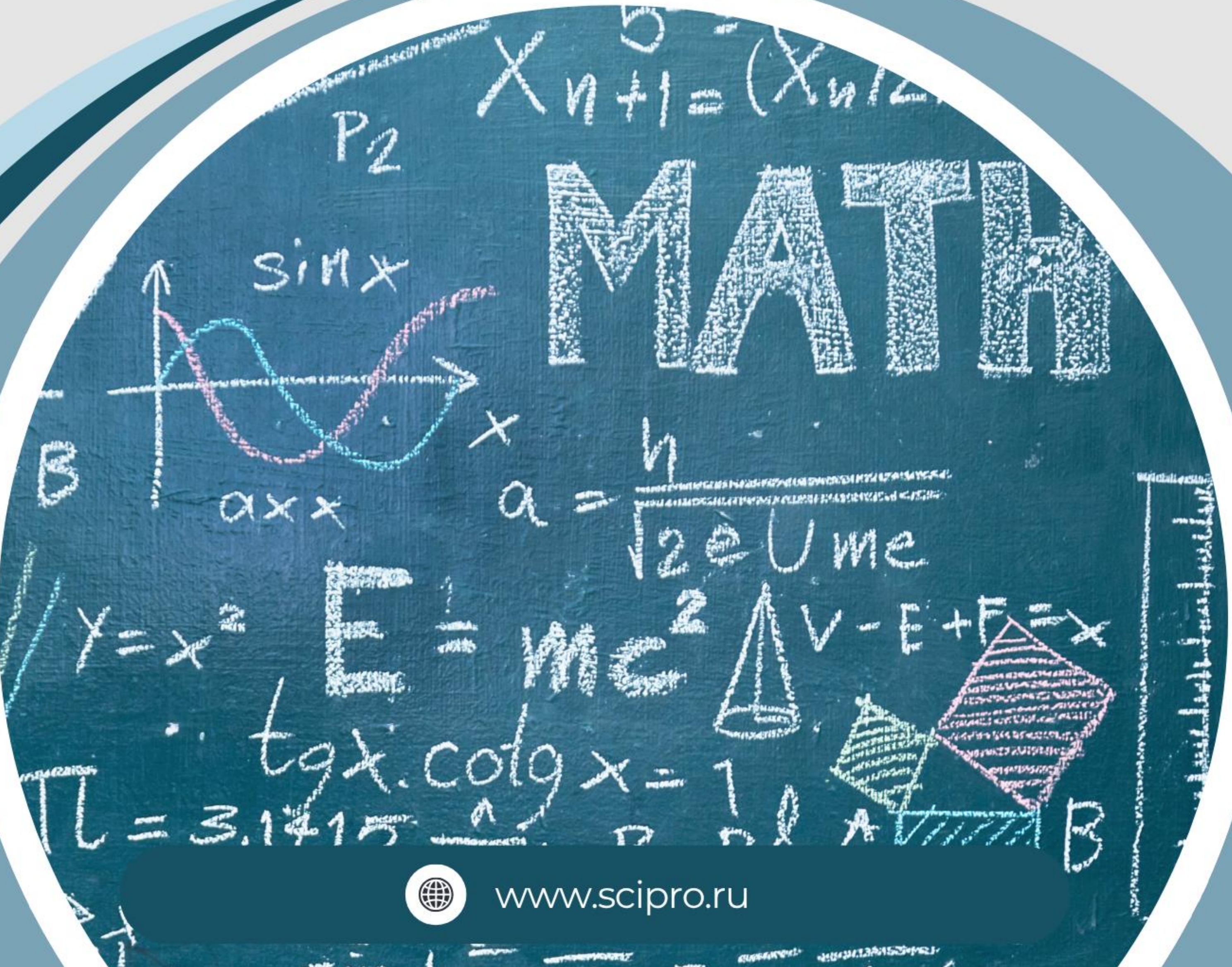
НОО "ПРОФЕССИОНАЛЬНАЯ НАУКА"

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебно-методическое пособие для студентов очной, очно–заочной и заочной формы обучения направления подготовки 38.03.01 «Экономика»

Н. С. Кошевая



www.scipro.ru

СОЧИНСКИЙ ИНСТИТУТ (ФИЛИАЛ)
Федерального государственного автономного образовательного учреждения высшего образования
«РОССИЙСКИЙ УНИВЕРСИТЕТ ДРУЖБЫ НАРОДОВ
ИМ. П. ЛУМУМБЫ»

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебно-методическое пособие для студентов
очной, очно–заочной и заочной формы обучения
направления подготовки 38.03.01 «Экономика»

Составитель Н.С. Кошевая

Сочи
2025

УДК 519.2
ББК 22.17

Главный редактор: Краснова Наталья Александровна – кандидат экономических наук, доцент, руководитель НОО «Профессиональная наука»

Технический редактор: Гусева Ю.О.

Автор-составитель:

Кошевая Наталья Сергеевна - ст. преподаватель кафедры «Математика и информационные технологии». Сочинский институт (филиал)
ФГАОУ ВО «Российский университет дружбы народов им. П. Лумумбы»

Рецензенты:

Воротников В.И. - д. ф.-м. н., проф.
Постников В. В. - к. м. н.

Теория вероятностей и математическая статистика [Электронный ресурс]: учебно-методическое издание – Эл. изд. - Электрон. текстовые дан. (1 файл pdf: 104 с.). - Н.С. Кошевая. 2025. – Режим доступа: http://scipro.ru/conf/mathematical_statistics08_25.pdf. Сист. требования: Adobe Reader; экран 10'.

ISBN 978-5-908003-08-7

В методическом пособии изложены основные понятия теории вероятности и математической статистики.

Здесь представлены разделы: случайные события, случайные величины, многомерные случайные величины, закон больших чисел, элементы математической статистики. Изложение иллюстрируется большим количеством разобранных примеров, в конце каждого раздела предлагаются примеры для самостоятельного разбора. Освоить материал помогают контрольные тесты и вопросы.

Для студентов очной, очно-заочной и заочной формы обучения направления подготовки 38.03.01 «Экономика», а также для студентов, изучающих дисциплину «Теория вероятности и математическая статистика»

ISBN 978-5-908003-08-7

9 785908 003087 >

Содержание

Тема 1. Вероятность случайного события

1.1. *Основные элементы комбинаторики*

Комбинаторика – раздел математики, в котором изучаются различные комбинации (соединения) элементов конечных множеств.

В комбинаторных задачах различают упорядоченные и неупорядоченные конечные множества.

Конечное множество называется упорядоченным, если для двух любых элементов «a» и «b» этого множества имеет место одно из следующих отношений порядка: либо $a \le b$ («a» не превосходит «b»); либо $b \le a$ («b» не превосходит «a»).

Свойства упорядоченного множества:

– рефлективность: любой элемент не превосходит самого себя;

– антисимметричность: если «a» не превосходит «b», а «b» не превосходит «a», то «a» и «b» совпадают;

– транзитивность: если «a» не превосходит «b», а «b» не превосходит «c», то «a» не превосходит «c».

Два множества, составленные из одних и тех же элементов, но с разными отношениями порядка, считаются различными упорядоченными множествами. Поэтому при задании упорядоченного множества через его элементы необходимо также задать их порядок (указать правило, по которому любые два элемента множества можно сравнить).

Многие комбинаторные задачи решаются с помощью двух правил: правило сложения и правило умножения.

Правило сложения: если некоторый объект «a» можно выбрать n способами, а объект «b» можно выбрать m способами, причем первый и второй способы не пересекаются, то любой из объектов «a» или «b» можно выбрать $n+m$ способами.

Правило умножения: если из некоторого конечного множества первый объект «a» можно выбрать n способами, а второй объект «b» можно выбрать m способами, то оба объекта в указанном порядке можно выбрать $n \cdot m$ способами.

Пример 1.1

Сколько различных трехзначных чисел можно составить из цифр: 0, 2, 3,5, 7, если:

а) цифры в числе не повторяются;

б) цифры в числе повторяются.

Решение.

а) Для составления трехзначного числа на первое место можно выбрать только 4 цифры – 2,5,3,7 (ноль не выбирается), на второе место можно поставить одно из оставшихся цифр – их также 4 (так как одно уже выбрали), на третье место можно поставить одно из оставшихся трех цифр.

Таким образом, согласно правилу умножения количество трехзначных чисел без повторения цифр будет равно $4 \cdot 4 \cdot 3$=48.

б) Если цифры в числе могут повторяться, то количество трехзначных чисел будет уже рано $4 \cdot 5 \cdot 5$ = 100.

Ответ. а) 48; б) 100.

Пример 1.2

Сколько чисел, содержащие не менее трех различных цифр, можно составить из цифр: 2,3,4,5,6?

Решение.

По правилу умножения трехзначных чисел можно составить: $5 \cdot 4 \cdot 3$ = 60 способами, четырехзначных: $5 \cdot 4 \cdot 3 \cdot 2$ = 120 способами, пятизначных: $5 \cdot 4 \cdot 3 \cdot 2$ = 120. По правилу сложения, всего можно составить: 60+120+120 = 300 чисел, состоящих не менее, чем из трех различных цифр.

Ответ. 300

В комбинаторике существует две схемы составления комбинаций m элементов из заданного множества: *без повторения*, когда при составлении комбинаций элементы не повторяются и *с повторением*, когда при составлении комбинаций осуществляется повторное использование элементов.

1.2. Комбинаторика без повторения

Рассмотрим множество, состоящее из n элементов, в котором, переставляя элементы данного множества, получим различные комбинации, называемые *перестановкой.*

Для нахождения числа перестановок используют формулу:

$$P(n) = n! \quad (1.1)$$

$n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n \qquad 0!=1 \qquad 1! =1$

В перестановке важен порядок.

Пример 1.3

Сколько различных пятизначных чисел можно составить из цифр 1,2,3,4,5?

Решение.

Для составления пятизначного числа поочередно отбирают цифры. Первую цифру можно отобрать и поставить на первое место пятью способами, тогда отобрать и поставить цифру на втрое место можно четырьмя способами, на третье место соответственно – тремя способами, на четвертое – двумя и пятое – одним способом.

Используя правило умножения, имеем: P(5) = $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$=5!=120.

Итак, 120 различных пятизначных чисел можно составить из цифр 1,2,3,4,5.

