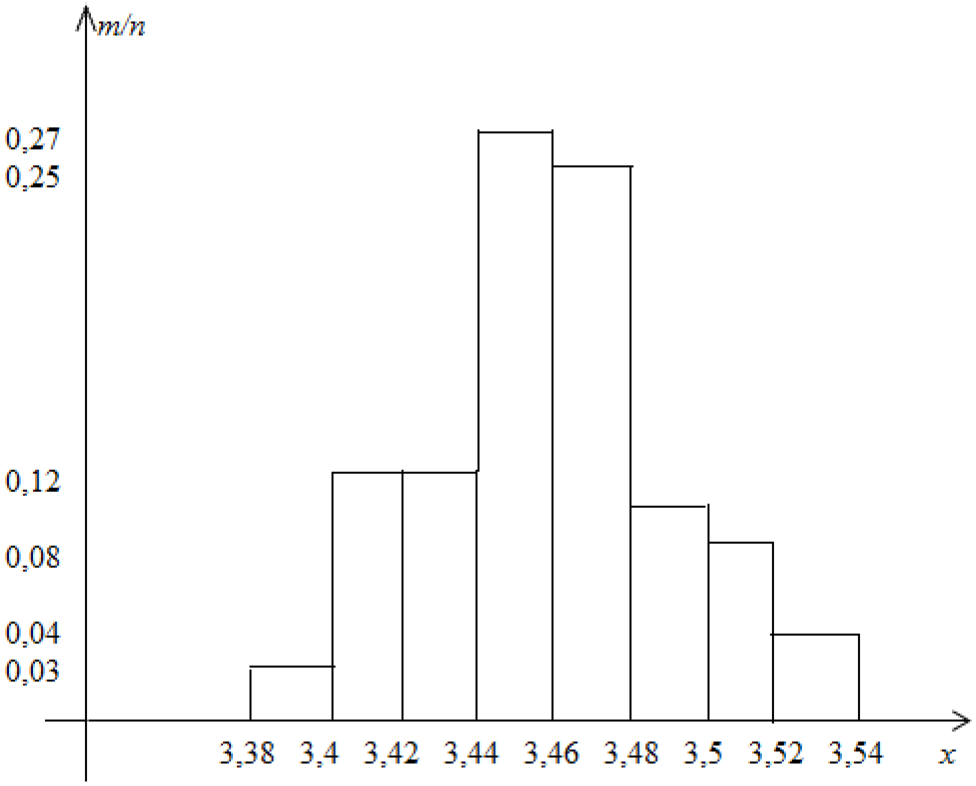
б)

$\bar{x}_i$	3,39	3,41	3,43	3,45	3,47	3,49	3,51	3,53
m_i	3	12	12	27	25	9	8	4
$\frac{m_i}{n}$	0,03	0,12	0,12	0,27	0,25	0,09	0,08	0,04

в)

$$\tilde{F}(x) = \begin{cases} 0 & \text{при} \quad x \le 3{,}39, \\ 0{,}03 & \text{при } 3{,}39 < x \le 3{,}41, \\ 0{,}15 & \text{при } 3{,}41 < x \le 3{,}43, \\ 0{,}27 & \text{при } 3{,}43 < x \le 3{,}45, \\ 0{,}54 & \text{при } 3{,}45 < x \le 3{,}47, \\ 0{,}79 & \text{при } 3{,}47 < x \le 3{,}49, \\ 0{,}88 & \text{при } 3{,}49 < x \le 3{,}51, \\ 0{,}96 & \text{при } 3{,}51 < x \le 3{,}53, \\ 1 & \text{при} \quad x > 3{,}53. \end{cases}$$

г)



7.4. а)

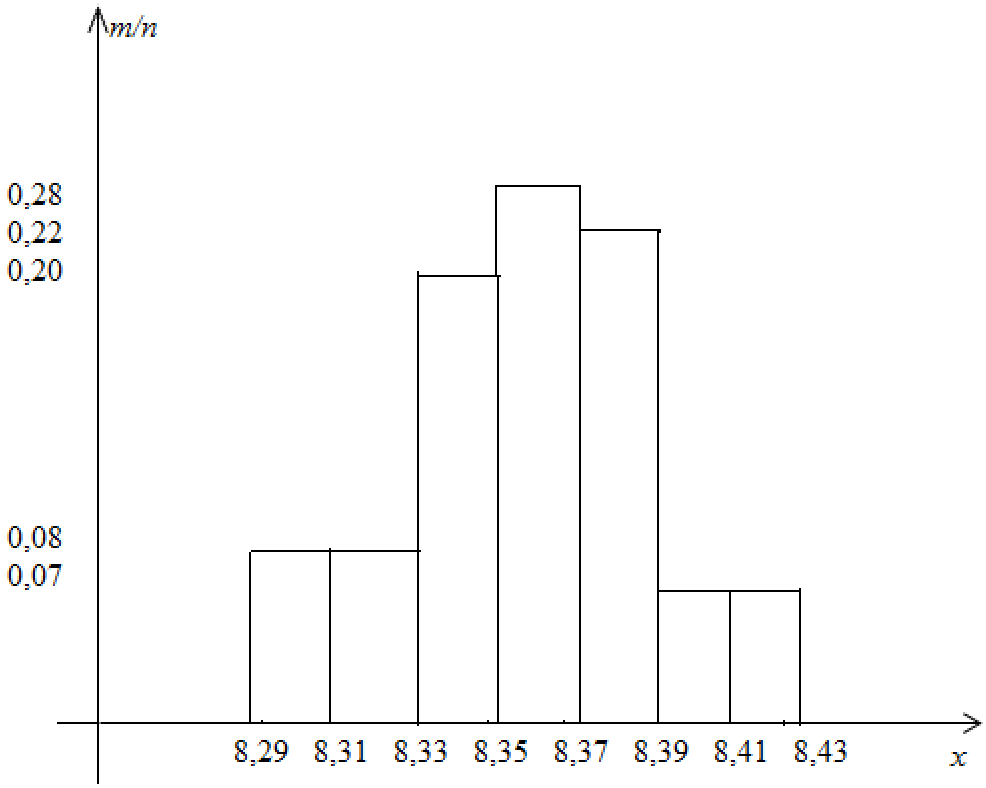
$[x_i; x_{i+1})$	m_i	$\frac{m_i}{n}$
[8,29; 8,31)	8	0,08
[8,31; 8,33)	8	0,08
[8,33; 8,35)	20	0,20
[8,35; 8,37)	28	0,28
[8,37; 8,39)	22	0,22
[8,39; 8,41)	7	0,07
[8,41; 8,43]	7	0,07

б)

$\bar{x}_i$	8,30	8,32	8,34	8,36	8,38	8,40	8,42
m_i	8	8	20	28	22	7	7
$\frac{m_i}{n}$	0,08	0,08	0,20	0,28	0,22	0,07	0,07

в) $$\tilde{F}(x) = \begin{cases} 0 & \text{при} \quad x \le 8{,}30, \\ 0{,}08 & \text{при } 8{,}30 < x \le 8{,}32, \\ 0{,}16 & \text{при } 8{,}32 < x \le 8{,}34, \\ 0{,}36 & \text{при } 8{,}34 < x \le 8{,}36, \\ 0{,}64 & \text{при } 8{,}36 < x \le 8{,}38, \\ 0{,}86 & \text{при } 8{,}38 < x \le 8{,}40, \\ 0{,}93 & \text{при } 8{,}40 < x \le 8{,}42, \\ 1 & \text{при} \quad x > 8{,}42. \end{cases}$$

г)



§8. Статистические оценки параметров

Точечные статистические оценки параметров распределения

Пусть собранный и обработанный статистический материал представлен в виде статистического ряда.

Определение 1. Точечной статистической оценкой параметра a распределения случайной величины называется приближенное значение a^* этого параметра, вычисленного по статистическим данным.

Замечание 1. Любая точечная статистическая оценка некоторого параметра, вычисляемая на основе статистического ряда, должна удовлетворять трём требованиям:

- при увеличении числа испытаний она должна сходиться по вероятности к оцениваемому параметру (свойство *состоятельности*);
- математическое ожидание статистической оценки (как случайной величины при изменении числа испытаний) равно оцениваемому параметру (свойство *несмещенности*);
- при заданном объёме выборки статистическая оценка имеет наименьшую дисперсию (свойство *эффективности*).

Определение 2. *Статистической оценкой математического ожидания* называется среднее арифметическое статистических значений изучаемой случайной величины:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i m_i ,$$

где $m_1+m_2+\ldots+m_k = n$.

Замечание 2. Эта оценка математического ожидания обладает всеми свойствами оценок: состоятельности, несмещенности, эффективности.

Определение 3. Смещенной оценкой дисперсии $D(x)$ называется выборочная дисперсия:

$$D_в = \frac{1}{n}\sum_{i=1}^{k} m_i (x_i - \bar{x})^2$$

Замечание 3. Эта оценка является смещенной, так как

$$M(D_в) = \frac{n-1}{n} D(x).$$

Определение 4. Несмещенной оценкой дисперсии $D(x)$ называется исправленная выборочная дисперсия:

$$s^2 = \frac{1}{n-1} D_в = \frac{1}{n-1} \cdot \sum_{i=1}^{k} m_i (x_i - \bar{x})^2$$

Замечание 4. При расчёте s^2 можно воспользоваться более удобной формулой:

$$s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{k} m_i \cdot x_i^2 - n \cdot \bar{x}^2\right)$$

Замечание 5. Выборочная дисперсия $D_в$ и исправленная выборочная дисперсия s^2 обладают свойством состоятельности. Оценка s^2 не обладает свойством эффективности, но обладает свойством несмещенности, поэтому ее чаще чем $D_в$ используют в качестве приближенного значения дисперсии $D(x)$.

Определение 5. *Оценкой среднего квадратического отклонения* $\sigma(x)$ называется квадратный корень из $D_в$ или s^2:

$$\sigma_в = \sqrt{D_в} \quad \text{или} \quad s = \sqrt{s^2}$$

Определение 6. Оценкой вероятности события A в n независимых испытаниях является относительная частота события A:

$$P^* = \frac{m}{n},$$

где m – число появления события A в n испытаниях.

Замечание 6. Эта оценка вероятности события A в n независимых испытаниях обладает свойствами несмещенности, состоятельности и эффективности.

Замечание 7. Если выборка состоит из вариант x_i громоздкого вида, то для упрощения расчета выборочных точечных оценок параметров следует перейти к *условным вариантам:*

$$u_i = \frac{x_i - c}{h},$$

где h – шаг между равноотстоящими вариантами; c – так называемый «ложный» нуль. Для них произвести расчет точечных оценок параметров:

$$\bar{u} = \frac{1}{n}\sum_{i=1}^{k} u_i m_i, \qquad D_в(u) = \frac{1}{n}\sum_{i=1}^{k} m_i (u_i - \bar{u})^2, \qquad s_u^2 = \frac{n}{n-1}\cdot D_в(u).$$

Затем вычислить искомые точечные оценки:

$$\bar{x} = \bar{u}\cdot h + c, \qquad D_в(x) = h^2\cdot D_в(u), \qquad s_x^2 = h^2\cdot s_u^2.$$

В качестве числа c обычно выбирают варианту x_{i0}, которая расположена в середине статистического ряда или имеет наибольшую частоту.

Интервальные оценки параметров нормального распределения

Для выборок небольшого объема вопрос точности оценок решается с помощью интервальных оценок.

При этом по вычисленной точечной оценке a^* параметра a при заданной вероятности γ, называемой *доверительной вероятностью*, а также по некоторому числу ε, зависящему от γ и a^*, строят интервал для истинного параметра a:

$$a^* - \varepsilon < a < a^* + \varepsilon ,$$

чтобы выполнялось равенство:

$$P\,(a^* - \varepsilon < a < a^* + \varepsilon) = \gamma .$$

Число ε называется *точностью* оценки a^*, границы интервала $a^* - \varepsilon$ и $a^* + \varepsilon$ называются *доверительными границами*, интервал ($a^* - \varepsilon$ *и* $a^* + \varepsilon$) – доверительным интервалом, вероятность γ – *доверительной вероятностью* или *надежностью* интервальной оценки.

Определение 7. *Интервальной оценкой математического ожидания* m нормального распределения *при известной* дисперсии σ^2 называется интервал

$$(\bar{x} - \varepsilon ;\ \bar{x} + \varepsilon) ,\ \varepsilon = z_\gamma \cdot \frac{\sigma}{\sqrt{n}} ,$$

удовлетворяющий равенству:

$$P\,(\bar{x} - \varepsilon < m < \bar{x} + \varepsilon) = \gamma ,$$

где γ – заданная доверительная вероятность; m – истинное математическое ожидание; $\bar{x}$ – точечная оценка математического ожидания; n – объем выборки; число z_γ находится из уравнения $\Phi\,(z_\gamma) = \frac{\gamma}{2}$ с помощью табл. П 2.2 функции Лапласа $\Phi\,(x)$, см. приложение 2.

Следовательно, интервальная оценка математического ожидания находится по формуле:

$$\bar{x} - z_\gamma \cdot \frac{\sigma}{\sqrt{n}} < m < \bar{x} + z_\gamma \cdot \frac{\sigma}{\sqrt{n}} .$$

Определение 8. *Интервальной оценкой математического ожидания* m нормального распределения *при неизвестной* дисперсии называется интервал

$$(\bar{x} - \varepsilon ;\ \bar{x} + \varepsilon) ,\ \varepsilon = t_\gamma \cdot \frac{s}{\sqrt{n}} ,$$

удовлетворяющий равенству:

$$P\,(\bar{x} - \varepsilon < m < \bar{x} + \varepsilon) = \gamma ,$$

где γ – заданная доверительная вероятность; m – истинное математическое ожидание; $\bar{x}$ – точечная оценка математического ожидания; s^2 – точечная оценка дисперсии; n – объем выборки; число t_γ вычисляется из уравнения

$$\int_0^{t\gamma} S\,(t;n)dt = \frac{\gamma}{2} ,$$

с помощью табл. П 2.3 распределения Стьюдента (см. приложение 2).

Следовательно, интервальная оценка математического ожидания с доверительной вероятностью γ вычисляется по формуле:

$$\bar{x} - t_\gamma \cdot \frac{s}{\sqrt{n}} < m < \bar{x} + t_\gamma \cdot \frac{s}{\sqrt{n}}.$$

Определение 9. *Интервальной оценкой среднего квадратического отклонения* σ нормального распределения называется интервал

$$(s - \varepsilon;\ s + \varepsilon),\ \varepsilon = q_\gamma \cdot s,$$

удовлетворяющий равенству:

$$P(s - \varepsilon < \sigma < s + \varepsilon) = \gamma,$$

где γ – заданная доверительная вероятность; s^2 – исправленная выборочная дисперсия; n – объем выборки; число q_γ определяется из табл. П 2.4 (см. приложение 2).

Следовательно, интервальная оценка среднего квадратического отклонения находится по формулам:

$$s(1 - q_\gamma) < \sigma < s(1 + q_\gamma),\ \text{если } q_\gamma < 1,$$

$$0 < \sigma < s(1 + q_\gamma),\ \text{если } q_\gamma > 1.$$

Задачи с решениями

Задача 8.1. По данным эксперимента построен статистический ряд:

x_i	11	12	13	14
m_i	20	5	15	10

Найти оценки математического ожидания, дисперсии и среднего квадратического отклонения случайной величины *X*.

Решение. 1) Число экспериментальных данных вычисляется по формуле:

$$n = \sum_{i=1}^{k} m_i = 20 + 5 + 15 + 10 = 50.$$

Значит, объем выборки $n = 50$.

2) Вычислим среднее арифметическое значение эксперимента:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{k} x_i \cdot m_i =$$

$$= \frac{1}{50} \sum_{i=1}^{4} x_i \cdot m_i = \frac{1}{50}(11\cdot20 + 12\cdot5 + 13\cdot15 + 14\cdot10) = \frac{615}{50} = 12{,}3.$$

Значит, найдена оценка математического ожидания $\bar{x}$ = 12,3.

3) Вычислим исправленную выборочную дисперсию:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{k}(x_i - \bar{x})^2 \cdot m_i =$$

$= \frac{1}{49}(20(11–12,3)^2 + 5(12–12,3)^2 + 15(13–12,3)^2 + 10(14–12,3)^2) = \frac{1}{49} \cdot (1,69\cdot 20+$
$+0,09\cdot 5 + 0,49\cdot 15 + 2,89\cdot 10) = \frac{1}{49}\cdot 70,5 = 1,44.$

Значит, найдена оценка дисперсии: $s^2 = 1,44$.

5) Вычислим оценку среднего квадратического отклонения:

$$s = \sqrt{1,44} = 1,2 .$$

Ответ: $\bar{x}$ = 12,3, s^2 = 1,44, s = 1,2 .

Задача 8.2. По данным эксперимента построен статистический ряд:

x_i	5,2	7,2	9,2	11,2	13,2	15,2
m_i	7	12	25	10	5	1

Найти оценки математического ожидания, дисперсии и среднего квадратического отклонения случайной величины X.

Решение. По формуле

$$u_i = \frac{x_i - 9,2}{2}$$

перейдем к условным вариантам:

u_i	– 2	– 1	0	1	2	3
m_i	7	12	25	10	5	1

Для них произведем расчет точечных оценок параметров:

$$\bar{u} = \frac{1}{60}\sum_{i=1}^{6} u_i m_i = -0,05 ,$$

$$D_в(u) = \frac{1}{60}\sum_{i=1}^{6} m_i (u_i + 0,05)^2 = 1,3143 ,$$

$$s_u^2 = \frac{60}{59} \cdot D_в(u) = 1,339.$$

Следовательно, вычисляем искомые точечные оценки:

$\bar{x} = 2\bar{u} + 9,2 = 9,1$,

$D_в(x) = 2^2 \cdot D_в(u) = 5{,}26$,

$s_x^2 = 2^2 \cdot s_u^2 = 5{,}36$.