Ответ. 120

Размещением из n элементов по k элементов называется любое упорядоченное подмножество данного множества, содержащее k элементов.

Два размещения различны, если они отличаются друг от друга либо составом элементов, либо порядком их расположения.

Число размещений из n элементов по k элементов обозначается A_n^k и вычисляется по формуле:

$$A_n^k = \frac{n!}{(n-k)!} \quad (1.2)$$

Пример 1.4

Сколько различных четырехзначных чисел можно составить из чисел 1, 2, 3, …, 9, если все цифры в четырехзначных числах различны?

Решение.

Для формирования каждого четырехзначного числа выбираем четыре цифры из девяти, поэтому существует:

$A_9^4 = \frac{9!}{(9-4)!} = \frac{9!}{5!} = 6 \cdot 7 \cdot 8 \cdot 9 = 3024$ различных чисел.

Ответ. 3 024

Сочетанием из n элементов по k элементов называется любое подмножество данного множества, содержащее k элементов.

Любые два сочетания отличаются друг от друга хотя бы одним элементом

Число сочетаний из n элементов по k элементов обозначается C_n^k и вычисляется по формуле:

$$C_n^k = \frac{n!}{k! \cdot (n-k)!} \quad (1.3)$$

Свойства сочетаний:

1. $C_n^k = C_n^{n-k}$
2. $C_n^0 + C_n^1 + ... + C_n^n = 2^n$
3. $C_n^k = C_{n-1}^k + C_{n-1}^{k-1}$
4. $C_n^0 = C_n^n = 1$

Пример 1.5

Если множество содержит 10 элементов, то, сколько оно содержит трехэлементных подмножеств?

Решение.

Поскольку множество неупорядоченно и выбираются любые три элемента из десяти, то поэтому всего имеется

$C_{10}^3 = \frac{10!}{3! \cdot (10-3)!} = \frac{10!}{3! \cdot 7!} = \frac{8 \cdot 9 \cdot 10}{1 \cdot 2 \cdot 3} = 4 \cdot 3 \cdot 10 = 120$ различных подмножеств.

Ответ. 120

1.3. Комбинаторика с повторением

Пусть в множестве из n элементов есть k различных типов элементов, при этом первый тип элементов повторяется n_1 раз, второй – n_2 раз, …, k-й тип – n_k раз, причем $n_1+n_2+\ldots+n_k = n$. Тогда перестановки элементов данного множества представляют собой *перестановки с повторениями.*

Число *перестановок с повторениями* обозначаются $P_n(n_1, n_{2,\ldots,}n_k)$ и вычисляется по формуле:

$$P_n(n_1, n_2, \ldots, n_k) = \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_k!} \qquad (1.4)$$

Пример 1.6

Сколько различных буквенных комбинаций можно составить, переставляя буквы в слове «ИНФЛЯЦИЯ»?

Решение.

Вообще из восьми букв можно составить P_8 =8! = 40320 различных восьмибуквенных комбинаций. В слове «инфляция» две одинаковые буквы «Л» и «И», перестановка которых не поменяет комбинацию. Поэтому число перестановок с повторением меньше числа перестановок без повторений во столько раз, сколько можно переставлять повторяющиеся буквы. Отсюда количество различных комбинаций при перестановке букв в слове «ИНФЛЯЦИЯ» будет равно:

$$P_8(2,2) = \frac{P_8}{P_2 \cdot P_2} = \frac{8!}{2! \cdot 2!} = \frac{40320}{4} = 10080$$

Ответ. 10080

Если при упорядоченной выборке k элементов из n элементы возвращаются обратно, то полученные выборки представляют собой *размещения с повторениями.*

Число всех размещений с повторениями из n элементов по k обозначается $\left(A_n^k\right)^*$ и вычисляется по формуле:

$$\left(A_n^k\right)^* = n^k \qquad (1.5)$$

Пример 1.7

Для того, чтобы попасть на свое рабочее место пять сотрудников сели в лифт девятиэтажного офиса. Сколько существует комбинаций выхода сотрудников на своем этаже.

Решение.

Каждый из пяти сотрудников может выйти на любом из восьми этажей в установленном порядке, причем на один этаж может выйти несколько сотрудников от одного до пяти, таким образом, общее число выходов на свой этаж сотрудников равно:

$$\left(A_n^k\right)^* = 8^5 = 32\ 768.$$

Ответ. 32 768

Если при упорядоченной выборке k элементов из n элементы возвращаются обратно без последующего упорядочивания, то полученные выборки представляют собой *сочетания с повторениями.*

Число всех сочетаний с повторениями из n элементов по k обозначается $\left(C_n^k\right)^*$ и вычисляется по формуле:

$$\left(C_n^k\right)^* = C_{n+k-1}^k \qquad (1.6)$$

Пример 1.8

В магазине имеется три вида холодильников. Сколькими способами можно приобрести семь холодильников.

Решение.

Поскольку имеет 3 вида холодильников, а приобрести нужно 7, то число возможных наборов равно:

$$\left(C_3^7\right)^* = C_{3+7-1}^7 = C_9^7 = \frac{9!}{7!\cdot(9-7)!} = \frac{9!}{7!\cdot 2!} = \frac{8\cdot 9}{1\cdot 2} = 4\cdot 9 = 36$$

Ответ. 36

Итоговую сводку формул представляем в таблице №1

Таблица №1

Формулы комбинаторики

Комбинаторика	Перестановка	Размещение	Сочетание
с повторением	*P (n) = n!*	$A_n^k = \dfrac{n!}{(n-k)!}$	$C_n^k = \dfrac{n!}{k!\cdot(n-k)!}$
без повторения	$P_n^*(n_1, n_2, ..., n_k) = \dfrac{n!}{n_1!\cdot n_2!\cdot \ldots \cdot n_k!}$	$\left(A_n^k\right)^* = n^k$	$\left(C_n^k\right)^* = C_{n+k-1}^k$

1.4. Вероятность случайного события.

Событие, которое в ходе испытания или эксперимента может наступить, а может не наступить, называется *случайным.*

Виды случайных событий:

- *совместные события* – появление одного события не исключает появления другого;

- *несовместные события* – появление одного события исключает появления другого, в одном и том же испытании;

- *противоположные события* – это два несовместных событий, одно из которых обязательно происходит;

- *равновозможные события* – в ходе эксперимента в одинаковых условиях ни одно событие не имеет преимущества перед другими событиями;

- *достоверное событие* – в ходе эксперимента имеет единственный возможный исход;

- *невозможное событие* – в ходе эксперимента заведомо не может произойти.

Классическое определение вероятности наступления случайного события

Классическое определение вероятность события определяется формулой:

$$P(A) = \frac{m}{n}, \tag{1.7}$$

где m – число исходов проводимого опыта, благоприятствующих появлению события A; n – общее число возможных исходов.

Свойство вероятности:

1. Вероятность невозможного события равна нулю.
2. Вероятность достоверного события равна единице.
3. Вероятность случайного события есть число положительное, заключенное в промежутке от нуля до единицы ($0<P(x)<1$)

Пример 1.9

В лотерее разыгрывается 100 билетов, 10 из которых являются выигрышными. Некто покупает один билет. Какова вероятность покупки билета не выигрышного.

Решение.

Обозначим событие А, куплен не выигрышный билет. Число не выигрышных билетов равно: 100 – 10 =90.

Вероятность того, что будет куплен не выигрышный билет, равна: $P(A) = \frac{90}{100} = 0{,}9$

Ответ. 0,9

1.5. Теоремы сложения и умножения вероятностей

Определение 1. Сумма двух событий A и B – это такое событие $A+B$, которое состоит в том, что произошло хотя бы одно из этих событий.

Определение 2. Произведение событий A и B – это такое событие $A \cdot B$, состоящее в том, что эти события произошли совместно: и A и B.

Теорема 1. Вероятность суммы двух несовместных событий равна сумме вероятностей этих событий:

$$P(A + B) = P(A) + P(B) \tag{1.8}$$

Следствие 1. Вероятность суммы попарно несовместных событий A_1, A_2, …, A_n равна сумме вероятностей этих событий:

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

Следствие 2. Сумма вероятностей событий, образующих полную группу (то есть, когда эти события попарно несовместны, но в результате испытания одно из них произойдет обязательно), равна 1.

Следствие 3. Сумма вероятностей противоположных событий равна единицы:

$P(A) + P(\overline{A}) = 1$, где $\overline{A}$ – событие противоположное событию A.

Пример 1.10

Очередь на оформление кредитов содержит 10 человек, четырем из которых одобрили кредиты под 12% годовых, трем – под 16 % годовых, а остальным под 19%. Специалист обслуживает клиента. Найти вероятность того, что обслуживается клиент с одобренным кредитом не более чем 16% годовых.

Решение.

Если обслуживается клиент с одобренным кредитом не более чем 16% годовых, то это может быть клиент либо с одобренным кредитом под 12% годовых, либо под 16%.

Обозначим два события: А – «специалист обслуживает клиента, с одобренным кредитом под 12% годовых», В -«специалист обслуживает клиента, с одобренным кредитом под 16% годовых». Вероятность того, что клиент с одобренным кредитом под 12% годовых» равна 0,4, вероятность того, что клиент с одобренным кредитом под 16% годовых» равна 0,3. События А и В несовместные, следовательно, искомую вероятность найдем по формуле (8).

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

Ответ. 0,7

Теорема 2. Вероятность произведения двух совместных независимых событий равна произведению вероятностей этих событий:

$$P(A \cdot B) = P(A) \cdot P(B) \tag{1.9}$$

Пример 1.11

Вероятность заключить экономический договор с фирмой А равна 0,6, фирмой В – 0,7. Найти вероятность того, что на экономическом форуме предприятие заключило договор с двумя фирмами.

Решение.

Обозначим событие D – предприятие заключило договор с двумя фирмами. Заключение договора с различными фирмами события независимые и совместные, поэтому искомую вероятность мы найдем по формуле (9):

$$P(D) = P(A \cdot B) = P(A)\ P(B) \Rightarrow \mathrm{P(D)} = 0{,}6 \cdot 0{,}7 = 0{,}42$$

Ответ. 0,42

Теорема 3. Вероятность произведения двух совместных зависимых событий равна произведению вероятностей одного из них на условную вероятность второго:

$$P(A\cdot B) = P(A)\cdot P(B/A) = P(B)\cdot P\,(A/B) \qquad (1.10)$$

где условная вероятность $P(B/A)$ – вероятность события B при условии, что A произошло.

Следствие 1. Если события A_1, A_2, …, A_n совместны и зависимы, то
$P\,(A_1\cdot A_2\,\ldots\cdot A_n) = P(A_1)\,\cdot p(A_2/A_1)\;P(A_3/A_1A_2)\,\ldots\cdot P\,(A_n\,/A_1A_2\ldots A_{n-1})$.

Следствие 2. Если события A_1, A_2, …, A_n независимы в совокупности, то
$P\,(A_1\cdot A_2\,\ldots\cdot A_n) = P(A_1)\,\cdot \mathrm{P}(A_2)\;P(A_3)\,\ldots\cdot P(A_n)$.

Пример 1.12

В городе N производится аудиторская проверка деятельности отелей, среди которых 12 отелей в муниципальной собственности, а остальные в частной собственности. Наудачу выбирают для проверки два отеля. Найти вероятность того, что проверили два отеля принадлежащих индивидуальным предпринимателям, если в городе всего 25 отелей.

Решение.

Обозначим событие A – проверили два отеля, принадлежащих индивидуальным предпринимателям.

Обозначим событие F – первым проверили отель, принадлежащий индивидуальному предпринимателю. Вероятность события F равна P(F) = 13:25 = 0,52.

Событие D – вторым проверили отель, принадлежащий индивидуальному предпринимателю. Тогда вероятность события D будет условной, поскольку уже произошло событие F. Поэтому вероятность события D равна P(D) = 12:24 = 0,5.

Поскольку события F и D зависимые события, то искомую вероятность события A найдем по формуле (10): $\mathrm{P(A)} = P(F\cdot D) = P(F)\cdot P(D/F) \Rightarrow P(A) = 0{,}52\cdot 0{,}5 = 0{,}26$

Ответ. 0,26

Теорема 4. Вероятность суммы двух совместных событий равна сумме вероятностей этих событий без вероятности их совместного появления:

$$P(A+B) = P(A)+P(B) - P(A\cdot B) \qquad (1.11)$$

Следствие 1. Если события A и B совместны и независимы, то
$P\,(A + B) = P(A) + P(B) - P(A)\cdot P(B)$.

Следствие 2. Если события A и B совместны и зависимы, то
$P\,(A + B) = P(A) + P(B) - P(A)\cdot P(B/A) = P(A) + P(B) - P(B)\;P(A/B)$.

Пример 1.13

На предприятие поступают заявки от нескольких торговых пунктов. Вероятности поступления заявки от пунктов A и B соответственно равны: 0,5 и

0, 4. Найти вероятность поступления заявок от пункта А или пункта В, считая события поступления заявок от этих пунктов независимыми, но совместными.

Решение.

Обозначим событие S – заявка поступила с пункта А и событие F – заявка поступила с пункта В. По условию задачи $P(S)= 0{,}5$, $P(F)= 0{,}4$. Так как события S и F независимыми, но совместными, то искомую вероятность наступления события «поступления заявок от пункта А или пункта В» найдем по формуле: $P(S+F) = P(S)+P(F) – P(SF) \Rightarrow P(S+F)=0{,}5+0{,}4 – 0{,}5\cdot 0{,}4 = 0{,}9 – 0{,}2 = 0{,}7$

Ответ. 0,7

1.6. Вероятность появления хотя бы одного события

Пусть события $A_1, A_2, \ldots, A_n$ независимы в совокупности, причем $P(A_1) =p_1$, $P(A_2) = p_2, \ldots, P(A_n) =p_n$. Пусть в результате испытания могут наступить либо все события, либо часть событий, либо ни одно из них.

Вероятность наступления события А, состоящего в появлении хотя бы одного их событий $A_1, A_2, \ldots, A_n$, независимых в совокупности, равна разности между единицей и произведением вероятностей противоположных событий: $\overline{A}_1, \overline{A}_2, \ldots, \overline{A}_n$:

$$P(A)= 1 – q_1\cdot q_2\cdot\ldots\cdot q_n \qquad (1.12)$$

где $q_1=1 – p_1$; $\quad q_2 = 1 – p_2; \ldots\ q_n= 1- p_n$

Пример 1.14

Предприятие обеспечивает регулярный выпуск продукции при безотказной поставке комплектующих от двух смежников. Вероятность отказа в поставке продукции от первого из смежников равна 0,05, от второго – 0,08. Найти вероятность сбоя в работе предприятия.

Решение.

Предприятие даст сбой в работе, если хотя бы один из смежников откажет в поставке комплектующих.

Искомая вероятность: $P(A) = 1 – (1 – 0{,}05)\cdot(1 – 0{,}08) = 1 – 0{,}874 = 0{,}126$

Ответ. 0,126

1.7. Формула полной вероятности. Формула Байеса

Пусть гипотезы B_1, B_2, …, B_n образуют полную группу событий и попарно несовместны, а событие A может наступить лишь в результате осуществления одной из гипотез $B_i (i=\overline{1,n})$. Тогда вероятность события A равна сумме произведений вероятностей каждой из гипотез на соответствующую условную вероятность события *A*:

$$P(A) = P(B_1) \cdot P_{B_1}(A) + P(B_2) \cdot P_{B_2}(A) + \ldots + P(B_n) \cdot P_{B_n}(A) \qquad (1.13)$$

где $P(B_1) + P(B_2) + \ldots + P(B_n) = 1$.

Допустим, произведено испытание, в результате которого появилось событие *A*. Вероятности гипотез B_i после опыта, т.е. условные вероятности: $p(A/B_1)$, $p(A/B_2)$, …, $p(A/B_n)$, вычисляются по формуле Байеса:

$$P_A(B_i) = \frac{P(B_i) \cdot P_{B_i}(A)}{P(A)}, \quad i = \overline{1,\ldots n} \qquad (1.14)$$

Эта формула позволяет оценить вероятности гипотез после того, как становится известным результат испытания, в итоге которого появилось событие *A*.

Пример 1.14

В ящике содержится 12 деталей завода №1, 20 деталей завода №2, 18 деталей завода №3. Завод №1 выпускает 90% продукции отличного качества, завод №2 – 60%, а завод №3 – 80% продукции отличного качества. Найти:

а) вероятность того, что извлеченная наудачу из ящика деталь оказалась отличного качества;

б) вероятность того, что она изготовлена на заводе №2.

Решение.

а) Обозначим событие *A* = {наудачу взятая из ящика деталь окажется отличного качества}. Возможны следующие гипотезы: B_1 = {деталь изготовлена на *i*-м заводе}, $i = \overline{1,3}$. Гипотезы B_1, B_2, B_3 попарно несовместны и образуют полную группу событий. Поскольку в ящике всего 22+20+18 = 50 деталей, то по классической формуле вероятности:

$$P(B_1) = \frac{12}{50} = 0{,}24;\ P(B_2) = \frac{20}{50} = 0{,}4;\ P(B_3) = \frac{18}{50} = 0{,}36$$

Условные вероятности того, что деталь окажется отличного качества, если она изготовлена на *i*-м заводе ($i = \overline{1,3}$) по условию задачи равны: $P(A/B_1) = 0{,}9$; $P(A/B_2) = 0{,}6$; $P(A/B_3) = 0{,}8$.

По формуле полной вероятности (2): $P(A) = 0{,}24 \cdot 0{,}9 + 0{,}4 \cdot 0{,}6 + 0{,}36 \cdot 0{,}8 = =0{,}744$.

б) По формуле Байеса (13) найдем вероятность того, что извлеченная деталь изготовлена на заводе №2:

$$P_A(B_2) = \frac{P(B_2) \cdot P_{B_2}(A)}{P(A)} = \frac{0{,}4 \cdot 0{,}6}{0{,}744} = 0{,}323$$

Ответ. а) 0,744; б) 0,323

1.8. Схема Бернулли

Несколько испытаний называются независимыми, если вероятность того или иного исхода в любом из этих испытаний не зависит от исхода других испытаний.

Схема Бернулли: производится в n независимых испытаний, в каждом из которых с одной и той же вероятностью p наступает некоторое событие A и, следовательно, с вероятностью

$q = 1 - p$ наступает событие $\overline{A}$, противоположное A.

Обозначим через $P_n(m)$ вероятность того, что в n испытаниях схемы Бернулли событие A наступит m раз ($m = \overline{1,n}$), которое определяется по формуле:

$$P_n(m) = C_n^m \cdot p^n \cdot q^{n-m} \qquad (1.15)$$

Вероятность того, что при n испытаниях в условиях схемы Бернулли событие A наступит не менее m_1 раз и не более m_2 раз, то есть $m_1 \le m \le m_2$.

$$P_n(m_1 \le m \le m_2) = P_n(m_1) + P_n(m_1+1) + \ldots + P_n(m_2)$$

Справедливы также следующие формулы:

$$P_n(m > m_1) = P_n(m_1+1) + P_n(m_1+2) + \ldots + P_n(n)$$
$$P_n(m < m_1) = P_n(0) + P_n(1) + \ldots + P_n(m_1-2) + P_n(m_1-1)$$
$$P_n(m \ge m_1) = P_n(m_1) + P_n(m_1+1) + \ldots + P_n(n)$$
$$P_n(m \le m_1) = P_n(0) + P_n(1) + \ldots + P_n(m_1-1) + P_n(m_1)$$
$$P_n(0) + P_n(1) + \ldots + P_n(n) = 1$$

Наивероятнейшее число m_0 появления события A в n испытаниях в условиях схемы Бернулли определяется из:

$$np - q \le m_0 \le np + p, \text{ где } m_0 \text{ – целое число.} \qquad (1.16)$$

Пример 1.15

Четыре покупателя приехали на оптовый склад. Вероятность того, что каждый покупатель приобретет товар А, равна 0,4. Найти вероятность того, что:

а) товар А приобретет ровно три покупателя;

б) товар приобретут не более двух покупателей.

Решение.

Вероятность покупки товара А каждым покупателем равна $p=0,4$, вероятность того. что товар не будет куплен $q= 1 - 0,4=0,6$.