Ответ: $\bar{x} = 9{,}1$, $D_в(x) = 5{,}26$, $s_x^2 = 5{,}36$.

Задача 8.3. По данным эксперимента построен интервальный статистический ряд:

$[a_i; a_{i+1})$	[0; 2)	[2; 4)	[4; 6)	[6; 8)	[8; 10]
m_i	3	4	10	5	3

Найти оценки математического ожидания, дисперсии и среднего квадратического отклонения.

Решение. 1) От интервального ряда перейдем к статистическому ряду, заменив интервалы их серединами $x_i = \frac{a_i + a_{i+1}}{2}$:

x_i	1	3	5	7	9
m_i	3	4	10	5	3

2) Объем выборки вычислим по формуле:

$$N = \sum_{i=1}^{5} m_i = 3 + 4 + 10 + 5 + 3 = 25.$$

3) Вычислим среднее арифметическое значений эксперимента:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i m_i = \frac{1}{25}(1\cdot 3 + 3\cdot 4 + 5\cdot 10 + 7\cdot 5 + 9\cdot 3) = \frac{1}{25}(3 + 12 + 50 + 35 + 27) = \\ = \frac{1}{25}\cdot 127 = 5{,}08.$$

3) Вычислим исправленную выборочную дисперсию:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{k}(x_i - \bar{x})^2 \cdot m_i =$$

$$= \frac{1}{24}((1-5{,}08)^2 \cdot 3 + (3-5{,}08)^2 \cdot 4 + (5-5{,}08)^2 \cdot 10 + (7-5{,}08)^2 \cdot 5 + (9-5{,}08)^2 \cdot 3) =$$

$$= \frac{1}{24}\cdot 131{,}87 \approx 5{,}49.$$

Можно было воспользоваться следующей формулой:

$s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{k} m_i \cdot x_i^{\,2} - n \cdot \bar{x}^{2}\right) = \frac{1}{24}\left(1 \cdot 3 + 3^2 \cdot 4 + 5^2 \cdot 10 + 7^2 \cdot 5 + 9^2 \cdot 3 - 25 \cdot (5{,}08)^2\right) =$

$= \frac{1}{24}(777 - 645{,}16) = \frac{1}{24} \cdot 131{,}84 \approx 5{,}49.$

5) Вычислим оценку среднего квадратического отклонения:

$$s = \sqrt{s^2} = \sqrt{5{,}49} \approx 2{,}34.$$

Ответ: $\bar{x} = 5{,}08$, $s^2 = 5{,}49$, $s = 2{,}34$.

Задача 8.4. Найти доверительный интервал с надежностью 0,95 для оценки математического ожидания $M(X)$ нормально распределенной случайной величины X, если известно среднее квадратическое отклонение $\sigma = 2$, оценка математического ожидания $\bar{x} = 10$, объем выборки $n = 25$.

Решение. Доверительный интервал для истинного математического ожидания с доверительной вероятностью $\gamma = 0{,}95$ при известной дисперсии σ находится по формуле:

$$\bar{x} - z_\gamma \frac{\sigma}{\sqrt{n}} < m < \bar{x} + z_\gamma \frac{\sigma}{\sqrt{n}},$$

где $m = M(X)$ – истинное математическое ожидание; $\bar{x}$ – оценка $M(X)$ по выборке; n – объем выборки; z_γ – находится по доверительной вероятности $\gamma = 0{,}95$ из равенства:

$$\Phi\,(z_\gamma) = 0{,}475.$$

Из табл. П 2.2 приложения 2 находим: $z_\gamma = 1{,}96$. Следовательно, найден доверительный интервал для $M(X)$:

$10 - 1{,}96\,\frac{2}{\sqrt{25}} < m < 10 + 1{,}96\,\frac{2}{\sqrt{25}}$, т.е. $9{,}216 < m < 10{,}784$.

Ответ: (9,216 ; 10,784).

Задача 8.5. По данным эксперимента построен статистический ряд:

x_i	– 3	– 2	– 1	0	1	2	3
m_i	1	3	4	5	3	2	2

Найти доверительный интервал для математического ожидания $M(X)$ с надежностью 0,95.

Решение. Воспользуемся формулой для доверительного интервала математического ожидания при неизвестной дисперсии:

$$\bar{x} - t_\gamma \frac{s}{\sqrt{n}} < m < \bar{x} + t_\gamma \frac{s}{\sqrt{n}}.$$

где n – объем выборки; $\bar{x}$ оценка $M(X)$; s – оценка среднего квадратического отклонения; t_γ – находится по доверительной вероятности γ = 0,95 из табл. П 2.3 приложения 2.

По числам γ = 0,95 и n = 20 находим: $t_\gamma = 2{,}093$.

Теперь вычисляем оценки для $M(X)$ и $D(X)$:

$$\bar{x} = \frac{1}{20}\sum_{i=1}^{7} x_i \cdot m_i = \frac{1}{20}(-3-6-4+0+3+4+6) = \frac{1}{20}\cdot 0 = 0.$$

$$s^2 = \frac{1}{19}\left(\sum_{i=1}^{7} m_i \cdot x_i^2 - 20\cdot \bar{x}^2\right) = = \frac{1}{19}(9\cdot 1 + 4\cdot 3 + 1\cdot 4 + 0 + 1\cdot 3 + 4\cdot 2 + + 9\cdot 2 - \frac{1}{20}\cdot 0^2)$$
$$= \frac{1}{19}(9 + 12 + 4 + 3 + 8 + 18) = \frac{1}{19}\cdot 54 \approx 2{,}84.$$

Следовательно, $s \approx 1{,}685$. Поэтому искомый доверительный интервал математического ожидания задается формулой:

$$0 - 2{,}093\,\frac{1{,}685}{\sqrt{20}} < m < 0 + 2{,}093\,\frac{1{,}685}{\sqrt{20}} \iff -0{,}76 < m < 0{,}76.$$

Ответ: (– 0,76; 0,76).

Задача 8.6. По данным десяти независимых измерений найдена оценка квадратического отклонения $s = 0{,}5$. Найти доверительный интервал точности измерительного прибора с надежностью 99 %.

Решение. Задача сводится к нахождению доверительного интервала для истинного квадратического отклонения, так как точность прибора характеризуется средним квадратическим отклонением случайных ошибок измерений.

Доверительный интервал для среднего квадратического отклонения находим по формуле:

$$s\,(1 - q_\gamma) < \sigma < s\,(1 + q_\gamma), \quad \text{если } q_\gamma < 1,$$

$$0 < \sigma < s\,(1 + q_\gamma), \text{ если } q_\gamma > 1,$$

где $s = 0{,}5$ – оценка среднего квадратического отклонения; q_γ – число, определяемое из табл. П 2.4 приложения 2 по заданной доверительной вероятности γ = 0,99 и заданному объему выборки n = 10.

Находим: $q_\gamma = 1{,}08 > 1$.

Тогда можно записать:

$0 < \sigma < 0{,}5\,(1 + 1{,}08) \iff 0 < \sigma < 1{,}04.$

Ответ: (0; 1,04).

Задачи

8.1. В результате испытаний получены следующие статистические значения случайной величины X:

– 0,9;	0,1;	– 2,9;	1,1;	5,1;	0,1;	– 6,9;	1,1;	– 3,9;	– 0,9;
5,1;	– 8,9;	–2,9;	0,1;	1,1;	6,1;	3,1;	0,1;	1,1;	– 2,9;
– 2,9;	0,1;	–3,9;	–0,9;	0,1;	3,1;	0,1;	1,1;	3,1;	0,1.

Найти оценки математического ожидания, дисперсии и среднего квадратического отклонения.

8.2. В результате испытаний получены следующие статистические значения случайной величины X:

1,36; 1,37; 1,35; 1,31; 1,34; 1,36; 1,38; 1,35; 1,39; 1,40;

1,33; 1,34; 1,36; 1,35; 1,37; 1,41; 1,36; 1,34; 1,39; 1,36;

1,35; 1,37; 1,38; 1,40; 1,37; 1,36; 1,35; 1,34; 1,37; 1,38.

Найти оценки математического ожидания, дисперсии и среднего квадратического отклонения.

8.3. В результате испытаний получены следующие статистические значения случайной величины X:

3,45; 3,47; 3,47; 3,43; 3,46; 3,44; 3,40; 3,45; 3,41; 3,42;

3,47; 3,49; 3,41; 3,48; 3,43; 3,40; 3,43; 3,47; 3,45; 3,44;

3,41; 3,40; 3,48; 3,46; 3,51; 3,39; 3,50; 3,50; 3,47; 3,38;

3,44; 3,40; 3,40; 3,44; 3,47; 3,53; 3,46; 3,46; 3,52; 3,47;

3,41; 3,44; 3,47; 3,45; 3,44; 3,45; 3,47; 3,42; 3,44; 3;50;

3,45; 3,50; 3,42; 3,48; 3,40; 3,45; 3,48; 3,48; 3,46; 3,47;

3,44; 3,44; 3,47; 3,43; 3,44; 3,47; 3,44; 3,45; 3,44; 3,46;

3,46; 3,44; 3,44; 3,44; 3,44; 3,46; 3,44; 3,42; 3,50; 3,46;

3,48; 3,43; 3,40; 3,46; 3,46; 3,47; 3,45; 3,48; 3,42; 3,46;

3,48; 3,38; 3,45; 3,43; 3,52; 3,43; 3,50; 3,51; 3,41; 3,52.

Найти оценки математического ожидания, дисперсии и среднего квадратического отклонения.

Указание: вместо интервального статистического ряда построить статистический, выбрав в качестве значений случайной величины середины интервалов.

8.4. При измерении диаметров ста подшипниковых шариков, выбранных из большой партии шариков для определения стандартности, получены следующие результаты:

8,31; 8,42; 8,37; 8,40; 8,40; 8,30; 8,30; 8,42; 8,32; 8,29;

8,33; 8,36; 8,34; 8,37; 8,32; 8,36; 8,38; 8,38; 8,33; 8,36;

8,40; 8,36; 8,32; 8,36; 8,36; 8,30; 8,30; 8,33; 8,35; 8,37;

8,37; 8,30; 8,41; 8,34; 8,33; 8,37; 8,34; 8,38; 8,29; 8,34;

8,31; 8,36; 8,37; 8,30; 8,41; 8,34; 8,34; 8,37; 8,35; 8,40;

8,34; 8,36; 8,37; 8,37; 8,41; 8,35; 8,38; 8,33; 8,36; 8,36;

8,36; 8,37; 8,36; 8,40; 8,37; 8,34; 8,37; 8,32; 8,35; 8,36;

8,37; 8,41; 8,36; 8,36; 8,36; 8,40; 8,34; 8,40; 8,34; 8,33;

8,35; 8,37; 8,34; 8,36; 8,37; 8,37; 8,35; 8,36; 8,34; 8,42;

8,36; 8,33; 8,34; 8,35; 8,36; 8,32; 8,38; 8,32; 8,36; 8,37.

Найти оценки математического ожидания, дисперсии и среднего квадратического отклонения.

Указание: вместо интервального статистического ряда построить статистический, выбрав в качестве значений случайной величины середины интервалов.

8.5. В результате независимых испытаний получены данные:

2,38; 2,41; 2,33; 2,45; 2,42; 2,46; 2,43; 2,41;

2,36; 2,31; 2,41; 2,48; 2,44; 2,35; 2,38; 2,49;

2,37; 2,42; 2,39; 2,40; 2,34; 2,42; 2,36; 2,45;

2,42; 2,41; 2,39; 2,38; 2,43; 2,46.

1) Найти с доверительной вероятностью $\gamma = 0{,}99$ доверительный интервал для истинного математического ожидания

а) с известной дисперсией $\sigma^2 = 0{,}0016$,

б) с неизвестной дисперсией.

2) Найти с доверительной вероятностью $\gamma = 0{,}95$ доверительный интервал для истинного среднего квадратического отклонения.

8.6. В результате независимых испытаний получены данные:

1,38; 1,41; 1,33; 1,45; 1,42; 1,46; 1,43; 1,41;

1,36; 1,31; 1,41; 1,48; 1,44; 1,36; 1,38; 1,40;

1,37; 1,42; 1,39; 1,40; 1,34; 1,42; 1,36; 1,46;

1,41; 1,39; 1,38; 1,43; 1,46; 1,42.

1) Найти с доверительной вероятностью $\gamma = 0{,}95$ доверительный интервал для истинного математического ожидания

а) с известной дисперсией $\sigma^2 = 0{,}0016$,

б) с неизвестной дисперсией.

2) Найти с доверительной вероятностью $\gamma = 0{,}99$ доверительный интервал для истинного среднего квадратического отклонения.

Ответы

8.1. $\bar{x} = -0{,}2$, $s^2 = 10{,}907$, $s = 3{,}302$.

8.2. $\bar{x} = 1{,}363$, $s^2 = 0{,}000498$, $s = 0{,}0223$.

8.3. $\bar{x} = 3{,}46$, $D_в = 0{,}0011$, $\sigma_в = 0{,}033$.

Указание: постройте таблицу вида:

№	α_i - α_{i+1}	m_i	$x_i = \frac{a_i + a_{i+1}}{2}$	$x_i \cdot m_i$	$x_i^2 \cdot m_i$
1. 2. 3.					
$\sum$		n		$\sum x_i \cdot m_i$	$\sum x_i^2 \cdot m_i$

8.4. $\bar{x} = 8{,}36$, $D_в = 0{,}00091$, $\sigma_в = 0{,}0302$.

8.5. 1) а) $2{,}381 < m < 2{,}419$;
б) $2{,}377 < m < 2{,}423$;
2) $0{,}032 < \sigma < 0{,}058$.

8.6. 1) а) $1{,}389 < m < 1{,}417$;
б) $1{,}371 < m < 1{,}434$;
2) $0{,}023 < \sigma < 0{,}059$.

§9. Проверка статистических гипотез

Статистические гипотезы

Определение 1. *Статистической гипотезой* называется предположение о виде распределения, о параметрах известных распределений.

Определение 2. Выдвинутая гипотеза называется *нулевой* (или *основной*) и обозначается H_0.

Определение 3. Гипотеза, которая противоречит нулевой, называется *конкурирующей гипотезой* (или *альтернативной гипотезой*) и обозначается H_1.

Определение 4. Гипотеза называется *простой*, если она содержит только одно предположение.

Определение 5. Гипотеза называется *сложной*, если она состоит из конечного или бесконечного числа предположений.

Проверка гипотез

Определение 6. Сопоставление выдвинутой гипотезы с выборочными данными называется *проверкой гипотезы*.

Определение 7. Правосторонней *критической областью* для проверки нулевой гипотезы с уровнем значимости α называется совокупность значений критерия проверки Z, для которых выполняется равенство:

$$P\,(Z > Z_{крит}) = \alpha\ .$$

При этом число $Z_{крит}$ называется *границей* критической области.

Замечание 1. Различают право –, лево –, двустороннюю критические области.

Правосторонняя критическая область определяется неравенством:

$$Z > Z_{крит}\ .$$

Левосторонняя критическая область определяется неравенством:

$$Z < -Z_{крит}\ .$$

Двусторонняя критическая область определяется неравенствами:

$$Z < Z_{1\,крит} \quad \text{и} \quad Z > Z_{2\,крит}\ .$$

Схема проверки гипотезы

а) Формирование нулевой гипотезы H_0, которая выдвигается на основе начального анализа выборочных данных, и конкурирующей гипотезы H_1;

б) Выбор некоторой вероятности α, называемой *уровнем значимости* нулевой гипотезы H_0;

в) Подбор по выборочным данным случайной величины Z, распределение которой называется *критерием для проверки нулевой гипотезы*;

г) Определение границы $Z_{крит}$ *критической области* для проверки нулевой гипотезы с уровнем значимости α;

д) Вычисление по данным выборки числа $Z_{наблюдаемое}$ (или $Z_{набл.}$).

Если оно попадает в критическую область нулевой гипотезы, т.е.

$$Z_{набл} > Z_{крит} ,$$

то гипотеза H_0 отвергается и принимается гипотеза H_1. Если же

$$Z_{набл} < Z_{крит} ,$$

то нет основания отвергать нулевую гипотезу.

Замечание 2. Если гипотеза принята, то не стоит думать, что она доказана. На практике для большей уверенности принятия гипотезы ее проверяют, повторяя эксперимент, увеличив объем выборки.

Определение 8. *Ошибкой первого рода* называется решение отвергнуть нулевую гипотезу H_0 и принять конкурирующую гипотезу H_1, если на самом деле гипотеза H_0 верна.

Замечание 3. Вероятность ошибки первого рода равна уровню значимости α.

Определение 9. *Ошибкой второго рода* называется решение принять нулевую гипотезу H_0 , то есть отвергнуть конкурирующую гипотезу H_1 , если на самом деле гипотеза H_1 верна.

Сравнение двух дисперсий нормальных генеральных совокупностей

Пусть по двум независимым выборкам, объемы которых равны n_1 и n_2 соответственно, полученным из нормально распределенных генеральных совокупностей X и Y , найдены исправленные выборочные дисперсии s_X^2 и s_Y^2. Требуется сравнить дисперсии генеральных совокупностей.

Схема сравнения $D(X)$ и $D(Y)$

1) Выдвинуть нулевую гипотезу: H_0: $D(X)=D(Y)$. Тогда конкурирующей гипотезой будет $H_1 : D(X) > D(Y)$;
2) Задать число α – уровень значимости нулевой гипотезы;
3) Найти из табл. П 2.7 распределения Фишера (см. приложение 2) значение $F_{крит}$ по заданному α и числам степеней свободы

$$k_1 = n_1 - 1 , k_2 = n_2 - 1,$$

где k_1 – число степеней свободы большей дисперсии, k_2 – число степеней свободы меньшей дисперсии.

4) Найти число

$$F_{\text{набл}} = \frac{S^2_{\text{наиб}}}{S^2_{\text{наим}}},$$

равное отношению большей из двух исправленных выборочных дисперсий s_X^2 и s_Y^2 к меньшей;

5) Сравнить числа $F_{\text{крит}}$ и $F_{\text{набл}}$:

- если $F_{\text{набл}} > F_{\text{крит}}$, то отвергнуть гипотезу H_0 и принять гипотезу H_1;
- если $F_{\text{набл}} < F_{\text{крит}}$, то нет основания отвергать гипотезу H_0.