а) Вероятность того, что из четырех покупателей товар приобретут ровно три покупателя, найдем по формуле Бернулли:

$$P_4(3) = C_4^3 \cdot p^3 \cdot q^{4-3} = \frac{4!}{3! \cdot 1!} \cdot 0{,}4^3 \cdot 0{,}6 = 4 \cdot 0{,}064 \cdot 0{,}6 = 0{,}1536.$$

б) Вероятность того, что товар приобретут не более двух покупателей, найдем по формуле:

$$\mathrm{P}_4(0) + \mathrm{P}_4(1) + \mathrm{P}_4(2) = \mathrm{C}_4^0 \cdot \mathrm{p}^0 \cdot \mathrm{q}^{4-0} + \mathrm{C}_4^1 \cdot \mathrm{p} \cdot \mathrm{q}^{4-1} + \mathrm{C}_4^2 \cdot \mathrm{p}^2 \cdot \mathrm{q}^2 =$$

$$= \frac{4!}{0! \cdot 4!} \cdot 0,4^0 \cdot 0,6^4 + \frac{4!}{1! \cdot 3!} \cdot 0{,}4 \cdot 0,6^3 + \frac{4!}{2! \cdot 2!} \cdot 0,4^2 \cdot 0,6^2 =$$

$$= 0{,}1296 + 4 \cdot 0{,}4 \cdot 0{,}126 + 6 \cdot 0{,}16 \cdot 0{,}36 = 0{,}1296 + 0{,}3456 + 0{,}3456 =$$

$$= 0{,}8208$$

Ответ. а) 0,1536; б) 0,8208

Пример 1.16

Отдел технического контроля проверяет партию из 10 деталей. Вероятность того, что деталь стандарта, равна 0,75. Найти наивероятнейшее число деталей, которые будут признаны стандартными.

Пример.

По условию *n=10; p=0,75; q=0,25*. Найдем наивероятнейшее число деталей, которые будут признаны стандартными из неравенства:

$np - q \le m_0 \le np + p$, $\Rightarrow$, $10 \cdot 0{,}75 - 0{,}25 \le m_0 \le 10 \cdot 0{,}75 + 0{,}75$, $\Rightarrow$,

$7{,}25 \le m_0 \le 8{,}25$.

Так как число m_0 целое, а между 7,25 и 8,25 заключено только одно целое число 8, то искомое наивероятнейшее число деталей, признанных стандартными, будет 8.

Ответ. 8

1.9. Локальная и интегральная теоремы Лапласа

Локальная теорема Лапласа. Вероятность того, что в n независимых испытаниях, в каждом из которых вероятность появления события равна p, событие наступит ровно k раз (безразлично, в какой последовательности), приближенно равна:

$$P_n(k) = \frac{1}{\sqrt{npq}} \cdot \varphi(x).$$

где
(1.17)

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot \ell^{-\frac{x^2}{2}}, \qquad x = \frac{k-np}{\sqrt{npq}}$$

Функция $\varphi(x)$ является четной $\varphi(-x) = \varphi(x)$ и определяется по таблице (Приложение 1).

Пример 1. 17

Найти вероятность того, что событие А наступит ровно 70 раз в 243 испытаниях, если вероятность появления этого события в каждом испытании равна 0,25.

Решение.

По условию, *n=243; k = 70; p =0,25; q= 0,75.* Так как *n=243* – достаточно большое число, воспользуемся локальной **теоремой Лапласа**:

$$P_{243}(70) = \frac{1}{\sqrt{0,25 \cdot 0,75 \cdot 243}} \cdot \varphi(x).$$

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

По таблице Приложения 1 найдем φ(1,37) =0,1561.

Искомая вероятность

$$P_{243}(70) = \frac{1}{\sqrt{0,25 \cdot 0,75 \cdot 243}} \cdot 0,1561 = \frac{0,1561}{6,75} = 0,0231.$$

Ответ. 0,0231

Пример 1.18

Найти вероятность того, что событие А наступит ровно 1400 раз в 2400 испытаниях, если вероятность появления этого события в каждом испытании равна 0,6.

Решение.

По условию, *n=2400; k = 1400; p =0,6; q= 0,4*. Так как *n=2400* – достаточно большое число, воспользуемся локальной **теоремой Лапласа**:

$$P_{2400}(1400) = \frac{1}{\sqrt{0,6 \cdot 0,4 \cdot 2400}} \cdot \varphi(x).$$

Найдем x: $x = \frac{k - np}{\sqrt{npq}} = \frac{1400 - 2400 \cdot 0,6}{\sqrt{2400 \cdot 0,6 \cdot 0,4}} = -\frac{40}{24} = -1,67$

По таблице Приложения 1 найдем φ(-1,67) = φ(1,67) =0,0989.

Искомая вероятность

$$P_{2400}(1400) = \frac{1}{\sqrt{0,6 \cdot 0,4 \cdot 2400}} \cdot 0,0989 = \frac{0,0989}{24} = 0,0041.$$

Ответ. 0,0041

Интегральная теорема Лапласа. Вероятность того, что в n независимых испытаниях, в каждом из которых вероятность появления события равна p, событие наступит не менее k_1 раз и не более k_2 раз, приближенно равна:

$$P(k_1; k_2) = \Phi(x'') - \Phi(x').$$

где

(1.18)

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^x \ell^{-\frac{x^2}{2}} dx, \qquad x' = \frac{k_1 - np}{\sqrt{npq}}; \quad x'' = \frac{k_2 - np}{\sqrt{npq}}$$

Значение функции определяется по таблице в Приложение 2, учитывая ее нечетность Ф(-x) = – Ф(x).

Пример. 1.19

Вероятность появления события в каждом из 100 независимых испытаний постоянна и равна p=0,8. Найти вероятность того, что событие появится:

а) не менее 75 раз и не более 90 раз,

б) не менее 75 раз,

в) не более 74 раз.

Решение.

Воспользуемся интегральной теоремой Лапласа: $P(k_1; k_2) = \Phi(x'') - \Phi(x')$, где Ф(x) – функция Лапласа.

$$x' = \frac{k_1 - np}{\sqrt{npq}}; \qquad x'' = \frac{k_2 - np}{\sqrt{npq}}.$$

а) По условию $n=100$; $k_1 = 75$; $k_2 = 90$; $p = 0,8$; $q = 0,2$. Определим x'' и x':

$$x' = \frac{k_1 - np}{\sqrt{npq}} = \frac{75 - 100 \cdot 0,8}{\sqrt{100 \cdot 0,8 \cdot 0,2}} = -1,25;$$

$$x'' = \frac{k_2 - np}{\sqrt{npq}} = \frac{90 - 100 \cdot 0,8}{\sqrt{100 \cdot 0,8 \cdot 0,2}} = 2,5$$

Учитывая, что интегральная функция Лапласа нечетная, получим:

P_{100} (75; 90) = Ф (2,5) – Ф (-1, 25) = Ф (2,5) + Ф (1,25)

Откуда Ф (2,5) = 0,4938 Ф (1,25) = 0,3944.

P_{100} (75; 90) = 0,4938 + 0,3944 = 0, 8882.

б) Требование, чтобы событие появилось не менее 75 раз, означает, что число появлений события может наступить ровно 75, либо 76, …, либо 100. Таким образом, в рассматриваемом случае следует принять $k_1 = 75; k_2 = 100.$ Тогда:

$$x' = \frac{k_1 - np}{\sqrt{npq}} = \frac{75 - 100 \cdot 0,8}{\sqrt{100 \cdot 0,8 \cdot 0,2}} = -1,25;$$

$$x'' = \frac{k_2 - np}{\sqrt{npq}} = \frac{100 - 100 \cdot 0,8}{\sqrt{100 \cdot 0,8 \cdot 0,2}} = 5$$

Учитывая, что интегральная функция Лапласа нечетная, получим:

P_{100} (75; 100) = Ф (5) – Ф (–1, 25) = Ф (5) + Ф (1,25)

Откуда Ф (5) = 0,5 Ф (1,25) = 0,3944.

P_{100} (75; 90) = 0,5 + 0,3944 = 0,8944.

в) События «А- появилось не менее 75 раз» и «А – появилось не более 74 раз» противоположны, поэтому сумма этих вероятностей равна 1. Следовательно, искомая вероятность равна: P_{100} (0; 74) = 1 – 0, 8944 = 0,1056.

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

Дополнительные задачи

Пример 1.20

Студент разыскивает нужную ему формулу в трех справочниках. Вероятность того, что формула содержится в первом, втором, третьем справочниках равна, соответственно, 0,6; 0,7; 0,8. Найти вероятность того, что формула содержится:

а) только в одном справочнике;

б) хотя бы в двух справочниках;

в) во всех трех справочниках;

г) хотя бы в одном справочнике.

Решение.

Введем обозначения событий: B_i = {нужная студенту формула находится в i-м справочнике}, i = 1, 2, 3.

По условию задачи $p(B_1)=0,6;\ p(B_2)=0,7;\ p(B_3)=0,8.$

Так как B_i – независимые события, то и противоположные им события $\overline{B_i}$ – независимы:

$$p(\overline{B_1})=1-p(B_1)=0,4;\ \ p(\overline{B_2})=1-p(B_2)=0,3;\ \ p(\overline{B_3})=1-p(B_3)=0,2.$$

а) Событие A ={формула содержится только в одном справочнике} эквивалентно событию {формула содержится только в первом справочнике и не содержится во 2-м и 3-м; или формула содержится только во втором справочнике, но не содержится в 1-м и 3-м; или формула содержится только в третьем справочнике, но не содержится в 1-м и 2-м}. Согласно определениям сложения и умножения событий алгеброй события A является:

$$A=B_1\cdot\overline{B_2}\cdot\overline{B_3}+\overline{B_1}\cdot B_2\cdot\overline{B_3}+\overline{B_1}\cdot\overline{B_2}\cdot B_3.$$

События – слагаемые в правой части последнего равенства несовместны, а события-сомножители – независимы.

По теоремам сложения и умножения вероятностей имеем:

$$p(A)=p\left(B_1\cdot\overline{B_2}\cdot\overline{B_3}\right)+p\left(\overline{B_1}\cdot B_2\cdot\overline{B_3}\right)+p\left(\overline{B_1}\cdot\overline{B_2}\cdot B_3\right)=$$

$$=p(B_1)\cdot p\left(\overline{B_2}\right)\cdot p\left(\overline{B_3}\right)+p\left(\overline{B_1}\right)\cdot p(B_2)\cdot p\left(\overline{B_3}\right)+p\left(\overline{B_1}\right)\cdot p\left(\overline{B_2}\right)\cdot p(B_3)=$$

$$=0,6\cdot 0,3\cdot 0,2+0,4\cdot 0,7\cdot 0,2+0,4\cdot 0,3\cdot 0,8=0,188.$$

б) Событие B ={формула содержится хотя бы в двух справочниках} эквивалентно событию {формула содержится только в 1-м и во 2-м; или в 1-м и 3-м; или во 2-м и в 3-м; или в 1-м, 2-м, 3-м справочниках}, то есть

$$B=B_1\cdot B_2\cdot\overline{B_3}+B_1\cdot\overline{B_2}\cdot B_3+\overline{B_1}\cdot B_2\cdot B_3+B_1\cdot B_2\cdot B_3.$$

Аналогично, по теоремам сложения и умножения вероятностей:

$$p(B)=p(B_1)\cdot p(B_2)\cdot p\left(\overline{B_3}\right)+p(B_1)\cdot p\left(\overline{B_2}\right)\cdot p(B_3)+p\left(\overline{B_1}\right)\cdot p(B_2)\cdot p(B_3)+$$

$$+p(B_1)\cdot p(B_2)\cdot p(B_3)=0,6\cdot 0,7\cdot 0,2+0,6\cdot 0,3\cdot 0,8+0,4\cdot 0,7\cdot 0,8+$$

$$+0,6\cdot 0,7\cdot 0,8=0,788.$$

в) Событие C ={формула содержится во всех трех справочниках} эквивалентно событию {формула содержится и в 1-м, и во 2-м, и в 3-м справочниках}; то есть $C=B_1\cdot B_2\cdot B_3.$

По теореме умножения независимых событий

$$p(C) = p(B_1) \cdot p(B_2) \cdot p(B_3) = 0,6 \cdot 0,7 \cdot 0,8 = 0,336.$$

г) Событию D = {формула содержится хотя бы в одном справочнике} противоположно событие $\overline{D}$ = {формулы нет ни в одном справочнике}, то есть $\overline{D} = \overline{B_1} \cdot \overline{B_2} \cdot \overline{B_3}$.

Тогда по теореме умножения независимых событий

$$p(\overline{D}) = p(\overline{B_1}) \cdot p(\overline{B_2}) \cdot p(\overline{B_3}) = 0,4 \cdot 0,3 \cdot 0,2 = 0,024.$$

Следовательно, $p(D) = 1 - p(\overline{D}) = 1 - 0,024 = 0,976.$

Примечание. $p(D) = p(B_1 + B_2 + B_3)$, но расчеты в этом случае более затруднительны.

Ответ. а) $p(A) = 0,188$; $p(B) = 0,788$; $p(C) = 0,336$; $p(D) = 0,976$.

Пример 1.21

Из 8 карточек, на каждой из которых написано по одной букве (Б, А, Н, К, О, М, А, Т), одну за другой выбирают наугад 4 карточки. Найти вероятность того, что получится слово «БАНК». Рассмотреть два случая:

а) выбранные карточки не возвращаются;

б) каждая выбранная карточка возвращается в совокупность, в которой все карточки перемешиваются, перед извлечение следующей.

Решение.

Введем следующие события:

A = {при извлечении трех карточек получится слово «ТОН»};

A_1 = {первая извлеченная буква – «Б»};

A_2 = {вторая извлеченная буква – «А»};

A_3 = {третья извлеченная буква – «Н»};

A_4 = {четвертая извлеченная буква – «К»};

тогда событие $A = A_1 \cdot A_2 \cdot A_3 \cdot A_4$

а) Если выбранные карточки не возвращаются, то события A_1, A_2, A_3, A_4 зависимы, и по теореме умножения для зависимых событий:

$$P(A) = P(A_1 \cdot A_2 \cdot A_3 \cdot A_4) = P(A_1) \cdot P_{A_1}(A_2) \cdot P_{A_1 A_2}(A_3) \cdot P_{A_1 A_2 A_3}(A_4)$$

$$\Rightarrow P(A) = \frac{1}{8} \cdot \frac{1}{7} \cdot \frac{1}{6} \cdot \frac{1}{5} = \frac{1}{1680}$$

б) Если выбранные карточки возвращаются, то события A_1, A_2, A_3, A_4 не зависимы и по теореме умножения для независимых событий:

$P(A) = P(A_1 \cdot A_2 \cdot A_3 \cdot A_4) = P(A_1) \cdot P(A_2) \cdot P(A_3) \cdot P(A_4) \ \Rightarrow$

$P(A) = \frac{1}{8} \cdot \frac{1}{8} \cdot \frac{1}{8} \cdot \frac{1}{8} = \frac{1}{4096}$

Ответ. а) $\frac{1}{1680}$; б) $\frac{1}{4096}$

Пример 1.22

В 1-й урне находится 7 белых и 5 черных шаров, а во второй 4 белых и 8 черных шаров. Из первой урны наудачу перекладывают во вторую 2 шара, а затем из второй урны извлекают один шар. Какова вероятность того, что он окажется белым?

Решение:

Обозначим событие А – из второй урны достаем белый шар, после того как в нее переложили два шара из первой урны. Рассмотрим три гипотезы:

1) H_1 – во вторую урну из первой переложили два белых шара, тогда вероятность этого события будет равна: $P(H_1) = \frac{7}{12} \cdot \frac{6}{11} = \frac{7}{22}$

2) H_2 – во вторую урну из первой переложили один белый, второй черный шар, тогда вероятность данного события будет равна: $P(H_2) = \frac{7}{12} \cdot \frac{8}{11} + \frac{8}{12} \cdot \frac{7}{11} = \frac{35}{66}$

3) H_3 – во вторую урну из первой переложили два черных шара, тогда вероятность этого события будет равна: $P(H_3) = \frac{5}{12} \cdot \frac{4}{11} = \frac{5}{33}$

Условные вероятности события А – «вынули из второй урны белый шар, после того как туда переложили два шара» будут соответственно равны:

$P_{H_1}(A) = \frac{6}{14} = \frac{3}{7}, \ \ P_{H_2}(A) = \frac{5}{14}, \ \ P_{H_3} = \frac{4}{14} = \frac{2}{7}$

Искомую вероятность события А – «достали белый шар из второй урны», находим по формуле полной вероятности:

$P(A) = \frac{7}{22} \cdot \frac{3}{7} + \frac{35}{66} \cdot \frac{5}{14} + \frac{5}{33} \cdot \frac{2}{7} = \frac{341}{924} = \frac{31}{84}$

Ответ: $\frac{31}{84}$

Примеры для самостоятельного решения.

1. В поступившей партии из 30 швейных машин 10 имеют внутренний дефект. Определить вероятность того, что из пяти наудачу взятых машин три окажутся бездефектными.

2. В урне находится 4 белых и 3 черных шара. Два игрока поочередно извлекают по шару (без возвращения). Выигрывает тот, кто 1-ым вытащит белый шар. Какова вероятность выигрыша для начинающего игру?

3. Какова вероятность получить выигрыш в игре Спортлото «5 из 6», полагающийся при угадывании: 3-х номеров из пяти; 4-х номеров из пяти; всех пяти номеров.

4. Устройство состоит из 3-х независимых элементов, работающих в течение времени T безотказно с вероятностями p_1=0,84; p_2=0,81; p_3=0,93. Найти вероятность того, что за время T выйдет из строя: а) хотя бы один элемент; б) только один элемент.

5. В жилом доме имеется n ламп, вероятность включения каждой из них в вечернее время равна 0,5. Найти вероятность того, что число одновременно включенных ламп будет между m1 и m2. Найти наивероятнейшее число включенных ламп среди n и его соответствующую вероятность. $n = 6400$, $m_1 = 3120$, $m_1 = 3200$.

6. Вычислительное устройство состоит из 1000 элементов, работающих независимо друг от друга. Вероятность отказа каждого элемента за смену равна p. Найти вероятность, что за смену откажут m элементов. $p = 0{,}024$, $m = 6$.

Тема 2. Случайные величины

2.1. Дискретная случайная величина

Определение. Случайной величиной *X* называется величина, принимающая то или иное заранее неизвестное числовое значение в зависимости от исхода испытания.

Различают дискретные и непрерывные случайные величины.

Дискретная случайная величина

Определение. *Дискретной случайной величиной* называется величина, которая может принимать только конечное или счетное множество значений, то есть такое множество, элементы которого можно пронумеровать.

Определение. Законом распределения дискретной случайной величины *X* называется соответствие между возможными значениями *X* и их вероятностями:

x	x_1	x_2	…	x_n
$p(x)$	$p(x_1)$	$p(x_2)$	…	$p(x_n)$

где $\sum_{i=1}^{n} p(x_i) = 1$.

Поскольку каждому значению x ДСВ ставиться в соответствие ее вероятность, то закон распределения можно задавать с помощью функции распределения ДСВ.

Функция распределения F(X) ДСВ X называется вероятность события того, что X< x: F(x) = P(X< x).

Свойства функции распределения:

1. функция распределения непрерывна при $x \neq x_i$ и имеет разрыв первого рода при $x=x_i$ равного p_i;
2. функция распределения постоянна на полуинтервале $(x_i; x_{i+1}]$;
3. функция распределения обладает свойством накопительной вероятности: $F(x) = \sum p_i$
4. функция распределения неубывающая, так как для любых $x_1 < x_2$ выполняется неравенство $f(x_1) \leq f(x_2)$

График функции распределения произвольной ДСВ представляет собой ступенчатую фигуру.

Основные характеристики дискретных случайных величин

Наиболее важными числовыми характеристиками случайной величины являются математическое ожидание *M*(*X*), дисперсия *D*(*X*) и среднее квадратическое отклонение σ(*X*).

Математическое ожидание дискретной случайной величины определяется соотношением:

$$M(X) = \sum_{i=1}^{n} x_i \cdot p\left(x_i\right) \tag{2.1}$$

Свойства математического ожидания:

1. *M(C) = C, где C = const;*
2. *M(CX) = C·M(X);*
3. *M (X ± Y) = M(X) ± M(Y);*
4. *M(X·Y) = M(X)·M(Y),* если *X* и *Y* – независимые случайные величины.