Замечание 4. Если для нулевой гипотезы H_0 : $D(X)=D(Y)$ в качестве конкурирующей гипотезы выбрана H_1 $D(X)\neq D(Y)$, то строят двустороннюю критическую область. Для этого по табл. П 2.7 (см. приложение 2) вычисляют правую границу $F_{2\,\text{крит}}$ критической области по уровню значимости $\frac{\alpha}{2}$ и числам степеней свободы $k_1 = n_1 - 1$, $k_2 = n_2 - 1$. Тогда, если $F_{\text{набл}} > F_{2\,\text{крит}}$, то гипотеза H_0 отвергается и принимается гипотеза H_1; если $F_{\text{набл}} < F_{2\,\text{крит}}$, то нет основания отвергать гипотезу H_0 .

Сравнение нескольких дисперсий нормальных генеральных совокупностей

Пусть генеральные совокупности $X_1, \ldots, X_l$ распределены нормально. Из этих совокупностей извлечены независимые выборки объемов n_1 , $n_2, \ldots, n_l$ соответственно. По выборкам найдены исправленные выборочные дисперсии $s_1^2, \ldots, s_l^2$. Требуется сравнить дисперсии генеральных совокупностей.

Схема сравнения $D(X_1), \ldots, D(X_l)$

1) Выдвинуть нулевую гипотезу: $D(X_1)=D(X_2)=\ldots= D(X_l)$;
2) Задать число α – уровень значимости нулевой гипотезы;
3) Найти из табл. П 2.5 распределения χ^2 (см. приложение 2) значение $\chi^2_{\text{крит}}$ по заданному α и числу степеней свободы $l - 1$;
4) Найти число $B_{\text{набл}} = \dfrac{V}{C}$,

где $V = 2{,}303 \cdot \left(k \cdot \lg \bar{s}^2 - \sum\limits_{i=1}^{l} k_i \cdot \lg s_i^2 \right)$, $C = 1 + \dfrac{1}{3(l-1)} \cdot \left(\sum\limits_{i=1}^{l} \dfrac{1}{k_i} - \dfrac{1}{k} \right)$,

$$k_i = n_i - 1, \qquad k = \sum_{i=1}^{l} k_i, \qquad \bar{s}^2 = \frac{1}{k} \sum_{i=1}^{l} k_i s_i^2 ;$$

5) Сравнить числа $\chi^2_{крит}$ и $B_{набл}$:
- если $B_{набл} > \chi^2_{крит}$, то отвергнуть гипотезу H_0,
- если $B_{набл} < \chi^2_{крит}$ то нет основания отвергать гипотезу H_0.

Замечание 5. В случае принятия гипотезы H_0 в качестве оценки дисперсии генеральной совокупности за дисперсию этой генеральной совокупности принимают число $\bar{s}^2$.

Сравнение двух средних нормальных генеральных совокупностей

Пусть даны две независимые выборки объемов n_1 и n_2 соответственно из нормально распределенных генеральных совокупностей X и Y. По выборкам найдены оценки математических ожиданий $\bar{x}$, $\bar{y}$ и исправленные выборочные дисперсии s_X^2, s_Y^2. Требуется сравнить $M(X)$ и $M(Y)$ генеральных совокупностей.

Схема сравнения M(X) и M(Y)

1) Выдвинуть нулевую гипотезу: $H_0 : M(X) = M(Y)$.
   В качестве конкурирующей гипотезы рассмотреть
$$H_1 : M(X) \neq M(Y);$$
2) Задать число α – уровень значимости нулевой гипотезы;
3) Найти по табл. П 2.6 распределения Стьюдента (см. приложение 2) значение $T_{крит}$ по заданному α и числу $k = n_1 + n_2 - 2$;
4) Найти число $T_{набл} = \dfrac{\bar{x}-\bar{y}}{\sqrt{(n_1-1)\cdot s_X^2 + (n_2-1)\cdot s_Y^2}} \cdot \sqrt{\dfrac{n_1 n_2 (n_1+n_2-2)}{n_1+n_2}}$;
5) Сравнить числа $T_{крит}$ и $T_{набл}$:
- если $|T_{набл}| > T_{крит}$, то отвергнуть гипотезу H_0,
- если $|T_{набл}| < T_{крит}$, то нет основания отвергать гипотезу H_0.

Замечание 6. Если необходимо проверить гипотезу H_0: $M(X) = M(Y)$ о равенстве средних двух нормальных генеральных совокупностей X и Y при условии известных дисперсий σ_X^2 и σ_Y^2, то в описанной выше схеме вместо $T_{крит}$ используют число $N_{крит}$, определяемое с помощью табл. П 2.2 (см. приложение 2) по заданному α из равенства:

$$\Phi(N_{крит}) = \frac{1-\alpha}{2}$$

Вместо $T_{набл}$ по данным выборок вычисляют число

$$N_{набл} = \frac{\bar{x}-\bar{y}}{\sqrt{\dfrac{\sigma_X^2}{n_1} + \dfrac{\sigma_Y^2}{n_2}}}.$$

Если $|N_{набл}| < N_{крит}$, то нет основания отвергать гипотезу H_0 .

Если $|N_{набл}| > N_{крит}$, то гипотезу H_0 отвергают.

Задачи с решениями

Задача 9.1. По двум независимым выборкам объемов $n_1 = 10$ и $n_2 = 15$, извлеченным из нормальных генеральных совокупностей X и Y, найдены оценки дисперсий: $s_X^2 = 8{,}42$, $s_Y^2 = 4{,}23$. При уровне значимости $\alpha = 0{,}05$ проверить гипотезу H_0: $D(X)=D(Y)$ при конкурирующей гипотезе H_1: $D(X)>D(Y)$.

Решение. 1) По данным выборки вычисляем

$$F_{набл} = \frac{S^2_{наиб}}{S^2_{наим}} = \frac{8{,}42}{4{,}23} = 1{,}99 .$$

2) По табл. П 2.7 (см. приложение 2), учитывая значения

$$\alpha = 0{,}05,\ k_1 = n_1 - 1 = 9 ,\ k_2 = n_2 - 1 = 14.$$

находим число: $F_{крит} = 2{,}65$.

3) Сравниваем: так как $1{,}99 < 2{,}65$, т.е. $F_{набл} < F_{крит}$, то нет основания отвергать гипотезу H_0.

Ответ: гипотеза $H_0 : D(X)=D(Y)$ принимается.

Задача 9.2. По двум независимым выборкам объемов $n_1=10$ и $n_2=15$, извлеченным из нормальных генеральных совокупностей X и Y, найдены оценки дисперсий: $S_X^2 = 8{,}42$, $S_Y^2 = 4{,}23$. При уровне значимости $\alpha = 0{,}1$ проверить гипотезу H_0: $D(X)=D(Y)$ при конкурирующей гипотезе H_1: $D(X) \neq D(Y)$.

Решение. 1) По данным выборки вычисляем

$$F_{набл} = \frac{S^2_{наиб}}{S^2_{наим}} = \frac{8{,}42}{4{,}23} = 1{,}99 .$$

2) По табл. П 2.7 (см. приложение 1), учитывая значения

$$\alpha = 0{,}1 \text{ и } \frac{\alpha}{2} = 0{,}05 ,\ k_1 = n_1 - 1 = 9 ,\ k_2 = n_2 - 1 = 14.$$

находим число: $F_{крит} = 2{,}65$.

3) Сравниваем: так как $1{,}99 < 2{,}65$, т.е. $F_{набл} < \mathrm{F}_{крит}$, то нет основания отвергать гипотезу H_0.

Ответ: гипотеза $H_0 : D(X)=D(Y)$ принимается.

Задача 9.3. По трем независимым выборкам объемов $n_1 = 10$ и $n_2 = 15$ и $n_3 = 20$, извлеченным из нормальных генеральных совокупностей X, Y и Z найдены оценки дисперсий: $S_1^2 = 3{,}62$, $S_2^2 = 4{,}23$, $S_3^2 = 7{,}45$. При уровне значимости $\alpha = 0{,}05$ проверить гипотезу $H_0 : D(X)=D(Y)=D(Z)$.

Решение. 1) По данным выборок вычисляем:

$$k_1 = n_1 - 1 = 9,\ k_2 = n_2 - 1 = 14,\ k_3 = n_3 - 1 = 19.$$

$$k = \sum_{i=1}^{l} k_i = 9 + 14 + 19 = 42,$$

$$\bar{s}^2 = \frac{1}{k}\sum_{i=1}^{l} k_i s_i^2 = \frac{1}{42}\cdot(9\cdot 3{,}62 + 14\cdot 4{,}23 + 19\cdot 7{,}45) = 5{,}56$$

$$V = 2{,}303\cdot\left(k\cdot \lg \bar{s}^2 - \sum_{i=1}^{l} k_i \cdot \lg s_i^2\right) =$$

$$= 2{,}303\cdot(42\cdot \lg 5{,}56 - 9\cdot \lg 3{,}62 - 14\cdot \lg 4{,}23 - 19\cdot \lg 7{,}45) = 2{,}13.$$

$$C = 1 + \frac{1}{3(3-1)}\cdot\left(\sum_{i=1}^{l}\frac{1}{k_i} - \frac{1}{k}\right) = 1 + \frac{1}{6}\cdot\left(\frac{1}{9} + \frac{1}{14} + \frac{1}{19} - \frac{1}{42}\right) = 1{,}035,$$

$$B_{\text{набл}} = \frac{2{,}13}{1{,}035} = 2{,}06.$$

2) По табл. П 2.5 (см. приложение 2), учитывая значения

$$\alpha = 0{,}05, \qquad k = 3 - 1 = 2,$$

находим число

$$\chi^2_{\text{крит}} = 6{,}0$$

3) Сравниваем: так как $2{,}06 < 6{,}0$, т.е. $B_{\text{набл}} < \chi^2_{\text{крит}}$, следовательно нет основания отвергать нулевую гипотезу.

Ответ: гипотезу $H_0 : D(X) = D(Y) = D(Z)$ принимают.

Задача 9.4. По двум независимым выборкам объемов $n_1=10$ и $n_2=16$, извлеченным из нормальных генеральных совокупностей X и Y, найдены оценки математических ожиданий $\bar{x}=2{,}5$, $\bar{y}=3{,}1$ и исправленные выборочные дисперсии $s_X^2 = 0{,}62$, $s_Y^2 = 0{,}43$. Проверить нулевую гипотезу: H_0 :

$M(X) = M(Y)$ при конкурирующей гипотезе $H_1 : M(X) \neq M(Y)$ и уровне значимости $\alpha = 0{,}01$.

Решение. 1) Так как $s_X^2 \neq s_Y^2$, то предварительно проверим гипотезу H_0: $D(X)=D(Y)$ о равенстве генеральных дисперсий при конкурирующей гипотезе H_1: $D(X)>D(Y)$. Для этого поступаем по аналогии с решением 1 задачи.

а) По данным выборки вычисляем

$$F_{\text{набл}} = \frac{S^2_{\text{наиб}}}{S^2_{\text{наим}}} = \frac{0{,}62}{0{,}43} = 1{,}44;$$

б) По табл. П 2.7 (см. приложение 2), учитывая значения

$$\alpha = 0{,}01,\ k_1 = n_1 - 1 = 9,\ k_2 = n_2 - 1 = 15.$$

находим число:

$$F_{\text{крит}} = 3{,}89.$$

в) Сравниваем: так как $1{,}44 < 3{,}89$, т.е. $F_{\text{набл}} < F_{\text{крит}}$, то гипотеза о равенстве генеральных дисперсий принимается, то есть различие между $s_X^2 = 0{,}62$ и $s_Y^2 = 0{,}43$ считаем незначительным.

2) Проверим гипотезу $H_0 : M(X) = M(Y)$ о равенстве средних при конкурирующей гипотезе

$$H_1 : M(X) \neq M(Y).$$

а) Найдем по табл. П 2.6 (см. приложение 2) значение $T_{\text{крит}}$ по заданному $\alpha = 0{,}01$ и числу $k = 10 + 16 - 2 = 24$:

$$T_{\text{крит}} = 2{,}8.$$

б) Найдем число $T_{\text{набл}}$:

$$T_{\text{набл}} = \frac{\bar{x} - \bar{y}}{\sqrt{(n_1 - 1)\cdot s_X^2 + (n_2 - 1)\cdot s_Y^2}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} =$$

$$= \frac{2{,}5 - 3{,}1}{\sqrt{9 \cdot 0{,}62 + 15 \cdot 0{,}43}} \cdot \sqrt{\frac{10 \cdot 16 (10 + 16 - 2)}{10 + 16}} =$$

$$= -\frac{0{,}6}{\sqrt{12{,}03}} \cdot 12{,}153 = -\frac{0{,}6}{3{,}47} \cdot 12{,}153 = -2{,}101.$$

в) Сравнить числа $T_{крит}$ и $|T_{набл}|$: так как 2,101 < 2,8 то $|T_{набл}| < T_{крит}$ и гипотеза $H_0 : M(X) = M(Y)$ о равенстве средних принимается.

Ответ: гипотеза $H_0 : M(X) = M(Y)$ принимается.

Задача 9.5. По двум независимым выборкам объемов n_1=10 и n_2=16, извлеченным из нормальных генеральных совокупностей X и Y, с дисперсиями σ_X^2=9, σ_Y^2=12, вычислены оценки математических ожиданий $\overline{x}$=12,7, $\overline{y}$=10,2. При уровне значимости α = 0,05 проверить гипотезу $H_0 : M(X) = M(Y)$ и конкурирующей гипотезе $H_1 : M(X) \neq M(Y)$.

Решение. Воспользуемся замечанием 6. 1) Вычислим:

$$N_{набл} = \frac{\overline{x}-\overline{y}}{\sqrt{\frac{\sigma_X^2}{n_1}+\frac{\sigma_Y^2}{n_2}}} = \frac{12{,}7-10{,}2}{\sqrt{\frac{9}{18}+\frac{12}{15}}} = \frac{2{,}5}{\sqrt{0{,}5+0{,}8}}\frac{2{,}5}{\sqrt{1{,}3}} = 2{,}19.$$

2) Находим $N_{крит}$ из уравнения

$$\Phi\left(N_{крит}\right) = \frac{1-\alpha}{2}, \text{ то есть } \Phi\left(N_{крит}\right) = 0{,}475$$

используя табл. П 2.2 (см. приложение 2).

Следовательно,

$$N_{крит} = 1{,}96.$$

3) Сравниваем: так как 2,19 > 1,96, т.е. $N_{набл} > N_{крит}$, то гипотезу H_0 отвергают. Значит, различие генеральных математических ожиданий значительное.

Ответ: гипотеза $H_0 : M(X) = M(Y)$ отвергается, т.е. $M(X) \neq M(Y)$.

Задачи

9.1. По двум независимым выборкам объемов n_1=10 и n_2=13, извлеченным из нормальных генеральных совокупностей X и Y, найдены исправленные выборочные дисперсии s_X^2=0,52 и s_Y^2=0,28. При уровне значимости α = 0,01 проверить гипотезу H_0: $D(X) = D(Y)$ при конкурирующей гипотезе H_1: $D(X) > D(Y)$.

9.2. По двум независимым выборкам объемов n_1=15 и n_2=10, извлеченным из нормальных генеральных совокупностей X и Y, найдены исправленные выборочные дисперсии s_X^2=1,92 и s_Y^2=3,21. При уровне значимости α = 0,1 проверить гипотезу H_0 : $D(X)=D(Y)$ при конкурирующей гипотезе H_1: $D(X) \neq D(Y)$.

9.3. По двум независимым выборкам объемов n_1=15 и n_2=10, извлеченным из нормальных генеральных совокупностей X и Y с дисперсиями $D(X)$=25, $D(Y)$=32, найдены выборочные средние $\bar{x}$=53, $\bar{y}$= 61. При уровне значимости α = 0,1 проверить нулевую гипотезу H_0 : $M(X)=M(Y)$ при конкурирующей гипотезе H_1 : $M(X) \neq M(Y)$.

9.4. По двум независимым выборкам объемов n_1=15 и n_2=12, извлеченным из нормальных генеральных совокупностей X и Y найдены выборочные средние $\bar{x}$=11,2, $\bar{y}$=15,7 и исправленные выборочные дисперсии s_X^2=0,58 , s_Y^2=0, 83. При уровне значимости α = 0,05 проверить нулевую гипотезу H_0 : $M(X)=M(Y)$ при конкурирующей гипотезе H_1: $M(X) \neq M(Y)$.

9.5. По трем независимым выборкам объемов n_1=10, n_2=12 и n_3=17, извлеченным из нормальных генеральных совокупностей X, Y и Z, найдены выборочные дисперсии $D_в(X)$=2,3, $D_в(Y)$=2,7 , $D_в(Z)$=4,5 При уровне значимости α = 0,05 проверить гипотезу H_0 : $D(X) = D(Y) = D(Z)$.

9.6. По двум независимым выборкам объемов n_1=12 и n_2=18, извлеченным из нормальных генеральных совокупностей X и Y найдены исправленные выборочные дисперсии s_X^2=25,31 , s_Y^2=10,23. При уровне значимости α = 0,05 проверить гипотезу H_0: $D(X) = D(Y)$ при конкурирующей гипотезе H_1 : $D(X) > D(Y)$.

9.7. По двум независимым выборкам объемов n_1=12 и n_2=10, извлеченным из нормальных генеральных совокупностей X и Y найдены исправленные выборочные дисперсии $D_в(X)$=12,3, $D_в(Y)$=18,5. При уровне значимости α = 0,02 проверить гипотезу H_0: $D(X) = D(Y)$ при конкурирующей гипотезе H_1 : $D(X) \neq D(Y)$.