Дисперсия дискретной случайной величины определяется соотношением:

$$D\left(X\right) = M\left(X - M\left(X\right)\right)^2 = \sum_{i=1}^{n}\left(x_i - M\left(X\right)\right)^2 \cdot p\left(x_i\right) \tag{2.2}$$

Свойства дисперсии:

1. *D(C) = 0;*
2. *D(CX) =* C^2*·D(X);*
3. *D (X ± Y) = D(X) + D(Y),* если *X* и *Y* - независимые случайные величины.

Среднее квадратическое отклонение дискретной случайной величины определяется по формуле

$$\sigma\left(X\right) = \sqrt{D\left(X\right)}. \tag{2.3}$$

С вероятностной точки зрения математическое ожидание случайной величины определяет среднее арифметическое значение, которое принимает случайная величина при очень большом числе испытаний.

Дисперсия же определяет среднее арифметическое квадратов отклонений случайной величины *X* от ее математического ожидания при очень большом числе испытаний.

И дисперсия, и среднее квадратическое отклонение характеризуют степень рассеяния случайной величины *X* в области ее математического ожидания.

Пример 2.1

В лотерее разыгрывается 10 билетов; из них 4 являются выигрышными. Некто покупает 3 билета. Составить закон распределения числа купленных выигрышных билетов. Найти все числовые характеристики ДСВ X.

Решение.

Обозначим случайную величину X – число купленных выигрышных билетов. Поскольку выигрышных билетов может быть куплено: 0,1,2,3, то найдем вероятность каждого из этого события:

$$P_3(0)=\frac{C_4^0\cdot C_6^3}{C_{10}^3}=\frac{\frac{4!}{0!\cdot 4!}\cdot\frac{6!}{3!\cdot 3!}}{\frac{10!}{3!\cdot 7!}}=\frac{20}{120}=\frac{1}{6}$$

$$P_3(1)=\frac{C_4^1\cdot C_6^2}{C_{10}^3}=\frac{\frac{4!}{1!\cdot 3!}\cdot\frac{6!}{2!\cdot 4!}}{\frac{10!}{3!\cdot 7!}}=\frac{60}{120}=\frac{1}{2}$$

$$P_3(2)=\frac{C_4^2\cdot C_6^1}{C_{10}^3}=\frac{\frac{4!}{2!\cdot 2!}\cdot\frac{6!}{1!\cdot 5!}}{\frac{10!}{3!\cdot 7!}}=\frac{36}{120}=\frac{3}{10}$$

$$P_3(3)=\frac{C_4^3\cdot C_6^0}{C_{10}^3}=\frac{\frac{4!}{3!\cdot 1!}\cdot\frac{6!}{0!\cdot 6!}}{\frac{10!}{3!\cdot 7!}}=\frac{4}{120}=\frac{1}{30}$$

Тогда закон распределения числа купленных выигрышных билетов будет иметь вид:

x_i	0	1	2	3
p_i	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{30}$

Найдем числовые характеристики:

1) $M(X)=0\cdot\frac{1}{6}+1\cdot\frac{1}{2}+2\cdot\frac{3}{10}+3\cdot\frac{1}{30}=0,5+0,6+0,1=1,2$

2) $D(X)=\left(0\cdot\frac{1}{6}+1\cdot\frac{1}{2}+4\cdot\frac{3}{10}+9\cdot\frac{1}{30}\right)-(1,2)^2=(0,5+1,2+0,3)-1,44=0,56$

3) $\sigma(X)=\sqrt{0,56}\approx 0,75$

4) $M_O(X)=1$

5) $M_E(X)=1$

Ответ. 1) 1,2; 2) 0,56; 3) 0,75; 4) 1; 5) 1.

Пример 2.2

Даны законы распределения двух независимых случайных величин:

X	2	4	6	8
p	0,4	0,2	0,1	0,3

Y	0	1	2
p	0,5	0,2	0,3

Найти математическое ожидание и дисперсию случайной величины Z= 2X – 3Y.

Решение.

Используя свойство математического ожидания и дисперсии. Имеем:

$M(Z) = M\,(2X - 3Y) = M(2X) - M(3Y) = 2M(X) - 3M(Y)$

$D(Z) = D\,(2X - 3Y) = D(2X) + D(3Y) = 4D(X) + 9D(Y)$

Найдем M(X), M(Y), D(X), D(Y).

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

$M(Y) = 0\cdot 0{,}5+1\cdot 0{,}2+2\cdot 0{,}3= 0{,}2+0{,}6 =0{,}8$

$M(Z) = 2\cdot 4{,}6+3\cdot 0{,}8=9{,}2+2{,}4=11{,}6$

$D(X)= (4\cdot 0{,}4+16\cdot 0{,}2+36\cdot 0{,}1+64\cdot 0{,}3) - 21{,}16=(1{,}6+3{,}2+3{,}6+19{,}2) - 21{,}16 =(4{,}8+22{,}8) - 21{,}16 =27{,}6 - 21{,}16=6{,}44$

$D(Y) = (0\cdot 0{,}5+1\cdot 0{,}2+4\cdot 0{,}3) - 0{,}64 = 1{,}4 - 0{,}64=0{,}76$

$D(Z) = 4\cdot 6{,}44+9\cdot 0{,}76=25{,}76+6{,}84 = 32{,}6$

Ответ. *M(Z) = 11,6; D(Z) = 32,6.*

2.2. Непрерывная случайная величина

Определение. Случайная величина называется *непрерывной,* если она может принимать все возможные значения из некоторого конечного или бесконечного интервала.

Определение. Интегральной функцией распределения случайной величины *X* называется функция, определяемая соотношением:

$$F(x) = P\,(X < x) = \int_{-\infty}^{x} f(x)\,dx. \qquad (2.4)$$

Для дискретной случайной величины

$F(x) = P\,(X < x) = \sum_{x_k < x} p(x_k)$.

Интегральная функция распределения обладает следующими свойствами:

1) $0 \le F(x) \le 1$, $F(-\infty) = 0$, $F(+\infty) = 0$;

2) $F(x)$ – неубывающая функция, т.е. $F(x_1) \le F(x_2)$ при $x_1 \le x_2$;

3) $P\,(x_1 < X < x_2) = F(x_2) - F(x_1)$.

Определение. Дифференциальной функцией распределения $f(x)$ или плотностью вероятности называется первая производная от интегральной функции распределения: $f(x) = F'(x)$. Понятие $f(x)$ применимо только для непрерывных случайных величин.

Функция $f(x)$ обладает следующими свойствами:

1) $f(x) \geq 0$;

2) $\int\limits_{-\infty}^{+\infty} f(x)dx = 1$; (2.5)

3) $P(x_1 \leq X \leq x_2) = \int\limits_{x_1}^{x_2} f(x)dx$. (2.6)

Математическое ожидание, дисперсия и среднее квадратическое отклонение

непрерывной случайной величины определяются формулами:

$$M(X) = \int\limits_{-\infty}^{+\infty} x \cdot f(x)dx;\ D(X) = \int\limits_{-\infty}^{+\infty} (X - M(X))^2 \cdot f(x)dx;\ \sigma(X) = \sqrt{D(X)}.$$

Замечание. Если значения непрерывной случайной величины X заполняют некоторый интервал (a, b), т.е. $f(x) \geq 0$ при $x \in (a, b)$, $f(x) \equiv 0$ при $x \notin (a, b)$, тогда

$$\int\limits_a^b f(x)dx = 1;\ M(X) = \int\limits_a^b x \cdot f(x)dx;\ D(X) = \int\limits_{-\infty}^{+\infty} (X - M(X))^2 \cdot f(x)dx.$$

Для вычисления дисперсии можно использовать формулу:

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

Пример 2.3

Плотность вероятности непрерывной случайной величины X определяется формулой:

$$f(x) = \begin{cases} 0, & при \quad x \leq 1, \\ x - \dfrac{1}{2}, & при \quad 1 < x \leq 2, \\ 0, & при \quad x > 2. \end{cases}$$

Найти функцию распределения $F(X)$.

Решение.

$$F(x)=\int_{-\infty}^{x} f(x)\cdot dx=0, \quad при \quad x\le 1,$$

$$F(x)=\int_{-\infty}^{x} f(x)\cdot dx=\int_{-\infty}^{1} f(x)dx+\int_{1}^{x} f(x)dx=0+\frac{x^2}{2}-\frac{x}{2}=\frac{(x^2-x)}{2}, \quad при \quad 1<x\le 2,$$

$$F(x)=\int_{-\infty}^{x} f(x)dx=\int_{-\infty}^{1} f(x)dx+\int_{1}^{2} f(x)dx+\int_{2}^{x} f(x)dx=\left.\frac{(x^2-x)}{2}\right|_{1}^{2}=1, \quad при \quad x>2.$$

Таким образом, функция распределения имеет вид:

$$F(X)=\begin{cases} 0, \quad при \quad x\le 1, \\ \dfrac{x^2-x}{2}, \quad при \quad 1<x\le 2, \\ 1, \quad при \quad x>2. \end{cases}$$

Пример 2.4

Плотность вероятности непрерывной случайной величины X задана в интервале (-п/2; п/2) функцией *f(x) =C cos x*. Вне этого интервала *f(x) =0*. Найти параметр C и определить вероятность попадания случайной величины X в интервал (0; п/4).

Решение.

На основании равенства: $\int_{-\infty}^{+\infty} f(x)dx=1$ имеем:

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} C\cos x\,dx=1 \;\Rightarrow C\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos x\,dx=1 \;\Rightarrow C\cdot(\sin x)\Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}}=1$$

$$\Rightarrow C\cdot\left(\sin\frac{\pi}{2}-\sin\left(-\frac{\pi}{2}\right)\right)=1 \;\Rightarrow 2C=1 \;\Rightarrow C=0{,}5$$

Найдем вероятность попадания случайной величины X в интервал (0; п/4) по формуле:

P(a<X<b) = F(b) – F(a).