9.8. По двум независимым выборкам объемов n_1=40 и n_2=30, извлеченным из нормальных генеральных совокупностей X и Y с дисперсиями $D(X)$ = 80, $D(Y)$ = 70, найдены выборочные средние $\bar{x}$ = 120, $\bar{y}$ = 115. При уровне значимости α = 0,01 проверить нулевую гипотезу H_0 : $M(X) = M(Y)$ при конкурирующей гипотезе H_1: $M(X) \neq M(Y)$.

9.9. Из двух партий изделий, изготовленных на двух одинаково настроенных станках, извлечены две выборки:

x_i	2,2	2,6	2,8	3,1
m_i	2	3	5	2

y_i	2,5	2,7	2,8	3,0
m_i	2	4	6	3

Проверить нулевую гипотезу H_0 : $M(X)=M(Y)$ при конкурирующей гипотезе H_1 : $M(X)\neq M(Y)$ при уровне значимости $\alpha = 0,1$.

9.10. По четырем независимым выборкам объемов $n_1=12$ и $n_2=15$, $n_3=18$ и $n_4=20$ извлеченным из нормальных генеральных совокупностей X, Y, Z и U, найдены исправленные выборочные дисперсии $S_1^2=0,27$, $S_2^2=0,52$, $S_3^2=0,85$ и $S_4^2=0,99$. При уровне значимости $\alpha = 0,01$ проверить гипотезу H_0: $D(X) = D(Y) = D(Z) = D(U)$.

Ответы

9.1. H_0 принимается: $F_{\text{набл}} = 1,86$, $F_{\text{крит}} = 4,39$.

9.2. H_0 принимается: $F_{\text{набл}} = 1,67$, $F_{\text{крит}} = 2,65$.

9.3. H_0 отвергается: $N_{\text{набл}} = -5,36$, $N_{\text{крит}} = 1,64$.

9.4. H_0 отвергается: $T_{\text{набл}} = -13,99$, $T_{\text{крит}} = 2,06$.

Указание. Предварительно проверить равенство дисперсий при заданном уровне значимости

9.5. H_0 принимается: $B_{\text{набл}} = 4,064$, $\chi^2_{\text{крит}} = 6,0$.

9.6. H_0 отвергается: $F_{\text{набл}} = 2,47$, $F_{\text{крит}} = 2,41$.

9.7. H_0 принимается: $F_{\text{набл}} = 1,53$, $F_{\text{крит}} = 4,63$.

9.8. H_0 принимается: $N_{\text{набл}} = 2,40$, $N_{\text{крит}} = 2,58$.

9.9. H_0 принимается: $T_{\text{набл}} = 0,73$, $T_{\text{крит}} = 1,71$.

Указание: Предварительно проверить равенство дисперсий при заданном уровне значимости

9.10. H_0 принимается: $B_{\text{набл.}} =2,918$, $\chi^2=11,3$.

§10. Критерий согласия Пирсона

Проверка гипотезы о нормальном распределении

Критерием согласия называется критерий проверки гипотезы о предполагаемом законе распределения.

Критерий согласия Пирсона (критерий проверки гипотезы о *нормальном* распределении генеральной совокупности):

1) По выборке объема n построить статистический ряд:

x_i	x_1	x_2	…	x_l
m_i	m_1	m_2	…	m_l

2) Вычислить по таблице оценку математического ожидания $\bar{x}$ и выборочное среднее квадратическое отклонение $\sigma_в$.

3) В предположении нормального распределения генеральной совокупности вычислить теоретические частоты $m_{1\text{ теор}}, \ldots, m_{l\text{ теор}}$ по формуле:

$$m_{1\text{ теор}} = n \cdot p_i ,$$

где $p_i = \Phi\left(\frac{x_{i+1} - \bar{x}}{\sigma_в}\right) - \Phi\left(\frac{x_i - \bar{x}}{\sigma_в}\right)$; $\Phi(x)$ – интегральная функция распределения Лапласа – табл. П 2.2(см. приложение 2).

4) Вычислить число $\chi^2_{\text{набл}}$ по формуле:

$$\chi^2_{\text{набл}} = \sum_{i=1}^{l} \frac{(m_i - m_{i\text{ теор}})^2}{m_{i\text{ теор}}} \quad \text{или} \quad \chi^2_{\text{набл}} = \sum_{i=1}^{l} \frac{m_i^2}{m_{i\text{ теор}}} - n .$$

5) По табл. П 2.5 (приложение 2) найти число $\chi^2_{\text{крит}}$, учитывая заданный уровень значимости α и число степеней свободы $k = l - 3$.

6) Сравнить числа $\chi^2_{\text{набл}}$ и $\chi^2_{\text{крит}}$:

- если $\chi^2_{\text{набл}} < \chi^2_{\text{крит}}$, то нет основания отвергать гипотезу о нормальном распределении генеральной совокупности.
- если $\chi^2_{\text{набл}} > \chi^2_{\text{крит}}$, то гипотезу о нормальном распределении генеральной совокупности следует отвергнуть.

Замечание 1. Объем выборки n должен быть достаточно велик (больше 100). Число l обычно выбирают в диапазоне от 7 до 15. Поэтому при составлении интервального статистического ряда не используют интервалы, содержащие малое число значений объединяя их в один и суммируя соответствующее число значений.

Замечание 2. В случае $\chi^2_{\text{набл}} < \chi^2_{\text{крит}}$, для избежания ошибки первого рода следует повторить опыт, увеличив число n.

Замечание 3. При использовании критерия Пирсона с целью систематизации записи рекомендуется записывать все промежуточные вычисления в виде следующей таблицы:

№	x_i	m_i	m_i^2	$m_{i\,\text{теор}}$	$\dfrac{(m_i - m_{i\,\text{теор}})^2}{m_{i\,\text{теор}}}$
1					
2					
…					
l					
$\sum$		n		n	$\chi^2_{\text{набл}}$

Задачи с решениями

Задача 10.1. Отделом технического контроля качества продукции произведен выбор 200 деталей для измерения отклонений их действительного диаметра от планируемого. Данные измерений приведены в табл. 10.1.

Таблица 10.1

$[a_i\,;\,a_{i+1})$	m_i
[–20 ; –15)	7
[–15 ; –10)	11
[–10 ; –5)	15
[–5 ; 0)	24
[0 ; 5)	49
[5 ; 10)	41
[10 ; 15)	26
[15 ; 20)	17
[20 ; 25)	7
[25 ; 30)	3

Оценить с помощью критерия Пирсона гипотезу о нормальном распределении генеральной совокупности при уровне значимости $\alpha = 0{,}05$.

Решение. Перейдем от интервального статистического ряда к статистическому ряду, заменив каждый промежуток $[a_i\;;\;a_{i+1})$ его средним значением $x_i = \dfrac{a_i + a_{i+1}}{2}$. Получаем табл. 10.2.

По табл. 10.2 вычислим математическое ожидание $\overline{x}$, дисперсию $D_в$ и среднее квадратическое отклонение $\sigma_в$:

Таблица 10.2

x_i	m_i
–17,5	7
–12,5	11
–7,5	15
–2,5	24
2,5	49
7,5	41
12,5	26
17,5	17
22,5	7
27,5	3

По табл. 10.2 вычислим математическое ожидание $\overline{x}$, дисперсию $D_в$ и среднее квадратическое отклонение $\sigma_в$:

$$\overline{x} = \frac{1}{200}\sum_{i=1}^{10} x_i m_i =$$

$$= \frac{1}{200}\cdot(-17{,}5\cdot 7 - 12{,}5\cdot 11 - 7{,}5\cdot 15 - 2{,}5\cdot 24 + 2{,}5\cdot 49 + 7{,}5\cdot 41 +$$

$$+12{,}5\cdot 26 + 17{,}5\cdot 17 + 22{,}5\cdot 7 + 27{,}5\cdot 3) = \frac{860}{200} = 4{,}3;$$

$$D_в = \frac{1}{200}\sum_{i=1}^{10} x_i^2 m_i - \overline{x}^{\,2} =$$

$$= \frac{1}{200}\cdot((-17{,}5)^2\cdot 7 + (-12{,}5)^2\cdot 11 + (-7{,}5)^2\cdot 15 + (-2{,}5)^2\cdot 24 + 2{,}5^2\cdot 49 + 7{,}5^2\cdot 41 +$$

$$+12{,}5^2\cdot 26 + 17{,}5^2\cdot 17 + 22{,}5^2\cdot 7 + 27{,}5^2\cdot 3) - 4{,}3^2 = \frac{22550}{200} - 18{,}49 = 94{,}26;$$

$$\sigma_{в.} = \sqrt{D_в} = \sqrt{94{,}26} = 9{,}71.$$

3) Дальнейшие вычисления выполним по алгоритму критерия согласия Пирсона и оформим их в виде табл. 10.3, причем $x_1 = -\infty$, $x_9 = +\infty$.

Таблица 10.3

№	x_i	m_i	m_i^2	p_i	$m_{i\,теор}=$ $=200 \cdot p_i$	$\dfrac{m_i^2}{m_{i\,теор}}$
1	–17,5	7	49	0,0233	4,66	10,52
2	–12,5	11	121	0,0475	9,5	12,74
3	–7,5	15	225	0,0977	19,54	11,52
4	–2,5	24	576	0,1615	32,3	17,83
5	2,5	49	2401	0,1979	39,58	60,66
6	7,5	41	1681	0,1945	38,9	43,22
7	12,5	26	676	0,1419	28,38	23,82
8	17,5	17	289	0,0831	16,62	17,39
9	27,5	10	100	0,0526	10,52	9,51
$\sum$		200			200	207,21

Следовательно, находим $\chi^2_{набл} = 207,21 - 200 = 7,21$.

Из табл. П 2.5 (приложение 2) с учетом значений $\alpha = 0,05$ $k = l - 3 = 6$ находим: $\chi^2_{набл} = 12,6$.

Так как $7,21 < 12,6$, т.е. $\chi^2_{набл} < \chi^2_{крит}$, то нет основания отвергать гипотезу о нормальном распределении генеральной совокупности.

Ответ: принимаем гипотезу о нормальном распределении генеральной совокупности.

Задача 10.2. В результате контрольных испытаний из генеральной совокупности взята выборка объема $n=200$:

Таблица 10.4

x_i	6	8	10	12	14	16	18	20	22
m_i	16	24	28	32	25	24	20	18	15

Оценить с помощью критерия Пирсона гипотезу о нормальном распределении генеральной совокупности при уровне значимости $\alpha = 0,01$.

Решение. 1) По табл. 10.4 вычислим выборочные математическое ожидание $\overline{x}$, дисперсию $D_в$ и среднее квадратическое отклонение $\sigma_в$:

$$\overline{x} = \frac{1}{200}\sum_{i=1}^{9} x_i m_i =$$

$$= \frac{1}{200} \cdot (6 \cdot 16 + 8 \cdot 24 + 10 \cdot 28 + 12 \cdot 32 + 14 \cdot 25 + 16 \cdot 24 +$$

$$+ 18 \cdot 20 + 20 \cdot 18 + 22 \cdot 15) = 13{,}68;$$

$$D_в = \frac{1}{200} \sum_{i=1}^{9} x_i^2 m_i - \bar{x}^2 =$$

$$= \frac{1}{200} \cdot (36 \cdot 16 + 64 \cdot 24 + 100 \cdot 28 + 144 \cdot 32 + 196 \cdot 25 + 256 \cdot 24 +$$

$$+ 324 \cdot 20 + 400 \cdot 18 + 484 \cdot 15) - 187{,}1424 = 20{,}3776;$$

$$\sigma_{в.} = \sqrt{D_в} = \sqrt{20{,}3776} = 4{,}51.$$

2) Дальнейшие вычисления выполним по алгоритму критерия согласия Пирсона и оформим их в виде табл. 10.5.

Таблица 10.5

№	x_i	m_i	m_i^2	p_i	$m_{i\,теор} = 200 \cdot p_i$	$\frac{m_i^2}{m_{i\,теор}}$
1	6	16	256	0,0446	8,92	28,70
2	8	24	576	0,0592	11,84	48,65
3	10	28	784	0,1023	20,46	38,32
4	12	32	1024	0,1496	29,92	34,23
5	14	25	625	0,1722	34,44	18,15
6	16	24	576	0,1671	33,42	17,24
7	18	20	400	0,1339	26,78	14,94
8	20	18	324	0,0903	18,06	17,94
9	22	15	225	0,0808	16,16	13,92
$\sum$		200		1	200	232,09

Следовательно, находим: $\chi^2_{набл} = 232{,}09 - 200 = 32{,}09$.

Из табл. 10.5 (приложение 2) с учетом значений $\alpha = 0{,}01$ $k = l - 3 = 6$ находим: $\chi^2_{крит} = 16{,}8$. Так как $32{,}09 > 16{,}8$, то $\chi^2_{набл} > \chi^2_{крит}$. Следовательно, отвергаем гипотезу о нормальном распределении генеральной совокупности.

Ответ: отвергаем гипотезу о нормальном распределении генеральной совокупности.

Задача 10.3. Используя критерий Пирсона, при уровне значимости $\alpha = 0{,}05$ установить, случайно или значимо расхождение между эмпирическими частотами m_i и теоретическими частотами $m_{i\ \text{теор}}$,которые вычислены, исходя их гипотезы о нормальном распределении генеральной совокупности:

m_i	6	10	36	56	32	20	12	8
$m_{i\ \text{теор}}$	4	9	25	60	35	24	14	9

Решение. 1) Найдем $\chi^2_{\text{набл}}$:

$$\chi^2_{\text{набл}} = \sum_{i=1}^{8} \frac{(m_i - m_{i\ \text{теор}})^2}{m_{i\ \text{теор}}} =$$

$$= \frac{2^2}{4} + \frac{1^2}{9} + \frac{11^2}{25} + \frac{(-4)^2}{60} + \frac{(-3)^2}{35} + \frac{(-4)^2}{24} + \frac{(-2)^2}{14} + \frac{(-1)^2}{9} =$$

$$= 1 + \frac{1}{9} + \frac{121}{25} + \frac{4}{15} + \frac{9}{35} + \frac{2}{3} + \frac{2}{7} + \frac{1}{9} = 7{,}54$$

Из табл. 10.5 (приложение 2) с учетом значений $\alpha = 0{,}05$ $k = 8 - 3 = 5$ находим: $\chi^2_{\text{крит}} = 11{,}1$.

Так как $7{,}54 < 11{,}1$ то $\chi^2_{\text{набл}} < \chi^2_{\text{крит}}$. Следовательно, нет основания отвергать гипотезу о нормальном распределении генеральной совокупности.

Ответ: расхождение случайное.

Задачи

10.1. Используя критерий Пирсона, при уровне значимости $\alpha = 0{,}05$ установить, случайно или значимо расхождение между эмпирическими частотами m_i и теоретическими частотами $m_{i\ \text{теор}}$,которые вычислены, исходя из гипотезы о нормальном распределении генеральной совокупности:

а)

m_i	7	11	31	14	7
$m_{i\ \text{теор}}$	6	15	30	14	5

б)

m_i	10	17	23	31	29	20	12	8
$m_{i\ \text{теор}}$	7	12	21	45	28	19	11	7

в)

m_i	8	18	35	28	22	18	11
$m_{i\ \text{теор}}$	5	11	28	43	31	16	6

10.2. Используя критерий Пирсона, при уровне значимости α = 0,05 проверить гипотезу о нормальном распределении генеральной совокупности по ее выборке объема n = 200 :

№	x_i	m_i
1	1,2	6
2	1,4	9
3	1,6	26
4	1,8	25
5	2,0	30
6	2,2	26
7	2,4	21
8	2,6	24
9	2,8	20
10	3,0	8
11	3,2	5

10.3. Используя критерий Пирсона, при уровне значимости α = 0,05 проверить гипотезу о нормальном распределении генеральной совокупности по ее выборке объема n=150 , собранной в интервальный статистический ряд:

$[a_i\,;\,a_{i+1})$	m_i
[0 ; 4)	8
[4 ; 8)	12
[8 ; 12)	19
[12 ; 16)	42
[16 ; 20)	24
[20 ; 24)	20
[24 ; 28)	16
[28 ; 32]	9

10.4. Используя критерий Пирсона, при уровне значимости α = 0,05 проверить гипотезу о нормальном распределении генеральной совокупности по ее выборке:

3,45;	3,47;	3,47;	3,43;	3,46;	3,44;	3,40;	3,45;	3,41;	3,42;
3,47;	3,49;	3,41;	3,48;	3,43;	3,40;	3,43;	3,47;	3,45;	3,44;
3,41;	3,40;	3,48;	3,46;	3,51;	3,39;	3,50;	3,50;	3,47;	3,38;
3,44;	3,40;	3,40;	3,44;	3,47;	3,53;	3,46;	3,46;	3,52;	3,47;
3,41;	3,44;	3,47;	3,45;	3,44;	3,45;	3,47;	3,42;	3,44;	3,50;
3,45;	3,50;	3,42;	3,48;	3,40;	3,45;	3,48;	3,48;	3,46;	3,47;
3,44;	3,44;	3,47;	3,43;	3,44;	3,47;	3,44;	3,45;	3,44;	3,46;
3,46;	3,44;	3,44;	3,44;	3,44;	3,46;	3,44;	3,42;	3,50;	3,46;
3,48;	3,43;	3,40;	3,46;	3,46;	3,47;	3,45;	3,48;	3,42;	3,46;
3,48;	3,38;	3,45;	3,43;	3,52;	3,43;	3,50;	3,51;	3,41;	3,52.

10.5. Используя критерий Пирсона, при уровне значимости $\alpha = 0{,}05$ проверить гипотезу о нормальном распределении генеральной совокупности по ее выборке:

8,31	8,42	8,37	8,40	8,40	8,30	8,30	8,42	8,32	8,29
8,33	8,36	8,34	8,37	8,32	8,36	8,38	8,38	8,33	8,36
8,40	8,36	8,32	8,34	8,36	8,30	8,30	8,33	8,35	8,37
8,37	8,30	8,41	8,36	8,43	8,37	8,34	8,38	8,29	8,34
8,31	8,36	8,37	8,30	8,41	8,34	8,34	8,37	8,35	8,40
8,34	8,36	8,37	8,37	8,31	8,35	8,34	8,33	8,36	8,36
8,36	8,37	8,36	8,40	8,37	8,34	8,387	8,32	8,35	8,36
8,37	8,41	8,36	8,36	8,36	8,40	8,34	8,30	8,34	8,33
8,35	8,37	8,34	8,36	8,37	8,37	8,35	8,36	8,34	8,42
8,36	8,33	8,34	8,35	8,36	8,32	8,38	8,32	8,36	8,37

Ответы

10.1. а) Расхождение частот случайное ($\chi^2_{\text{набл}} = 2{,}07$; $\chi^2_{\text{крит}} = 6$);

б) Расхождение частот случайное ($\chi^2_{\text{набл}} = 8{,}23$; $\chi^2_{\text{крит}} = 11{,}1$);

в) Расхождение частот значимое ($\chi^2_{\text{набл}} = 20{,}26$; $\chi^2_{\text{крит}} = 9{,}5$).

10.2. Гипотеза о нормальном распределении принимается

($\chi^2_{\text{набл}} = 7{,}71$; $\chi^2_{\text{крит}} = 15{,}5$).

10.3. Гипотеза о нормальном распределении принимается

($\chi^2_{\text{набл}} = 8{,}065$; $\chi^2_{\text{крит}} = 15{,}1$).

10.4. Гипотеза о нормальном распределении отвергается

($\chi^2_{набл} = 15{,}97;\ \chi^2_{крит} = 7{,}8$).

10.5. Гипотеза о нормальном распределении отвергается

($\chi^2_{набл} = 15{,}46;\ \chi^2_{крит} = 13{,}3$).

§11. Элементы теории корреляции

Функциональная, статистическая и корреляционная зависимости

Определение 1. *Функциональной зависимостью* между случайными величинами X и Y называется зависимость, при которой изменение величины X влечет изменение значений Y, т.е. Y является функцией случайного аргумента X.

Определение 2. *Статистической зависимостью* между случайными величинами называется зависимость, при которой изменение одной из величин влечет изменение распределения другой

Определение 3. *Корреляционной зависимостью* между случайными величинами называется статистическая зависимость между ними, при которой изменение одной из величин влечет изменение среднего значения другой.

Определение 4. *Функцией регрессии* называется зависимость среднего значения одной из коррелированных случайных величин от другой, т.е. функция:

$$\overline{y} = f(x) \ \text{(регрессия } Y \text{ на } X) \quad \text{или} \quad \overline{x} = \varphi(y) \ \text{(регрессия } X \text{ на } Y).$$

Замечание 1. На практике наибольший интерес представляет анализ следующих вопросов:

- существует ли корреляционная зависимость между двумя случайными величинами;
- если корреляционная зависимость существует, то какой вид имеет функция регрессии (будет ли она линейной, параболической или какой-нибудь другой).

Коэффициент корреляции

Определение 5. Коэффициентом корреляции случайных величин X и Y называется число, вычисляемое по формуле:

$$r_{xy} = \frac{\mu_{xy}}{\sigma_x \cdot \sigma_y},$$

где $\mu_{xy} = M((X - M_x)\cdot(Y - M_y))$ – корреляционный момент; M_x , M_y – математические ожидания x и y ; σ_x, σ_y – средние квадратические отклонения величин X и Y соответственно.

Свойства коэффициента корреляции

1. $|r_{xy}| \le 1$;

2. Коэффициент корреляции служит для оценки «тесноты» *линейной* связи между X и Y:

- чем ближе абсолютная величина числа r_{xy} к единице, тем корреляционная связь между случайными величинами X и Y сильнее;
- чем ближе абсолютная величина r_{xy} к нулю, тем корреляционная связь между случайными величинами X и X слабее.

3. Если случайные величины связаны линейной зависимостью

$$Y = aX + b,$$

то абсолютная величина коэффициента корреляции равна 1.

Схема вычисления выборочного коэффициента корреляции

1) Провести n испытаний над случайными величинами X и Y , после чего получить выборки объемов n из генеральных совокупностей X и Y :

x_1 , x_2 , x_3 , …, x_n и y_1 , y_2 , y_3 , …, y_n ;

2) Найти: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$;

3) Найти: $\sigma_{вx}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2$, $\sigma_{вy}^2 = \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \bar{y}^2$;

4) Найти выборочный корреляционный момент:

$$\bar{\mu}_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})\cdot(y_i - \bar{y}) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y} \quad ;$$

5) Найти выборочный коэффициент корреляции:

$$\bar{r}_{xy} = \frac{\bar{\mu}_{xy}}{\sigma_{bx}\cdot\sigma_{by}} \ .$$

Замечание 2. Если число n велико, то выборки представляют интервальными статистическими рядами, причем число интервалов для X и Y может быть различным (например, k_1 и k_2) . По этим интервальным статистическим рядам подсчитывают частоты m_x , m_y и m_{xy} . Результаты вычислений заносят в корреляционную таблицу:

$y \backslash x$	x_1	x_2	…	x_{k_1}	m'_y
y_1	m_{11}	m_{21}	…	$m_{k_1 1}$	m'_1
y_2	m_{12}	m_{22}	…	$m_{k_1 2}$	m'_2
⋮	…	…	…	…	…
y_{k_2}	m_{1k_2}	m_{2k_2}	…	$m_{k_1 k_2}$	m'_{k_2}
m_x	m_1	m_2	…	m_{k_1}	n

Далее вычисляем:

1) $\bar{x} = \frac{1}{n}\sum_{i=1}^{k_1} x_i m_i$, $\bar{y} = \frac{1}{n}\sum_{j=1}^{k_2} y_j m'_j$,

2) $\sigma_{bx}^2 = \frac{1}{n}\sum_{i=1}^{k_1} x_i^2 m_i - \bar{x}^2$, $\sigma_{by}^2 = \frac{1}{n}\sum_{j=1}^{k_2} y_j^2 m'_j - \bar{y}^2$,

3) $\bar{\mu}_{xy} = \frac{1}{n}\sum_{i=1}^{k_1} x_i \cdot \left(\sum_{j=1}^{k_2} y_j m_{ji}\right) - \bar{x}\cdot\bar{y}$, 4) $\bar{r}_{xy} = \frac{\bar{\mu}_{xy}}{\sigma_{bx}\cdot\sigma_{by}}$.

Проверка значимости выборочного коэффициента корреляции

1) По двум выборкам объемов n из нормальных генеральных совокупностей X и Y найти выборочный коэффициент корреляции

$\bar{r}_{xy}$($\bar{r}_{xy} \neq 0$);

2) Проверить для генеральных совокупностей X и Y нулевую гипотезу $H_0 : r_{xy} = 0$:

- Если H_0 принимается, то нет корреляционной зависимости между случайными величинами X и Y ;
- Если H_0 отвергается, то существует корреляционная зависимость между случайными величинами X и Y .

Правило проверки гипотезы $H_0 : r_{xy} = 0$:

Чтобы при уровне значимости α проверить для генеральных совокупностей X и Y нулевую гипотезу $H_0 : r_{xy} = 0$, необходимо :

1) Вычислить наблюдаемое значение критерия

$$T_{\text{набл}} = \frac{\overline{r}_{xy} \cdot \sqrt{n-2}}{\sqrt{1-\left(\overline{r}_{xy}\right)^2}};$$

2) По табл. П 2.6 критических точек распределения Стьюдента (см. приложение 2) по уровню значимости α и числу степеней свободы $k = n - 2$ найти число $T_{\text{крит}}$;

3) Сравнить числа $|T_{\text{набл}}|$ и $T_{\text{крит}}$: если $|T_{\text{набл}}| < T_{\text{крит}}$, то нет основания отвергать гипотезу H_0 , если $|T_{\text{набл}}| > T_{\text{крит}}$, то гипотеза H_0 отвергается.

Линейная корреляция. Уравнение регрессии

Определение 6. *Корреляция* случайных величин X и Y называется *линейной*, если являются линейными функциями регрессии Y на X (т.е. $\overline{y} = f(x)$) и X на Y (т.е. $\overline{x} = \varphi(y)$).

Определение 7. *Выборочным уравнением линейной регрессии* Y на X называется уравнение вида :

$$y_x - \overline{y} = r_{xy} \cdot \frac{\sigma_{by}}{\sigma_{bx}}\left(x - \overline{x}\right) \;,$$

где, y_x – условная средняя; $\overline{x}, \overline{y}$ – выборочные средние; σ_{bx}, σ_{by} – выборочные средние квадратические отклонения; $\overline{r}_{xy}$ – выборочный коэффициент корреляции.

Выборочным уравнением линейной регрессии X на Y называется уравнение вида:

$$x_y - \overline{x} = r_{xy} \cdot \frac{\sigma_{bx}}{\sigma_{by}}\left(y - \overline{y}\right) \;,$$

Замечание 3. Следует иметь в виду, что прямые

$$y_x = ax + b \quad \text{и} \quad x_y = cy + d$$

в общем случае не совпадают.

Для их нахождения необходимо:

1) по выборочным данным найти числа $\overline{x}, \overline{y}$, σ_{bx}, $\sigma_{by}, \overline{r}_{xy}$:

2) Проверить гипотезу о существовании корреляционной связи между

X и Y ;

3) Составить уравнения обеих линий регрессии.

Замечание 4. Числа

$$\bar{r}_{xy} \cdot \frac{\sigma_{bx}}{\sigma_{by}} = \bar{p}_{yx} \qquad \text{и} \qquad \bar{r}_{xy} \cdot \frac{\sigma_{by}}{\sigma_{bx}} = \bar{p}_{xy}$$

называются *выборочным коэффициентом регрессии*, соответственно Y на X, или X на Y.

Ранговая корреляция

Рассмотрим выборку объема n из генеральной совокупности X, объекты которой обладают двумя качественными признаками: A и B.

1) Объекты выборки расположим в порядке ухудшения качества по каждому признаку

A: $x_1 \quad x_2 \quad x_3 \quad \ldots \quad x_n$

B: $y_1 \quad y_2 \quad y_3 \quad \ldots. \quad y_n$

2) присвоим объектам полученных двух выборок порядковые номера (ранги):

$$A: \begin{matrix} 1 & 2 & 3 & \ldots & n \\ x_1 & x_2 & x_3 & \ldots & x_n \end{matrix}$$

$$B: \begin{matrix} 1 & 2 & 3 & \ldots & n \\ y_1 & y_2 & y_3 & \cdots & y_n \end{matrix}$$

3) Составим две последовательности соответствующих друг другу в выборках рангов:

$$x_i: \quad 1 \quad 2 \quad 3 \quad \ldots \quad n$$

$$y_i: \quad a_1 \quad a_2 \quad a_3 \quad \ldots \quad a_n,$$

где a_k – ранг (т.е. порядковый номер) объекта y_{a_k}, соответствующего объекту x_k.

4) Вычислим разности $d_1 = 1 - a_1$, $d_2 = 2 - a_2$, …, $d_n = n - a_n$.

Определение 8. *Коэффициентом ранговой корреляции Спирмана* называется число:

$$\rho_в = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n^3 - n} \quad . \tag{1}$$

5) Найдем число R_1 – количество рангов справа от a_1, больших a_1. Аналогично R_2 – количество рангов справа от a_2, больших a_2. И так далее.

Обозначим : $R = R_1 + R_2 + ... + R_{n-1}$

Определение 9. *Коэффициентом ранговой корреляции Кендалла* называется число:

$$\tau_в = \frac{4R}{n \cdot (n-1)} - 1 \quad .$$

Замечание 5. Абсолютные величины коэффициентов ранговой корреляции не превышают единицы:

$$|\rho_в| \le 1 \qquad \text{и} \qquad |\tau_в| \le 1$$

Правило проверки наличия связи между качественными признаками

Для обоснованного суждения о наличии связи между качественными признаками A и B генеральной совокупности X следует проверить, значим ли выборочный коэффициент их ранговой корреляции.

1) Пусть при уровне значимости α нулевая гипотеза H_0 : между признаками A и B нет значимой связи.

- По табл. П 2.6 критических точек распределения Стьюдента (см. приложение 2) по уровню значимости α и числу степеней свободы $k = n - 2$ найти число $t_{крит}$;
- Найти число $T_{крит} = t_{крит} \cdot \sqrt{\dfrac{1 - p_в^2}{n - 2}}$,

где n – объем выборки;

$\rho_в$ – выборочный коэффициент ранговой корреляции Спирмена;

- Сравнить числа $T_{крит}$, $|\rho_в|$:

если $|\rho_в| < T_{крит}$, то нет основания отвергать гипотезу H_0 , т.е. ранговая корреляция между признаками не значима;

если $|\rho_в| > T_{крит}$ то нулевая гипотеза H_0 отвергается, т.е. между качественными признаками существует значимая корреляционная связь.

2) Пусть при уровне значимости α нулевая гипотеза H_0 : между признаками A и B генеральной совокупности нет значимой связи.

- По табл. П 2.2 функции Лапласа (см. приложение 2) по уровню значимости α из равенства $\Phi(z_{крит}) = \frac{1-\alpha}{2}$ найти число $z_{крит}$;

- Найти число $T_{крит} = z_{крит} \cdot \sqrt{\frac{2\cdot(2n+5)}{9n\cdot(n-1)}}$,

где n – объем выборки;

- Сравнить числа $T_{крит}$, $|\tau_в|$:

если $|\tau_в| < T_{крит}$, то нет основания отвергать гипотезу H_0 , т.е. ранговая корреляция между признаками не значима ;

если $|\tau_в| > T_{крит}$, то нулевая гипотеза H_0 отвергается, т.е. между качественными признаками существует значимая корреляционная связь.

Задачи с решениями

Задача 11.1. Найти выборочный коэффициент корреляции и уравнение линейной регрессии Y на X по данным пяти наблюдений:

X: 2 2,5 3 3,5 4

Y: 1,25 1,45 1,65 1,85 2,05

Решение. Используем формулы:

1) Выборочный коэффициент корреляции:

$$\bar{r}_{xy} = \frac{\bar{\mu}_{xy}}{\sigma_{bx}\cdot\sigma_{by}} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2}\cdot\sqrt{\frac{1}{n}\sum_{i=1}^{n} y_i^2 - \bar{y}^2}};$$

2) линейное уравнение регрессии Y на X :

$$y_x - \bar{y} = r_{xy}\cdot\frac{\sigma_{by}}{\sigma_{bx}}\left(x - \bar{x}\right),$$

где $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$, $\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$

$\sigma_{вх}^2=\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2$, $\sigma_{вy}^2=\frac{1}{n}\sum_{i=1}^{n}y_i^2-\bar{y}^2$, $\bar{\mu}_{xy}=\frac{1}{n}\sum_{i=1}^{n}x_iy_i-\bar{x}\bar{y}$.

Проведем необходимые вычисления, для чего составим расчетную таблицу:

№	x_i	y_i	x_i^2	y_i^2	x_iy_i
1	2	1,25	4	1,5625	2,5
2	2,5	1,45	6,25	2,1025	3,625
3	3	1,65	9	2,7225	4,95
4	3,5	1,85	12,25	3,4225	6,475
5	4	2,05	16	4,2025	8,2
$\sum$	15	8,25	47,5	14,0125	25,75

Тогда получаем:

$$\bar{x}=\frac{15}{5}=3,\quad \bar{y}=\frac{8,25}{5}=1,65,$$

$$\sigma_{вх}^2=\frac{1}{5}\cdot 47,5-3^2=9,5-9=0,5,\quad \sigma_{вх}=0,708$$

$$\sigma_{вy}^2=\frac{1}{5}\cdot 14,0125-1,65^2=2,8025-2,7225=0,08,\quad \sigma_{вх}=0,284,$$

$$\bar{\mu}_{xy}=\frac{1}{5}\cdot 25,75-3\cdot 1,65=5,15-4,95=0,2,$$

$$\bar{r}_{xy}=\frac{0,2}{0,708\cdot 0,284}=0,99.$$

Запишем уравнение линейной регрессии Y на X :

$$y_x-1,65=\frac{0,2}{0,5}(x-3),\quad \text{т.е. } y_x=0,4x+0,45.$$

Ответ: $\bar{r}_{xy}=0,99$, $y_x=0,4x+0,45$.

Задача 11.2. Найти выборочный коэффициент корреляции и выборочные уравнения линейных регрессий Y на X и X на Y по данным выборки X и Y , сведенным в корреляционную таблицу:

$Y \backslash X$	5	10	15	20	25	30	35	40	m_y
100	2	1							3
120	3	4	3						10
140			5	10	8				23
160				1		6	1	1	9
180							4	1	5
m_x	5	5	8	11	*8*	*6*	5	2	50

Р е ш е н и е .