Найдем функцию распределения:

$$F(x)=\int_{-\infty}^{x} f(x)\cdot dx=0, \quad при \quad x\le -\frac{\pi}{2},$$

$$F(x)=\int_{-\infty}^{x} f(x)\cdot dx=\int_{-\infty}^{-\frac{\pi}{2}} f(x)dx+\int_{-\frac{\pi}{2}}^{x} f(x)dx=0+0{,}5\sin x, \quad при \quad -\frac{\pi}{2}<x\le\frac{\pi}{2},$$

$$F(x)=\int_{-\infty}^{x} f(x)dx=\int_{-\infty}^{-\frac{\pi}{2}} f(x)dx+\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(x)dx+\int_{\frac{\pi}{2}}^{x} f(x)dx=0{,}5\sin x\Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}}=0{,}5(1-(-1))=1, \quad при \quad x>\frac{\pi}{2}.$$

Таким образом, функция распределения имеет вид:

$$F(X)=\begin{cases} 0, \quad при \quad x\le -\frac{\pi}{2}, \\ 0{,}5\sin x, \quad при \quad -\frac{\pi}{2}<x\le\frac{\pi}{2}, \\ 1, \quad при \quad x>\frac{\pi}{2}. \end{cases}$$

P(0≤X≤ п/4) = F(п/4) – F(0) = 0,5 sin (п/4) – 0,5 sin 0 = $\frac{\sqrt{2}}{4}\approx$ *0,35*

Ответ. *C= 0,5; P(0≤X≤ п/4) ≈0,35*

Пример 2.5

Случайная величина X задана в интервале (0; п) плотностью вероятности *f(x)= 0,5 sin x,* вне этого интервала *f(x) = 0*. Найти математическое ожидание и дисперсию величины X.

Решение.

Для нахождения математического ожидания воспользуемся формулой:

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

$$M(X)=\int_0^{\pi} 0{,}5x\cdot \sin x\, dx=\begin{vmatrix} u=x & dv=\sin x\, dx \\ du=dx & v=-\cos x\end{vmatrix}=$$

$$=0{,}5\left(-x\cdot\cos x\Big|_0^{\pi}+\int_0^{\pi}\cos x\, dx\right)=0{,}5\left(-x\cdot\cos x\Big|_0^{\pi}+\sin x\Big|_0^{\pi}\right)=$$

$$=0{,}5((-\pi\cdot\cos\pi-0)+(\sin\pi-\sin 0))=0{,}5\pi$$

Для нахождения дисперсии воспользуемся формулой: $D(X)=\int_{-\infty}^{\infty} x^2 \cdot f(x)dx - M^2(X)$

$$D(X) = 0{,}5\int_0^{\pi} x^2 \cdot \sin x\, dx - 0{,}25\pi^2 = \begin{vmatrix} u = x^2 & dv = \sin x\, dx \\ du = 2xdx & v = -\cos x \end{vmatrix} =$$

$$= 0{,}5\left(-x^2 \cos x \Big|_0^{\pi} + 2\int_0^{\pi} x \cos x\, dx\right) - 0{,}25\pi^2 = \begin{vmatrix} u = x & dv = \cos x\, dx \\ du = dx & v = \sin x \end{vmatrix} =$$

$$= \left(-0{,}5x^2 \cos x \Big|_0^{\pi} + \left(x \sin x \Big|_0^{\pi} - \int_0^{\pi} \sin x\, dx\right)\right) =$$

$$= \left(-0{,}5x^2 \cos x \Big|_0^{\pi} + x \cdot \sin x \Big|_0^{\pi} + \cos x \Big|_0^{\pi}\right) - 0{,}25\pi^2 =$$

$$= 0{,}5\pi^2 + (-1-1) - 0{,}25\pi^2 = \frac{\pi^2}{4} - 2$$

Ответ. *M(X)=0,5π; D(X)= 0,25π² – 2.*

2.3. Основные законы распределения случайных величин

Дискретная случайная величина			
/п	Закон распределения	Формулы для составления закона распределения	Числовые характеристики
1	Биноминальный закон	$P_n(m) = C_n^m \cdot p^m \cdot q^{n-m}$	$M(X) = np$ $D(X) = npq$ $\sigma = \sqrt{npq}$
2	Геометрический закон	$P_k = p \cdot q^{k-1}$	$M(X) = \frac{1}{p}$ $D(X) = \frac{q}{p^2}$ $\sigma(X) = \frac{\sqrt{q}}{p}$
3	Гипергеометрический закон	$P(A) = \frac{C_N^n \cdot C_{M-N}^{m-n}}{C_M^m}$	$M(X) = m\frac{N}{M}$ $D(X) = m\frac{N}{M-1}\left(1-\frac{N}{M}\right)\left(1-\frac{m}{M}\right)$
4	Закон Пуассона	$P_n(k) = \frac{\lambda^k}{k!}\ell^{-\lambda}$	$M(X) = \lambda$ $D(X) = \lambda$

Пример 2.6 (Биноминальный закон)

Вероятность того, что аудитор допустит ошибку при проверке бухгалтерского баланса, равна 0,05. Аудитору на заключение представлено 2 баланса. Составить закон распределения числа правильных заключений на проверяемые балансы. Найти математическое ожидание, дисперсию и среднее квадратическое отклонение.

Решение.

Пусть X – случайная величина числа правильных заключений на проверяемые балансы. Аудитор может не допустить ошибки при проверке, может допустить одну ошибку и две ошибки, т.е. $x_1 = 0$; $x_2 = 1$; $x_3 = 2$. Вероятности вычисляем по формуле Бернулли, при этом n= 2, p=0,95, q= 0,05:

$$P_2(0) = C_2^0 \cdot p^0 \cdot q^2 = \frac{2!}{0! \cdot 2!} \cdot (0{,}95)^0 \cdot (0{,}05)^2 = 0{,}0025$$

$$P_2(1) = C_2^1 \cdot p^1 \cdot q^1 = \frac{2!}{1! \cdot 1!} \cdot (0{,}95) \cdot (0{,}05) = 2 \cdot 0{,}05 \cdot 0{,}95 = 0{,}095$$

$$P_2(2) = C_2^2 \cdot p^2 \cdot q^0 = \frac{2!}{2! \cdot 0!} \cdot (0{,}95)^2 \cdot (0{,}05)^0 = 0{,}9025$$

Закон распределения числа правильных заключений на проверяемые балансы будет иметь вид:

x	0	1	2
p	0,0025	0,0950	0,9025

По формулам биноминального закона распределения, найдем числовые характеристики:

M(X) = n⋅p *M(X) = 2⋅0,95 = 1,9*

D(X) =n⋅p⋅q *D(X) = 2⋅0,95⋅0,05=0,095*

$\sigma(X) = \sqrt{npq}$ $\sigma(X) = \sqrt{0{,}095} \approx 0{,}31$

Пример 2.7 (Геометрический закон)

В партии, состоящей из 10 холодильников, 8 – стандартных. Товаровед проверяет качество полученного товара. Составить закон распределения числа возможных проверок холодильников до первого нестандартного, если товаровед перестал проверять товар после четвертого холодильника. Найти математическое ожидание, дисперсию и среднее квадратическое отклонение.

Решение.

Вероятность того, что товаровед проверит стандартный холодильник, равна 0,8. Случайная величина X – число проверенных холодильников до первого нестандартного. Товаровед может проверить сразу нестандартный холодильник, либо нестандартный холодильник может быть после одного стандартного, либо после двух, либо после трех.

Найдем вероятности данных событий, если $n = 4$, $p= 0,8$, $q= 0,2$.

$P_4(1) = q$ $\quad P_4(1) = 0,2$

$P_4(2) = q \cdot p$ $\quad P_4(2) = 0,8 \cdot 0,2 = 0,16$

$P_4(3) = p^2 \cdot q$ $\quad P_4(3) = 0,8^2 \cdot 0,2 = 0,64 \cdot 0,2 = 0,128$

$P_4(4) = p^3 \cdot q + p^4$ $\quad P_4(4) = 0,8^3 \cdot 0,2 + 0,8^4 = 0,1024 + 0,4096 = 0,512$

Закон распределения число проверенных холодильников до первого нестандартного будет иметь вид:

x	1	2	3	4
p	0,200	0,160	0,128	0,512

По формулам геометрического закона распределения, найдем числовые характеристики:

$$M(X) = \frac{1}{p} \qquad M(X) = \frac{1}{0,8} = 1,25$$

$$D(X) = \frac{q}{p^2} \qquad D(X) = \frac{0,2}{0,64} = 0,3125$$

$$\sigma(X) = \frac{\sqrt{q}}{p} \qquad \sigma(X) = \frac{\sqrt{0,2}}{0,8} = 0,56$$

Пример 2.8 (Гипергеометрический закон)

Среди 10 лотерейных билетов имеется 4 билета с выигрышем. Наудачу покупают два билета. Написать закон распределения вероятностей числа выигрышных билетов среди купленных. Найти математическое ожидание, дисперсию и среднее квадратическое отклонение.

Решение.

Пусть X – случайная величина числа выигрышных билетов среди купленных 2 билетов. Очевидно, что оно может принимать значения: $x_1 = 0$; $x_2 = 1$; $x_3 = 2$. Для определения вероятности появления каждого из этих значений воспользуемся формулой:

$$P(A) = \frac{C_N^n \cdot C_{M-N}^{m-n}}{C_M^m}$$

$$P_2(0)=\frac{C_4^0\cdot C_6^2}{C_{10}^2}=\frac{\frac{4!}{0!\cdot4!}\cdot\frac{6!}{2!\cdot4!}}{\frac{10!}{2!\cdot8!}}=\frac{15}{45}=\frac{1}{3}$$

$$P_3(1)=\frac{C_4^1\cdot C_6^1}{C_{10}^2}=\frac{\frac{4!}{1!\cdot3!}\cdot\frac{6!}{1!\cdot5!}}{\frac{10!}{2!\cdot8!}}=\frac{24}{45}=\frac{8}{15}$$

$$P_2(2)=\frac{C_4^2\cdot C_6^0}{C_{10}^2}=\frac{\frac{4!}{2!\cdot2!}\cdot\frac{6!}{0!\cdot6!}}{\frac{10!}{2!\cdot8!}}=\frac{6}{45}=\frac{2}{15}$$

Закон распределения числа выигрышных билетов среди купленных 2 билетов будет иметь вид:

x	0	1	2
p	$\frac{1}{3}$	$\frac{8}{15}$	$\frac{2}{15}$

По формулам гипергеометрического закона распределения, найдем числовые характеристики:

$$M(X)=m\frac{N}{M} \qquad M(X)=2\cdot\frac{4}{10}=0{,}8$$

$$D(X)=m\frac{N}{M-1}\left(1-\frac{N}{M}\right)\left(1-\frac{m}{M}\right) \qquad D(X)=2\cdot\frac{4}{10-1}\cdot\left(1-\frac{4}{10}\right)\cdot\left(1-\frac{2}{10}\right)=0{,}43$$

$$\sigma(X)=\sqrt{D(X)} \qquad \sigma(X)=\sqrt{0{,}43}\approx0{,}66$$

Пример 2.9 (Закон Пуассона)

Завод отправил в торговую сеть 500 изделий. Вероятность повреждения изделия в пути равна 0,003. Найти вероятность того, что при транспортировке будет повреждено: а) ни одного изделия, б) ровно три изделия.