1) Найдем оценки математических ожиданий X и Y :

$$\bar{x}=\frac{1}{50}(5(2+3)+10(1+4)+15(3+5)+20(10+1)+25\cdot 8+30\cdot 6+35(1+4)+40(1+1))=19\ ;$$

$$\bar{y}=\frac{1}{50}(100(2+1)+120(3+4+3)+140(5+10+8)+160(1+6+1+1)+180(4+1))=141{,}2$$

2) Найдем выборочные дисперсии:

$$D_{вx}=\frac{1}{50}(5^2(2+3)+10^2(1+4)+15^2(3+5)+20^2(10+1)+25^2\cdot 8+30^2\cdot 6+$$

$$+35^3(1+4)+40^2(1+1))-19^2=531-361=170\ ,$$

$$D_{вy}=\frac{1}{50}(100^2(2+1)+120^2(3+4+3)+140^2(5+10+8)+$$

$$+160^2\cdot 9+180^2\cdot 5)-141{,}2^2=20344-19937{,}44=406{,}56\ .$$

3) Найдем выборочные средние квадратические отклонения:

$$\sigma_{вx}=\sqrt{D_{вx}}=13{,}04\,,\quad \sigma_{вy}=\sqrt{D_{вy}}=20{,}16\ .$$

4) Найдем выборочный корреляционный момент:

$$\bar{\mu}_{xy}=\frac{1}{50}(5(100\cdot 2+120\cdot 3)+10(100\cdot 1+120\cdot 4)+15(120\cdot 3+140\cdot 5)+$$

$$+20(140\cdot 10+160\cdot 1)+25(140\cdot 5)+30(160\cdot 6)+35(160\cdot 1+180\cdot 4)+$$

$$+40(160\cdot 1+180\cdot 1))-19\cdot 141{,}2=2928-2682{,}8=245{,}2\ .$$

5) Найдем выборочный коэффициент корреляции:

$$r_{xy}=\frac{\mu_{xy}}{\sigma_x\cdot\sigma_y}=\frac{245{,}2}{13{,}04\cdot 20{,}16}=\frac{245{,}2}{262{,}9}=0{,}93\ .$$

6) Напишем выборочное уравнение линейной регрессии Y на X:

$$y_x-\overline{y}=\frac{\overline{\mu}_{xy}}{\sigma_{bx}^2}\left(x-\overline{x}\right)\quad\Rightarrow\quad y_x-141{,}2=\frac{245{,}2}{170}(x-19)\quad\Rightarrow$$

$$\Rightarrow\quad y_x=141{,}2+1{,}44x-27{,}4\quad\Rightarrow\quad y_x=1{,}44x+116{,}8$$

7) Напишем выборочное уравнение линейной регрессии X на Y:

$$x_y-\overline{x}=\frac{\overline{\mu}_{xy}}{\sigma_{by}^2}\left(y-\overline{y}\right)\quad\Rightarrow\quad x_y-19=\frac{245}{406{,}56}(y-141{,}2)\quad\Rightarrow$$

$$\Rightarrow\quad x_y-19=0{,}6(y-141{,}2)\quad\Rightarrow\quad x_y=0{,}6y-65{,}72$$

Ответ: $\overline{r}_{xy}=0{,}93$, $\quad y_x=1{,}44x+116{,}8$, $\quad x_y=0{,}6y-65{,}72$.

Задача 11.3. Знания 10 студентов проверены по двум тестам A и B. Оценки по стобальной системе оказались следующими:

По A: 92 96 90 50 75 83 65 70 62 55

По B: 94 98 84 52 70 87 62 74 59 50.

Найти выборочный коэффициент ранговой корреляции: а) Спирмена; б) Кендалла и оценить их значимость при уровне значимости α=0,1.

Решение. 1) Присвоим ранги a_i оценкам x_i по тесту A, расположив эти оценки в порядке убывания:

a_i	1	2	3	4	5	6	7	8	9	10
x_i	96	92	90	83	75	70	65	62	55	50

2) Присвоим ранги b_i оценкам y_i по тесту B, расположив их в порядке убывания:

b_i	1	2	3	4	5	6	7	8	9	10
y_i	98	94	87	84	74	70	62	59	52	50

3) Рангу a_1=1 оценки 96 по тесту A соответствует ранг b_1 оценки 98 (первого студента) по тесту B

Рангу a_2=2 оценки 92 по тесту A соответствует ранг b_2=2 оценки 94 по тесту B.

Рангу a_3=3 оценки 90 по тесту A соответствует ранг b_4=4 оценки 84.

Аналогично получаем:

b_4= 3, b_5= 6, b_6=5, b_7=7, b_8=8, b_9=10, b_{10}=9.

4) Выпишем последовательности рангов a_i и b_i:

a_i	1	2	3	4	5	6	7	8	9	10
b_i	1	2	4	3	6	5	7	8	10	9

и получим разности рангов:

$$d_1=a_1-b_1=0;\quad d_2=2-2=0;\quad d_3=-1; d_4=1; d_5=-1;\quad d_6=1; d_7=0;\ d_8=0; d_9=-1; d_{10}=1.$$

5) Вычислим выборочный коэффициент ранговой корреляции Спирмена:

$$\rho_b=1-\frac{6\sum d_i^2}{n^3-n}=1-\frac{6\cdot 6}{1000-10}=1-\frac{36}{990}=1-0{,}04=0{,}96.$$

6) Вычислим выборочный коэффициент ранговой корреляции Кендалла:

$$\tau_b=\frac{4R}{n(n-1)}-1$$

где $R=R_1+R_2+\ldots+R_{n-1}$ = 9 + 8 + 6 + 6 + 4 + 4 + 3 + 2 + 0 = 42.

Тогда получаем:

$$\tau_в=\frac{4\cdot 42}{10(10-1)}-1=\frac{168}{90}-1=1{,}87-1=0{,}87.$$

7) При уровне значимости $\alpha=0{,}1$ находим число

$$T_{крит}=t_{крит}\cdot\sqrt{\frac{1-\rho_b^2}{n-2}}=1{,}86\sqrt{\frac{1-0{,}92}{8}}=0{,}186\,,$$

Сравниваем числа $T_{крит}$ и $|\rho_b|$: так как 0,96 > 0,186, то $|\rho_b|>T_{крит}$ и ранговая корреляция между признаками является значимой.

8) При уровне значимости α = 0,1 находим:

$$T_{\text{крит}} = z_{\text{крит}} \sqrt{\frac{2 \cdot (2n+5)}{9n \cdot (n-1)}} = 1{,}645 \cdot \sqrt{\frac{2 \cdot 25}{90 \cdot 9}} = 1{,}645 \cdot \frac{\sqrt{5}}{9} = 0{,}43.$$

Сравним числа $T_{\text{крит}}$ и $|\tau_b|$:

так как 0,87 > 0,43 , то $|\tau_b| > T_{\text{крит}}$

Значит корреляционная связь между сравниваемыми оценками значимая.

Ответ: $\rho_b = 0{,}96$; $\tau_b = 0{,}87$; гипотеза о наличии корреляционной связи между оценками принимается.

Задачи

11.1. Найдите выборочный коэффициент корреляции и выборочное линейное уравнение *Y* на *X* по данным семи наблюдений:

x_i	4,0	4,25	4,5	4,75	5,0	5,25	5,5
y_i	1,25	1,35	1,50	1,65	1,80	2,05	2,30

11.2. Найдите выборочный коэффициент корреляции и выборочное линейное уравнение *Y* на *X* по данным пяти наблюдений:

x_i	1,25	2,05	3,1	3,95	5,0
y_i	4,2	2,5	3,5	1,0	2,1

11.3. Даны результаты 50-ти наблюдений, собранные в корреляционную таблицу:

Y \ X	18	23	28	33	38	43	48	m_y
125		1						1
150	1	2	5					8
175		3	2	12				17
200			1	8	7			16
225					3	3		6
250						1	1	2
m_i	1	6	8	20	10	4	1	50

Найти выборочный коэффициент корреляции и выборочные линейные уравнения регрессий Y на X и X на Y , проверив гипотезу значимости выборочного коэффициента корреляции при уровне значимости $\alpha = 0{,}1$.

11.4. По данным 50-ти наблюдений, собранным в корреляционную таблицу:

Y \ X	5	10	15	20	25	30	35	m_y
100						4	1	5
120						6	2	8
140			8	10	5			23
160	3	4	3					10
180	2	1		1				4
m_i	5	5	11	11	5	10	3	50

Найти выборочный коэффициент корреляции и выборочные линейные уравнения регрессий Y на X и X на Y, проверив гипотезу значимости выборочного коэффициента корреляции при уровне значимости $\alpha = 0{,}05$.

11.5. В результате 79 опытов получена корреляционная таблица:

Y \ X	0,5	0,6	0,7	0,8	0,9	m_y
0,5			2	21	1	24
0,6	2	4	12	14		32
0,7		2	3			5
0,8	8	9	1			18
m_i	10	15	18	35	1	79

Определить выборочный коэффициент корреляции, проверить гипотезу значимости коэффициента корреляции при уровне значимости $\alpha = 0{,}1$, написать выборочные уравнения регрессий Y на X и X на Y.

11.6. В результате 60 опытов получена корреляционная таблица величин X и Y:

Y \ X	7,0	7,5	8,0	8,5	9,0	9,5	m_y
0	2	5					7
5		3	10	2			15
10			4	6	10	3	23
15				4	5	2	11
20					3	1	4
m_i	2	8	14	12	18	6	60

Определить выборочный коэффициент корреляции, проверить гипотезу значимости коэффициента корреляции при уровне значимости $\alpha = 0{,}05$, написать выборочные уравнения регрессий Y на X и X на Y.

11.7. Знания 10 студентов оценены двумя преподавателями по стобальной системе и выставлены следующие оценки:

98	94	88	80	76	70	63	61	60	58
99	91	93	74	78	65	64	66	52	53

Найти выборочные коэффициенты ранговой корреляции Спирмена и Кендалла и проверить их значимость при уровне значимости $\alpha = 0{,}05$.

11.8. Два контролера расположили 10 деталей в порядке ухудшения их качества. В результате получены две последовательности рангов:

1	2	3	4	5	6	7	8	9	10
1	2	4	3	6	5	7	10	9	8

Найти выборочные коэффициенты ранговой корреляции Спирмена и Кендалла и проверить их значимость при уровне значимости $\alpha = 0{,}1$.

11.9. Три арбитра A, B и C оценили мастерство 10 спортсменов. В итоге были получены три последовательности рангов:

A:	1	2	3	4	5	6	7	8	9	10;
B:	3	10	7	2	8	5	6	9	1	4;
C:	6	2	1	3	9	4	5	7	10	8.

Определите пару арбитров, оценки которых наиболее согласуются, используя:

а) выборочный коэффициент ранговой корреляции Спирмена;

б) выборочный коэффициент ранговой корреляции Кендалла.

Ответы

11.1. $\overline{r}_{xy} = 0{,}99$; $y_x = 0{,}69x - 1{,}59$.

11.2. $\overline{r}_{xy} = -0{,}7$; $x_y = 5{,}54 - 0{,}93y$.

11.3. $\overline{r}_{xy} = 0{,}84$; $y_x = 3{,}68x + 66{,}12$; $x_y = 0{,}19y - 3$.

11.4. $\overline{r}_{xy} = -0{,}85$; $y_x = 180{,}7 - 2{,}06x$; $x_y = 69 - 0{,}35y$.

11.5. $\overline{r}_{xy} = -0{,}86$; $y_x = -0{,}828x + 1{,}204$; $x_y = -0{,}908y + 1{,}268$.

11.6. $\overline{r}_{xy} = 0{,}71$; $y_x = 5{,}61x - 38{,}2$; $x_y = 0{,}9y + 7{,}62$.

11.7. $\rho_b = 0{,}927;\ \tau_b = 0{,}82$; гипотеза о наличии корреляционной связи между оценками принимается.

11.8. $\rho_b = 0{,}927;\ \tau_b = 0{,}78$; гипотеза о наличии корреляционной связи между наблюдаемыми величинами принимается.

11.9. а) $\rho_{b(AB)} = -0{,}21;\quad \rho_{b(AC)} = 0{,}64;\quad \rho_{b(BC)} = -0{,}3.$

Наиболее согласуются оценки арбитров A и C;

б) $\tau_{b(AB)} = -0{,}16;\ \tau_{b(AC)} = 0{,}51;\ \tau_{b(BC)} = -0{,}16.$

Наиболее согласуются оценки арбитров A и C.

Приложение 1

Контрольные работы и контрольные вопросы по теории

1. Элементы теории вероятностей

Задача П 1.1. На полке стоят 10 книг, из них 3 словаря, 4 справочника и 3 учебника. Какова вероятность того, что из пяти наудачу взятых книг окажется 2 словаря, 2 справочника и один учебник?

Решение. В данном случае общее число книг равно 10. Из них 5 книг можно выбрать n различными способами, где

$$n = C_{10}^5 = \frac{10!}{5!\cdot(10-5)!} = \frac{10!}{5!\cdot 5!} = \frac{5!\cdot 6\cdot 7\cdot 8\cdot 9\cdot 10}{5!\cdot 1\cdot 2\cdot 3\cdot 4\cdot 5} = 7\cdot 4\cdot 9 = 252.$$

Найдем число m событий, благоприятствующих выбору 2-х словарей (из 3-х имеющихся), 2-х справочников (из 4-х имеющихся) и одного учебника (из 3-х имеющихся). Получим

$$m = C_3^2\cdot C_4^2\cdot C_3^1 = \frac{3!}{2!\cdot 1!}\cdot\frac{4!}{2!\cdot 2!}\cdot\frac{3!}{1!\cdot 2!} = \frac{3\cdot 3\cdot 4\cdot 3}{2} = 54.$$

Следовательно, искомая вероятность вычисляется по формуле:

$$p = \frac{m}{n} = \frac{54}{252} = \frac{3}{14}.$$

Ответ: $\frac{3}{14}$.

Задача П 1.2. Баскетболист бросает мяч пять раз. Вероятность попадания при каждом броске равна 0,7. Найти вероятность того, что он попадет в корзину: а) три раза; б) менее трех раз; в) более трех раз.

Решение. Для решения задачи воспользуемся формулой Бернулли:

$$P_n(m) = C_n^m\cdot p^m\cdot(1-p)^{n-m},$$

где n – число выполненных бросков; m – число попаданий мяча из этих n бросков; p – вероятность попадания при одном броске.

В данной задаче n=5, p=0,7.

а) m=3. Следовательно,

$$P = P_5(3) = C_5^3 \cdot 0{,}7^3 \cdot 0{,}3^2 = \frac{5!}{3! \cdot 2!} \cdot 0{,}7^3 \cdot 0{,}3^2 = 0{,}3087.$$

б) m<3 $\Rightarrow$ $m = 0$ или m=1, или m=2. Следовательно, получаем:

$$P = P_5(0) + P_5(1) + P_5(2) =$$
$$= C_5^0 \cdot 0{,}7^0 \cdot 0{,}3^5 + C_5^1 \cdot 0{,}7^1 \cdot 0{,}3^4 + C_5^2 \cdot 0{,}7^2 \cdot 0{,}3^3 =$$
$$= 0{,}3^5 + 5 \cdot 0{,}7 \cdot 0{,}3^4 + 10 \cdot 0{,}7^2 \cdot 0{,}3^3 = 0{,}16308.$$

в) m>3 $\Rightarrow$ $m = 4$ или m=5. Следовательно, получаем:

$$P = P_5(4) + P_5(5) = C_5^4 \cdot 0{,}7^4 \cdot 0{,}3 + C_5^5 \cdot 0{,}7^5 =$$

$=5 \cdot 0{,}7^4 \cdot 0{,}3 + 0{,}7^5 = 0{,}52822.$

Ответ: а) 0,3087; б) 0,16308; в) 0,52822.

Задача П 1.3. В первой урне лежат 5 белых и 10 черных шаров, во второй – 3 белых и 7 черных шаров. Из второй урны в первую переложили какой-то один шар, а затем из первой урны вынули наугад один шар. Определить вероятность того, что вынутый шар – белый.

Решение. После того, как из второй урны в первую был переложен шар, в первой урне оказалось 16 шаров:

1) или 6 белых и 10 черных, если добавленный шар был белым (одним из тех 3-х, что лежали во второй урне);
2) или 5 белых и 11 черных, если добавленный шар был черным (одним их тех семи, что лежали во второй урне).

   Обозначим события: H_1 – взяли из второй урны белый шар, H_2 – взяли из второй урны черный шар.

H_1 и H_2 предшествуют событию A. Они являются попарно несовместными и $H_1 + H_2 = E$, т.е. образуют полную группу. Вычислим:

$$P(H_1) = \frac{3}{10}, \quad P(H_2) = \frac{7}{10}, \quad P_{H_1}(A) = \frac{6}{16} = \frac{3}{8}, \quad P_{H_2}(A) = \frac{5}{16}.$$

Поэтому по формуле полной вероятности находим:

$$P(A) = P(H_1) \cdot P_{H_1}(A) + P(H_2) \cdot P_{H_2}(A) =$$
$$= \frac{3}{10} \cdot \frac{3}{8} + \frac{7}{10} \cdot \frac{5}{16} = \frac{18 + 35}{160} = \frac{53}{160}.$$

Ответ: $\frac{53}{160}$.

Задача П 1.4. В партии из 10 деталей имеется 8 новых и две бывших в употреблении. Наудачу отобраны две детали.

а) Составить закон распределения случайной величины X – числа новых деталей среди отобранных.

б) Вычислить числовые характеристики случайной величины X.

Решение. а) X – дискретная случайная величина. Она имеет следующие возможные значения: $x_1 = 0$, $x_2 = 1$, $x_3 = 2$. Вероятность этих значений вычислим по формуле:

$$p_k = P(X = k) = \frac{C_n^k \cdot C_{s-n}^{m-k}}{C_s^m},$$

где $s = 10$ – общее число деталей в партии; $n = 8$ – число новых деталей в партии; $m = 2$ – число отобранных деталей; k – число новых деталей среди отобранных.

Тогда получаем:

$$p_1 = P(X = 0) = \frac{C_8^0 \cdot C_2^2}{C_{10}^2} = \frac{1}{\frac{10!}{2! \cdot 8!}} = \frac{2}{10 \cdot 9} = \frac{1}{45};$$

$$p_2 = P(X = 1) = \frac{C_8^1 \cdot C_2^1}{C_{10}^2} = \frac{8 \cdot 2}{45} = \frac{16}{45};$$

$$p_3 = P(X = 2) = \frac{C_8^2 \cdot C_2^0}{C_{10}^2} = \frac{\frac{8!}{2! \cdot 6!}}{45} = \frac{7 \cdot 8}{2 \cdot 45} = \frac{28}{45}.$$

Контроль: $p_1 + p_2 + p_3 = \frac{1}{45} + \frac{16}{45} + \frac{28}{45} = 1$.

Следовательно, искомый закон распределения случайной величины X задается табл. П 1.1:

Таблица П 1.1

X	0	1	2
p	$\frac{1}{45}$	$\frac{16}{45}$	$\frac{28}{45}$

б) По определению:

$$M(X)=\sum_{i=1}^{n} x_i\,p_i,\qquad D(X)=\sum_{i=1}^{n} x_i^2\,p_i-(M(X))^2,\qquad \sigma(X)=\sqrt{D(X)}.$$

Тогда, пользуясь табл. П 1.1, вычисляем:

$$M(X)=0\cdot\frac{1}{45}+1\cdot\frac{16}{45}+2\cdot\frac{28}{45}=\frac{16+56}{45}=\frac{72}{45}=\frac{8}{5};$$

$$D(X)=M(X^2)-(M(X))^2=0\cdot\frac{1}{45}+1\cdot\frac{16}{45}+4\cdot\frac{28}{45}-\left(\frac{8}{5}\right)^2=$$

$$=\frac{16+112}{45}-\frac{64}{25}=\frac{640-576}{25\cdot 9}=\frac{64}{225};$$

$$\sigma(X)=\sqrt{D(X)}=\frac{8}{15}.$$

Ответ: а) табл. П 1.1; б) $M(X)=\frac{8}{5}$; $D(X)=\frac{64}{225}$; $\sigma(X)=\frac{8}{15}$.

Задача П 1.5. Случайная величина X задана функцией распределения вероятностей:

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

Найти: а) плотность распределения вероятностей $f(x)$;

б) числовые характеристики случайной величины X;

в) вероятность попадания величины X в интервал [1; 2,5).

Решение. Рассматриваемая случайная величина X является непрерывной, так как функция $F(x)$ непрерывна на $(-\infty;\ +\infty)$,

Её график изображен на рис. П 1.1.

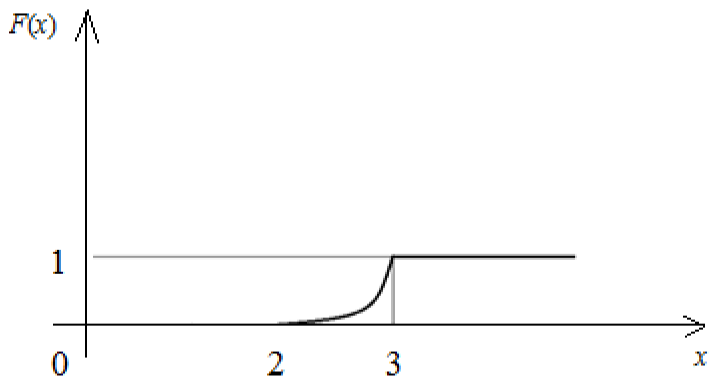


Рис. П 1.1

а) Так как функция $f(x)$ и $F(x)$ связаны равенством

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

То получаем:

$$f(x) = \begin{cases} 0 & \text{при } x \le 2 \\ 2x - 4 & \text{при } 2 < x \le 3, \\ 0 & \text{при } x > 3. \end{cases} \qquad (1)$$

График функции $f(x)$ изображен на рис. П 1.2.

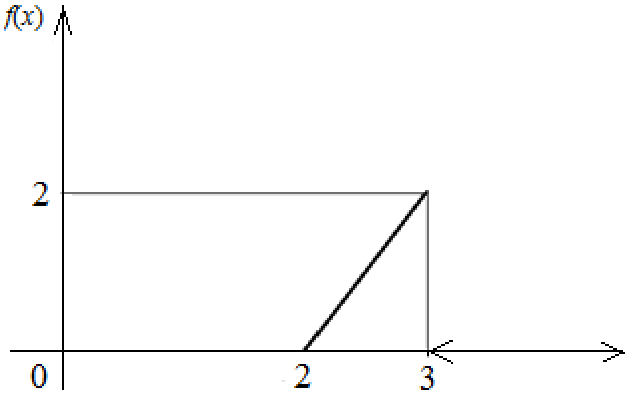


Рис. П 1.2

б) Вычисляем:

$$M(X) = \int_{-\infty}^{+\infty} x \cdot f(x)dx = \int_{2}^{3} x(2x-4)dx = \int_{2}^{3} (2x^2 - 4x)dx =$$

$$= \left(\frac{2x^3}{3} - 2x^2\right)\Bigg|_2^3 = 18 - 18 - \frac{16}{3} + 8 = \frac{8}{3};$$

$$D(X) = \int_{-\infty}^{+\infty} x^2 f(x)dx - (M(X))^2 = \int_{2}^{3} (2x^3 - 4x^2)dx - \left(\frac{8}{3}\right)^2 =$$

$$= \left(\frac{x^4}{2} - \frac{4x^3}{3}\right)\Bigg|_2^3 - \frac{64}{9} = \frac{81}{2} - 36 - 8 + \frac{32}{3} - \frac{64}{9} = \frac{1}{18};$$

$$\sigma(X) = \sqrt{D(X)} = \sqrt{\frac{1}{18}} = \frac{\sqrt{2}}{6}.$$

в) $P(1 \le X < 2{,}5) = F(2{,}5) - F(1) = (2{,}5 - 2)^2 - 0 = \frac{1}{4}.$

Ответ: а) формула (1); б) $M(X) = \frac{8}{3}$; $D(X) = \frac{1}{18}$; $\sigma(X) = \frac{\sqrt{2}}{6}$; в) $\frac{1}{4}$.

2. Элементы математической статистики

Задача П 1.6. Дана выборка объема n=30:

42	51	36	43	52	37	45	49
42	51	45	43	45	48	44	40
45	46	44	43	47	38	47	48
46	40	44	37	39	46.		

Требуется:

1) Найти статистический ряд и построить полигон частот;

2) Составить интервальный статистический ряд, взяв 7–10 интервалов, и построить гистограмму частот;

3) Найти оценки математического ожидания $\bar{x}$, выборочную дисперсию D_b, исправленную выборочную дисперсию S^2, выборочное среднее квадратическое отклонение $\sigma_в$, исправленное среднее квадратическое отклонение S;

4) С доверительной вероятностью $\gamma = 0{,}99$ найти доверительный интервал

а) для математического ожидания $M(X)$ в случае известной дисперсии, предполагая $D(X)= S^2$,

б) для математического ожидания $M(X)$ в случае неизвестной дисперсии,

в) для среднего квадратического отклонения $\sigma = \sqrt{D(X)}$;

5) Проверить критерий Пирсона о нормальном распределении наблюдаемой случайной величины X при уровне значимости $\alpha = 0{,}01$.

Решение. 1) По данной выборке находим: $x_{max} = 52$, $x_{min} = 36$.

Строим статистический ряд:

Нанесем на плоскости Oxy точки $M_i\left(x_i; \frac{m_i}{n}\right)$,

где i – порядковый номер варианты x_i, $i = 1, \ldots, 15$. Соединив эти точки последовательно, получим ломаную линию – полигон частот задачи (рис. П 1.3).

2) Найдем «размах» выборки: $x_{max} - x_{min} = 52 - 36 = 16$. Поэтому для составления интервального статистического ряда выберем число интервалов из условия:

Таблица П 1.2

№	x_i	m_i	m_i /n
1	36	1	1/30
2	37	2	1/15
3	38	1	1/30
4	39	1	1/30
5	40	2	1/15
6	42	2	1/15
7	43	3	1/10
8	44	3	1/10
9	45	4	2/15
10	46	3	1/10
11	47	2	1/15
12	48	2	1/15
13	49	1	1/30
14	51	2	1/15
15	52	1	1/30
$\sum$		30	1

$$7 \leq \frac{x_{max} - x_{min}}{l} \leq 10,$$

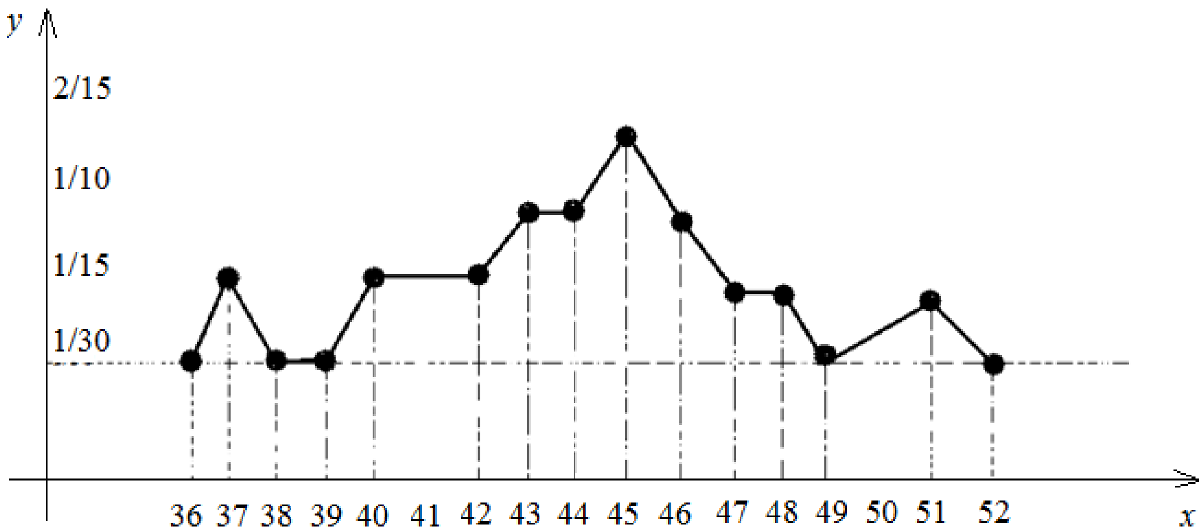


Рис. П 1.3

где l – длина интервала. Отсюда находим:

$$7 \leq \frac{16}{l} \leq 10 \Leftrightarrow 1{,}6 \leq l \leq 2\frac{2}{7}.$$

Следовательно, выберем $l = 2$, тогда число интервалов будет равно 8.

Интервальный статистический ряд указан в табл. П 1.2.

Таблица П 1.2

№	$[a_i\,;\,a_{i+1})$	m_i
1	[36;38)	3
2	[38;40)	2
3	[40;42)	2
4	[42;44)	5
5	[44;46)	7
6	[46;48)	5
7	[48;50)	3
8	[50;52]	3

В системе координат Oxy на оси Ox отложим точки $a_1,\dots,a_9$.

Построим прямоугольники с основанием $[a_i\,;\,a_{i+1})$ и высотой $\frac{m_i}{30}$,

где $i = 1,\dots,8$. Построенное ступенчатое тело – гистограмма частот задачи (рис. П 1.4).

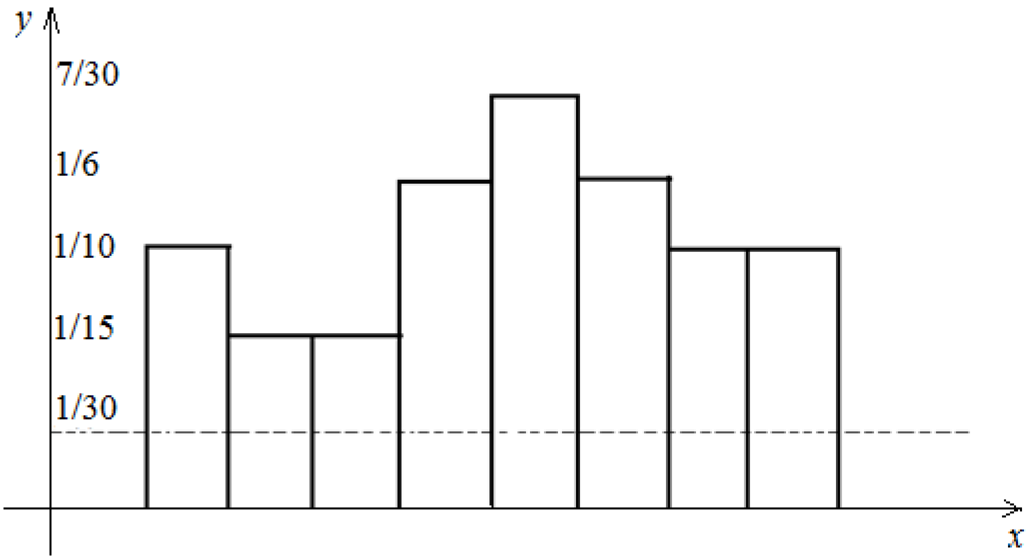


Рис. П 1.4

3) Для нахождения оценок параметров выборки составим по интервальному статистическому ряду расчетную табл. П 1.3, заменив в ней каждый интервал его средним значением $x_i = \frac{a_i + a_{i+1}}{2}$.

Таблица П 1.3

№	$[a_i\,;\,a_{i+1})$	x_i	m_i	$x_i\,m_i$	x_i^2	$x_i^2\,m_i$
1	[36;38)	37	3	111	1369	4107
2	[38;40)	39	2	78	1521	3042
3	[40;42)	41	2	82	1681	3362
4	[42;44)	43	5	215	1849	9245
5	[44;46)	45	7	315	2025	14175
6	[46;48)	47	5	235	2209	11045
7	[48;50)	49	3	147	2401	7203
8	[50;52]	51	3	153	2601	7803
$\sum$			30	1336		59982

Тогда получаем:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{8} x_i\,m_i = \frac{1336}{30} = 44{,}5;$$

$$D_b = \frac{1}{n}\sum_{i=1}^{8} x_i^2\,m_i - \overline{x^2} = \frac{59982}{30} - 44{,}5^2 = 19{,}15;$$

$$s^2 = \frac{n}{n-1}D_b = \frac{30}{29}\cdot 19{,}15 = 19{,}81;$$

$$\sigma_b = \sqrt{D_b} = 4{,}38;\;\; s = \sqrt{s^2} = 4{,}45.$$

4) а) При построении доверительного интервала для математического ожидания $M(X)=m$ с известной дисперсией $D(X)=s^2=19{,}81$ воспользуемся формулой:

$$\overline{x} - z_\gamma\frac{\sigma}{\sqrt{n}} < m < \overline{x} + z_\gamma\sigma/\sqrt{n},$$

где $\overline{x} = 44{,}5$; $\sigma = s = 4{,}45$; $n = 30$. Число z_γ находим с помощью табл. П 2.2 (приложение 2) из уравнения:

$$\Phi(z_\gamma) = \frac{\gamma}{2} \Leftrightarrow \Phi(z_\gamma) = 0{,}495 \Leftrightarrow z_\gamma = 2{,}58.$$

Следовательно, искомый доверительный интервал имеет вид:

$$44{,}5 - 2{,}58\cdot\frac{4{,}45}{\sqrt{30}} < m < 44{,}5 + 2{,}58\cdot\frac{4{,}45}{\sqrt{30}}$$

т.е. $42{,}41 < m < 46{,}59$.

б) При построении доверительного интервала для математического ожидания $M(X)=m$ с неизвестной дисперсией воспользуемся формулой:

$$\overline{x} - t_\gamma \frac{s}{\sqrt{n}} < m < \overline{x} + t_\gamma \frac{s}{\sqrt{n}},$$

где $\overline{x} = 44{,}5$; $s = 4{,}45$; $n = 30$. Число t_γ находим с помощью табл. П 2.3 (приложение 2) при $n = 30$ и $\gamma = 0{,}99$: $t_\gamma = 2{,}756$.

Следовательно, искомый интервал имеет вид:

$$44{,}5 - 2{,}756 \cdot \frac{4{,}45}{\sqrt{30}} < m < 44{,}5 + 2{,}756 \cdot \frac{4{,}45}{\sqrt{30}}$$

т.е. $42{,}27 < m < 46{,}73$.

в) При построении доверительного интервала среднего квадратического отклонения σ воспользуемся формулой:

$$s(1 - q_\gamma) < \sigma < s(1 + q_\gamma),$$

где s = 4,45 – исправленное выборочное среднее квадратическое отклонение, q_γ= 0,43 – число, которое находим с помощью табл. П 2.4 (приложение 2) при $n = 30$ и $\gamma = 0{,}99$.

Следовательно, искомый доверительный интервал имеет вид:

$$4{,}45 \cdot (1 - 0{,}43) < \sigma < 4{,}45 \cdot (1 + 0{,}43)$$

т.е. $2{,}54 < \sigma < 6{,}36$.

5) Для проверки гипотезы о нормальном распределении генеральной совокупности, из которой взята данная в примере выборка, составим расчетную таблицу, используя интервальный статистический ряд (табл. П 1.4).

Для нахождения чисел $m_{i\,\text{теор}}$ воспользуемся формулой:

$$m_{i\,\text{теор}} = n \cdot p_i\,,$$

где

$$p_i = P(a_i < x_i < a_{i+1}) = \Phi\left(\frac{a_{i+1} - \overline{x}}{s}\right) - \Phi\left(\frac{a_i - \overline{x}}{s}\right),$$

причем

$$a_1 = -\infty, \qquad a_9 = +\infty.$$

Таблица П 1.4

№	$[a_i\ ;\ a_{i+1})$	m_i	$\dfrac{a_i - \overline{x}}{s}$	$\dfrac{a_{i+1} - \overline{x}}{s}$	$m_{i\ \text{теор}}$	$\dfrac{(m_i - m_{i\ \text{теор}})^2}{m_{i\ \text{теор}}}$
1	[36;38)	3	–∞	–1,46	2,16	0,33
2	[38;40)	2	–1,46	–1,01	2,52	0,11
3	[40;42)	2	–1,01	–0,56	3,94	0,96
4	[42;44)	5	–0,56	–0,11	5,06	0,001
5	[44;46)	7	–0,11	0,34	5,31	0,54
6	[46;48)	5	0,34	0,79	4,56	0,04
7	[48;50)	3	0,79	1,24	3,23	0,02
8	[50;52)	3	1,24	+∞	3,22	0,02
∑		30			30	2,02

Следовательно, $\chi^2_{\text{набл}} = 2{,}02$.

Число $\chi^2_{\text{крит}}$ находим из табл. П 2.5 (приложение 2) по уровню значимости $\alpha = 0{,}01$ и числу степеней свободы $k = 8 - 3 = 5$:

$$\chi^2_{\text{крит}} = 15{,}1.$$

Сравним числа: $\chi^2_{\text{набл}} = 2{,}02$ и $\chi^2_{\text{крит}} = 15{,}1$:

Так как $2{,}02 < 15{,}1$, то

$$\chi^2_{\text{набл}} < \chi^2_{\text{крит}}$$

и гипотеза о нормальном распределении генеральной совокупности принимается.

Ответ: 1) табл. П 1.2 и рис. П 1.3;

2) табл. П 1.3 и рис. П 1.4;

3) $\overline{x}$=44,5 , $D_b = 19{,}15$, $s^2 = 19{,}81$, $\sigma_b = 4{,}38$, $s = 4{,}45$;

4) а) $42{,}41 < m < 46{,}59$,

б) $42{,}27 < m < 46{,}73$,

в) $2{,}54 < \sigma < 6{,}36$;

5) генеральная совокупность распределяется нормально.

Задача П 1.7. Найти выборочный коэффициент корреляции и выборочные уравнения линейных регрессий Y на X и X на Y по данным выборки для величин X и Y, сведенным в корреляционную таблицу:

$Y \setminus X$	15	16	18	19	21	m_y
12	2	7				9
20	1	5	12	2	3	23
28			6	10	2	18
m_x	3	12	18	12	5	50

Решение. Для данной в примере выборки объема n=50 вычислим выборочные параметры:

$$\overline{x} = \frac{1}{50}\sum_{i=1}^{5} x_i\, m_i =$$

$$= \frac{1}{50} \cdot (15 \cdot 3 + 16 \cdot 12 + 18 \cdot 18 + 19 \cdot 12 + 21 \cdot 5) = 17{,}88\,,$$

$$\overline{y} = \frac{1}{50}\sum_{i=1}^{3} y_i\, m_i = \frac{1}{50} \cdot (12 \cdot 9 + 20 \cdot 23 + 28 \cdot 18) = 21{,}44\,,$$

$$D_{bx} = \frac{1}{50}\sum_{i=1}^{5} x_i^2\, m_i - \overline{x^2} =$$

$$= \frac{1}{50} \cdot (15^2 \cdot 3 + 16^2 \cdot 12 + 18^2 \cdot 18 + 19^2 \cdot 12 + 21^2 \cdot 5) - 17{,}88^2 = 2{,}63\,,$$

$$D_{by} = \frac{1}{50}\sum_{i=1}^{3} y_i^2\, m_i - \overline{y^2} =$$

$$= \frac{1}{50} \cdot (12^2 \cdot 9 + 20^2 \cdot 23 + 28^2 \cdot 18) - 21{,}44^2 = 32{,}49$$

$$\sigma_{bx} = \sqrt{D_{bx}} = 1{,}62, \quad \sigma_{by} = \sqrt{D_{by}} = 5{,}7\,,$$

$$\bar{\mu}_{xy} = \frac{1}{50}\sum_{i=1}^{5} x_i \cdot \left(\sum_{j=1}^{3} y_j\, m_j\right) - \overline{x} \cdot \overline{y} =$$

$$= \frac{1}{50} \cdot (\,(15 \cdot (12 \cdot 2 + 20 \cdot 1) + 16 \cdot (12 \cdot 7 + 20 \cdot 5) + 18 \cdot (20 \cdot 12 + 28 \cdot 6) + 19 \cdot (20 \cdot 2 + 28 \cdot 10) + 21 \cdot (20 \cdot 3 + 28 \cdot 2)) -$$

$$-17{,}88 \cdot 21{,}44 = 5{,}93\,,$$

$$\bar{r}_{xy} = \frac{\bar{\mu}_{xy}}{\delta_{bx} \cdot \delta_{by}} = \frac{5{,}93}{1{,}62 \cdot 5{,}7} = 0{,}64$$

Запишем выборочное линейное уравнение регрессии *Y* на *X:*

$$y_x - \overline{y} = \frac{\overline{\mu_{xy}}}{\sigma_{bx}^2} \cdot (x - \overline{x}),$$

$$y_x = 2{,}25x - 18{,}87.$$

Запишем выборочное линейное уравнение регрессии *X* на *Y:*

$$x_y - \overline{x} = \frac{\overline{\mu_{xy}}}{\sigma_{by}^2} \cdot (y - \overline{y}),$$

$$x_y = 0{,}18y - 13{,}97.$$

Ответ: выборочный коэффициент корреляции:

$$\bar{r}_{xy} = 0{,}64;$$

выборочное линейное уравнение регрессии *Y* на *X*:

$$y_x = 2{,}25x - 18{,}87;$$

выборочное линейное уравнение регрессии *X* на *Y*:

$$x_y = 0{,}18y - 13{,}97.$$

3. Контрольные вопросы по теории

1. Случайное событие : определение, виды событий, полная группа событий, алгебра событий.
2. Классическое определение вероятности случайного события.
3. Теорема сложения вероятностей.
4. Условная вероятность. Теорема умножения вероятностей.
5. Теорема полной вероятности.
6. Формула Бернулли. Локальная и интегральная теоремы Лапласа.

7. Числовые характеристики дискретной случайной величины.
8. Числовые характеристики непрерывной случайной величины.
9. Функция распределения дискретной и непрерывной случайных величин: определение, свойства.
10. Равномерное распределение непрерывной случайной величины: математическое ожидание, дисперсия, графики плотности распределения вероятностей и функции распределения.
11. Нормальное распределение непрерывной случайной величины: математическое ожидание, дисперсия, графики плотности распределения вероятностей и функции распределения.
12. Интеграл вероятности. Вероятность попадания нормальной случайной величины в заданный промежуток.
13. Биномиальное распределение, его параметры.
14. Распределение Пуассона, его параметры.
15. Неравенство Чебышева.
16. Статистическое определение вероятности. Относительная частота. Выборка.
17. Статистический ряд и интервальный статистический ряд. Полигон и гистограмма относительных частот.
18. Точечные статистические оценки параметров распределения. Их свойства.
19. Интервальные статистические оценки параметров распределения: доверительный интервал для математического ожидания генеральной совокупности при известной и неизвестной дисперсии.
20. Интервальные статистические оценки параметров распределения: доверительный интервал для среднего квадратического отклонения генеральной совокупности.
21. Сравнение двух дисперсий и двух средних нормальных генеральных совокупностей.
22. Критерий согласия Пирсона: проверка гипотезы о нормальном распределении генеральной совокупности.
23. Определение корреляционной зависимости двух генеральных совокупностей. Коэффициент корреляции: определение, свойства, алгоритм вычисления.
24. Функция и линия регрессии случайной величины Y на величину X и наоборот. Уравнения линейных регрессий.
25. Выборочные оценки корреляционных параметров: корреляционный момент и коэффициент корреляции. Проверка гипотезы о значимости выборочного коэффициента корреляции.
26. Ранговая корреляция. Выборочные коэффициенты ранговой корреляции Спирмена и Кендалла. Проверка гипотезы об их значимости.

Приложение 2

Вероятностные таблицы

Таблица П 2.1

Таблица значений функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3652	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127

Продолжение табл. П 2.1

	0	1	2	3	4	5	6	7	8	9
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3.9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Таблица П 2.2

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

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000	0,32	0,1255	0,64	0,2389	0,96	0,3315
0,01	0,0040	0,33	0,1293	0,65	0,2422	0,97	0,3340
0,02	0,0080	0,34	0,1331	0,66	0,2454	0,98	0,3365
0,03	0,0120	0,35	0,1368	0,67	0,2486	0,99	0,3389
0,04	0,0160	0,36	0,1406	0,68	0,2517	1,00	0,3413
0,05	0,0199	0,37	0,1443	0,69	0,2549	1,01	0,3438
0,06	0,0239	0,38	0,1480	0,70	0,2580	1,02	0,3461
0,07	0,0279	0,39	0,1517	0,71	0,2611	1,03	0,3485
0,08	0,0319	0,40	0,1554	0,72	0,2642	1,04	0,3508
0,09	0,0359	0,41	0,1591	0,73	0,2673	1,05	0,3531
0,10	0,0398	0,42	0,1628	0,74	0,2703	1,06	0,3554
0,11	0,0438	0,43	0,1664	0,75	0,2734	1,07	0,3577
0,12	0,0478	0,44	0,1700	0,76	0,2764	1,08	0,3599
0,13	0,0517	0,45	0,1736	0,77	0,2794	1,09	0,3621
0,14	0,0557	0,46	0,1772	0,78	0,2823	1,10	0,3643
0,15	0,0596	0,47	0,1808	0,79	0,2852	1,11	0,3665
0,16	0,0636	0,48	0,1844	0,80	0,2881	1,12	0,3686
0,17	0,0675	0,49	0,1879	0,81	0,2910	1,13	0,3708
0,18	0,0714	0,50	0,1915	0,82	0,2939	1,14	0,3729
0,19	0,0753	0,51	0,1950	0,83	0,2967	1,15	0,3749
0,20	0,0793	0,52	0,1985	0,84	0,2995	1,16	0,3770
0,21	0,0832	0,53	0,2019	0,85	0,3023	1,17	0,3790
0,22	0,0871	0,54	0,2054	0,86	0,3051	1,18	0,3810
0,23	0,0910	0,55	0,2088	0,87	0,3078	1,19	0,3830
0,24	0,0948	0,56	0,2123	0,88	0,3106	1,20	0,3849
0,25	0,0987	0,57	0,2157	0,89	0,3133	1,21	0,3869
0,26	0,1026	0,58	0,2190	0,90	0,3159	1,22	0,3883
0,27	0,1064	0,59	0,2224	0,91	0,3186	1,23	0,3907
0,28	0,1103	0,60	0,2257	0,92	0,3212	1,24	0,3925
0,29	0,1141	0,61	0,2291	0,93	0,3238	1,25	0,3944
0,30	0,1179	0,62	0,2324	0,94	0,3264		
0,31	0,1217	0,63	0,2357	0,95	0,3289		

Продолжение табл. П 2.2

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,26	0,3962	1,59	0,4441	1,92	0,4726	2,50	0,4938
1,27	0,3980	1,60	0,4452	1,93	0,4732	2,51	0,4941
1,28	0,3997	1,61	0,4463	1,94	0,4738	2,54	0,4945
1,29	0,4015	1,62	0,4474	1,95	0,4744	2,56	0,4948
1,30	0,4032	1,63	0,4484	1,96	0,4750	2,58	0,4951
1,31	0,4049	1,64	0,4495	1,97	0,4756	2,60	0,4953
1,32	0,4066	1,65	0,4505	1,98	0,4761	2,62	0,4956
1,33	0,4082	1,66	0,4515	1,99	0,4767	2,64	0,4959
1,34	0,4099	1,67	0,4525	2,00	0,4772	2,66	0,4961
1,35	0,4115	1,68	0,4535	2,02	0,4783	2,68	0,4963
1,36	0,4131	1,69	0,4545	2,04	0,4793	2,70	0,4965
1,37	0,4147	1,70	0,4554	2,06	0,4803	2,72	0,4967
1,38	0,4162	1,71	0,4564	2,08	0,4812	2,74	0,4969
1,39	0,4177	1,72	0,4573	2,10	0,4821	2,76	0,4971
1,40	0,4192	1,73	0,4582	2,12	0,4830	2,78	0,4973
1,41	0,4207	1,74	0,4591	2,14	0,4838	2,80	0,4974
1,42	0,4222	1,75	0,4499	2,16	0,4846	2,82	0,4976
1,43	0,4236	1,76	0,4608	2,18	0,4854	2,84	0,4977
1,44	0,4251	1,77	0,4616	2,20	0,4861	2,86	0,4979
1,45	0,4265	1,78	0,4625	2,22	0,4868	2,88	0,4980
1,46	0,4279	1,79	0,4633	2,24	0,4875	2,90	0,4981
1,47	0,4292	1,80	0,4641	2,26	0,4881	2,92	0,4982
1,48	0,4306	1,81	0,4649	2,28	0,4887	2,94	0,4984
1,49	0,4319	1,82	0,4656	2,30	0,4893	2,96	0,4985
1,50	0,4332	1,83	0,4664	2,32	0,4898	2,98	0,4986
1,51	0,4345	1,84	0,4671	2,34	0,4904	3,00	0,49865
1,52	0,4357	1,85	0,4678	2,36	0,4909	3,20	0,49931
1,53	0,4370	1,86	0,4686	2,38	0,4913	3,40	0,49966
1,54	0,4382	1,87	0,4693	2,40	0,4918	3,60	0,499841
1,55	0,4394	1,88	0,4699	2,42	0,4922	3,80	0,499928
1,56	0,4406	1,89	0,4706	2,44	0,4927	4,00	0,499968
1,57	0,4418	1,90	0,4713	2,46	0,4931	4,50	0,499997
1,58	0,4429	1,91	0,4719	2,48	0,4934	5,00	0,499997

Таблица П 2.3

Таблица значений $t_\gamma = t(\gamma, n)$

n	γ			n	γ		
	0,95	0,99	0,999		0,95	0,99	0,999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,720	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,001	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,68	3,92				

Таблица П 2.4

Таблица значений $q = q(\gamma, n)$

n	γ			n	γ		
	0,95	0,99	0,999		0,95	0,99	0,999
5	1,37	2,67	5,64	20	0,37	0,58	0,88
6	1,09	2,01	3,88	25	0,32	0,49	0,73
7	0,92	1,62	2,98	30	0,28	0,43	0,63
8	0,80	1,38	2,42	35	0,26	0,38	0,56
9	0,71	1,20	2,06	40	0,24	0,35	0,50
10	0,65	1,08	1,80	45	0,22	0,32	0,46
11	0,59	0,98	1,60	50	0,21	0,30	0,43
12	0,55	0,90	1,45	60	0,188	0,269	0,38
13	0,52	0,83	1,33	70	0,174	0,245	0,34
14	0,48	0,78	1,23	80	0,161	0,226	0,31
15	0,46	0,73	1,15	90	0,151	0,211	0,29
16	0,44	0,70	1,07	100	0,143	0,298	0,27
17	0,42	0,66	1,01	150	0,115	0,160	0,211
18	0,40	0,63	0,96	200	0,099	0,136	0,185
19	0,39	0,60	0,92	250	0,089	0,120	0,162

Таблица П 2.5

Критические точки распределения χ^2

Число степеней свободы k	Уровень значимости α					
	0,01	0,025	0,05	0,95	0,975	0,89
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21 ,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26,1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	8,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63
20	37,6	34,2	31,4	10,9	9,59	8,26
21	38,9	35,5	32,7	11,6	10,3	8,90
22	40,3	36,8	33,9	12,3	11,0	9,54
23	41,6	38,1	35,2	13,1	11,7	10,2
24	43,0	39,4	36,4	13,8	12,4	10,9
25	44,3	40,6	37,7	14,6	13,1	11,5
26	45,6	41,9	38,9	15,4	13,8	12,2
27	47,0	43,2	40,1	16,2	14,6	12,9
28	48,3	44,5	41,3	16,9	15,3	13,6
29	49,6	45,7	42,6	17,7	16,0	14,3
30	50,9	47,0	43,8	18,5	16,8	15,0

Таблица П 2.6

Критические точки распределения Стьюдента

Число степеней свободы k	Уровень значимости α (двусторонняя критическая область)					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,95
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,59	3,77
24	1,71	2,06	2,49	2,80	3,47	3,74
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,46	2,76	3,40	3,66
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3.23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,37
∞	1,64	1,96	2,33	2,58	3,09	3,29

Продолжени табл. П 2.6

Число степеней свободы k	Уровень значимости α (двусторонняя критическая область)					
	0,10	0,05	0,02	0,01	0,002	0,001
	0,05	0,025	0,01	0,005	0.\,001	0,0005
	Уровень значимости α (односторонняя критическая область)					

Таблица П 2.7

Критические точки распределения *F* Фишера – Снедекора

(k_1— число степеней свободы большей дисперсии,

k_2—число степеней свободы меньшей дисперсии)

Уровень значимости $\alpha = 0{,}01$												
	k_1											
k_2	1	2	3	4	5	6	7	8	9	10	11	12
1	4052	4999	5403	5625	5764	5889	5928	5981	6022	6056	6082	6106
2	98,49	99,01	90,17	99,25	99,33	99,30	99,34	99,36	99,36	99,40	99,41	99,42
3	34,12	30,81	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,13	27,05
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,54	14,45	14,37
5	16,26	13,27	12,06	11,39	10,97	10,67	10,45	10,27	10,15	10,05	9,96	9,89
6	13,74	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72
7	12,25	9,55	8,45	7,85	7,46	7,19	7,00	6,84	6,71	6,62	6,54	6,47
8	11,26	8,65	7,59	7,01	6,63	6,37	6,19	6,03	5,91	5,82	5,74	5.67
9	10,56	8,02	6,99	6,42	6,06	5,80	5,62	5,47	5,35	5,26	5,18	5,11
10	10,04	7,56	6,55	5,99	5,64	5,39	5,21	5,06	4,95	4,85	4,78	4,71
11	9,86	7,20	6,22	5,67	5,32	5,07	4,88	4,74	4,63	4,54	4,46	4,40
12	9,33	6,93	5,95	5,41	5,06	4,82	4,65	4,50	4,39	4,30	4,22	4,16
13	9,07	6,70	5,74	5,20	4,86	4,62	4,44	4,30	4,19	4,10	4,02	3,96
14	8,86	6,51	5,56	5,03	4,69	4,46	4,28	4,14	4,03	3,94	3,86	3,80
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73	3,67
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,61	3,55
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3.52	3,45

Продолжение табл. П 2.7

Уровень значимости $\alpha = 0{,}05$												
	k_1											
k_2	1	2	3	4	5	6	7	8	9	10	11	12
1	161	200	216	225	230	234	237	239	241	242	243	244
2	18,5	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39	19,40	19,41
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,76	8,74
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93	5,91
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,70	4,68
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,60	3,57
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,31	3,28
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,10	3,07
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,94	2,91
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,82	2,79
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,72	2,69
13	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,63	2,60
14	4,60	3,74	3.34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,56	2,53
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,51	2,48
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,41	2,38

Для заметок