Решение:

Для решения используем формулу Пуассона: $P_m\approx\frac{\lambda^m}{m!}\cdot\ell^{-\lambda}$

а) Найдем вероятность того, что ни одно изделие не повреждено.

Имеем: $\lambda=n\cdot p=500\cdot0{,}003=1{,}5$. Отсюда $P_0\approx\frac{1{,}5^0}{0!}\cdot\ell^{-1{,}5}\approx\ell^{-1{,}5}\approx0{,}2231$.

б) Найдем вероятность того, что повреждено ровно три изделия: $P_0 \approx \frac{1{,}5^3}{3!} \cdot \ell^{-1{,}5} \approx 0{,}1255$

Ответ. а) 0,2231; б) 0,2155.

Непрерывная случайная величина				
№ п\п	Законы распределения	Функция распределения	Плотность распределения вероятности	Основные формулы
1	Нормальный закон распределения	$F(X) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} \ell^{-\frac{(x-m)^2}{2\sigma^2}} dx$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\ell^{-\frac{(x-m)^2}{2\sigma^2}}$	$M(X)=m$ $D(X)=\sigma^2$ $P(\alpha < x < \beta) = \Phi\left(\frac{\beta-m}{\sigma}\right) - \Phi\left(\frac{\alpha-m}{\sigma}\right)$ $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \ell^{-\frac{t^2}{2}} dt$ функция Лапласа
2	Равномерный закон распределения	$F(x) = \begin{cases} 0, \text{ при } x < a; \\ \frac{x-a}{b-a}, \text{ при } a \le x \le b; \\ 1, \text{ при } x > b \end{cases}$	$f(x) = \begin{cases} 0, \text{ при } x < a; \\ \frac{1}{b-a}, \text{ при } a \le x \le b; \\ 0, \text{ при } x > b. \end{cases}$	$M(X) = \frac{b+a}{2}$ $D(X) = \frac{(b-a)^2}{12}$ $\sigma(X) = \frac{b-a}{2\sqrt{3}}$
3	Показательный (экспоненциальный) закон распределения	$F(x) = \begin{cases} 0, \text{ при } x < 0; \\ 1-\ell^{-\lambda x}, \text{ при } x \ge 0. \end{cases}$	$f(x) = \begin{cases} 0, \text{ при } x < 0; \\ \lambda\ell^{-\lambda x}, \text{ при } x \ge 0. \end{cases}$	$M(X) = \frac{1}{\lambda}$ $D(X) = \frac{1}{\lambda^2}$ $\sigma(X) = \frac{1}{\lambda}$

Пример 2.10 (Нормальный закон распределения)

Математическое ожидание нормально распределенной случайной величины X равно 5, дисперсия равна 9. Найти выражение для плотности вероятности.

Решение.

Случайная величина X распределена по нормальному закону распределения, если ее функция плотности распределения вероятностей имеет вид:

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

По условию задачи: m= 5, $\sigma^2 = 9$, тогда плотность вероятности по данным задачи будет иметь вид: $f(x) = \frac{1}{3\sqrt{2\pi}} \ell^{-\frac{(x-5)^2}{18}}$.

Ответ. $f(x) = \frac{1}{3\sqrt{2\pi}} \ell^{-\frac{(x-5)^2}{18}}$

Пример 2.11

Математическое ожидание нормально распределенной случайной величины X равно 12, а среднее квадратическое отклонение 2. Найти вероятность того, что случайная величина X примет значение, заключенное в интервале (14; 16).

Решение.

Используя формулу: $P(\alpha < x < \beta) = \Phi\left(\frac{\beta - m}{\sigma}\right) - \Phi\left(\frac{\alpha - m}{\sigma}\right)$ и учитывая, что *m =12, σ =2*, получим: $P(14 < x < 16) = \Phi\left(\frac{16-12}{2}\right) - \Phi\left(\frac{14-12}{2}\right) = \Phi(2) - \Phi(1)$.

По таблице значений функции Лапласа находим *Ф (2)* = 0, 4772;

Ф (1) = 0,3413

P(14<X<16) = 0, 4772 − 0,3413 = 0,1359.

Ответ. *P(14<X<16)* = 0,1359.

Пример 2. 12

Известно, что средний расход удобрений на один гектар пашни составляет 80 кг, а среднее квадратическое отклонение расхода равно 5 кг. Считая расход удобрений нормально распределенной случайной величиной, определить диапазон, в который вносимая доза удобрений попадает с вероятностью 0,98.

Решение.

Так как

$$P(\alpha < x < \beta) = \Phi\left(\frac{\beta - m}{\sigma}\right) - \Phi\left(\frac{\alpha - m}{\sigma}\right) = p \Rightarrow 2\Phi\left(\frac{X - m}{\sigma}\right) = p \Rightarrow \Phi\left(\frac{X - m}{\sigma}\right) = \frac{p}{2}$$

Отсюда:

$$\Phi\left(\frac{X - m}{\sigma}\right) = 0{,}49 \Rightarrow X - m = 2{,}33 \cdot \sigma \Rightarrow X - m = 2{,}33 \cdot 5 \Rightarrow X - m = 11{,}65$$

$$\beta = 80 + 11{,}65 = 91{,}65; \quad \alpha = 80 - 11{,}65 = 68{,}35$$

Таким образом, искомый интервал будет иметь вид (68,35; 91,65).

Ответ. (68,35; 91,65).

Пример 2.13

Математическое ожидание нормально распределенной величины – количества сыра, используемого для приготовления 100 бутербродов равно 1 кг. Известно, что с вероятностью 0,96 расход сыра на изготовления 100 бутербродов составляет от 900 г. до 1100 г. Определить среднее квадратическое отклонение расхода сыра на 100 бутербродов.

Решение.

По условию задачи:

$$P(1100 < x < 900) = \Phi\left(\frac{1100-1000}{\sigma}\right) - \Phi\left(\frac{900-1000}{\sigma}\right) = 0{,}96$$

$$\Rightarrow 2\Phi\left(\frac{100}{\sigma}\right) = 0{,}96 \;\Rightarrow\; \Phi\left(\frac{100}{\sigma}\right) = 0{,}48 \;\Rightarrow$$

$\Rightarrow \frac{100}{\sigma} = 2{,}05 \;\Rightarrow \sigma = \frac{100}{2{,}05} \Rightarrow \sigma = 48{,}8$. **Ответ.** 48,8 г.

Пример 2.14 (Равномерный закон распределения)

Автобусы подходят к остановке с интервалом в 5 минут. Считая, что случайная величина X – время ожидания автобуса – распределена равномерно, найти функцию распределения, функцию плотности вероятности, математическое ожидание, дисперсию и среднеквадратическое отклонение случайной величины.

Решение.

Время ожидание автобуса на остановке определенно интервалом [0;5]. Следовательно,

– функция распределения случайной величины X будет иметь вид:

$$F(X) = \begin{cases} 0, & при \quad x \le 0, \\ \dfrac{x}{5}, & при \quad 0 < x \le 5, \\ 1, & при \quad x > 5. \end{cases}$$

– плотность вероятности: $f(X) = \begin{cases} 0, & при \quad x \le 0, \\ \dfrac{1}{5}, & при \quad 0 < x \le 5, \\ 0, & при \quad x > 5. \end{cases}$

– математическое ожидание: $M(X) = \dfrac{b+a}{2} = \dfrac{5+0}{2} = 2{,}5$

– дисперсия: $D(X) = \dfrac{(b-a)^2}{12} = \dfrac{(5-0)^2}{12} = \dfrac{25}{12} = 2{,}08$

– среднее квадратическое отклонение: $\sigma(X) = \dfrac{b-a}{2\sqrt{3}} = \dfrac{5}{2\sqrt{3}}$

Ответ. M(X) =2,5; D(X) = 2,08; $\sigma(X) = \dfrac{5}{2\sqrt{3}}$

Пример 2.15

Поезда метрополитена идут с интервалом в 2 минуты. Время ожидания поезда – равномерно распределенная случайная величина. С какой вероятностью пассажир, пришедший на платформу, будет ожидать поезд более одной минуты?

Решение.

Обозначим случайную величину X – время ожидания пассажиром поезда в метрополитене.

Функция распределения случайной величины X имеет вид:

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

Вероятность того, что пассажир будет ожидать поезд более одной минуты, являются противоположной для вероятности ожидания поезда менее одной минуты, поэтому сначала найдем вероятность того, что пассажир будет ожидать поезд менее одной минуты по формуле:

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

Искомая вероятность того, что пассажир ждет поезд более одной минуты, равна: 1 – 0,5 = 0,5.

Ответ. 0,5

2.4. Функция надежности

Назовем элементом устройство вне зависимости от его сложности. Рассмотрим время работы данного устройства в системах массового обслуживания.

Пусть t=0 начальный момент времени работы элемента, тогда T – период времени, через который произошел отказ в работе.

Величина T – непрерывная случайная величина – безотказной работы устройства.

Вероятность отказа работы элемента за некоторое время продолжительностью t определяется функцией распределения $F(t)= P(T<t)$.

Поэтому *вероятность безотказной работы* за время t как событие противоположное вероятности отказа можно найти по формуле:

$R(t) = P(T>t) = 1 - F(t)$, где функция $R(t)$ – называется **функцией надежности.**

Продолжительность безотказной работы случайная величина, которая соответствует показательному закону распределения НСВ: $F(t)=1-\ell^{-\lambda t}$, где λ - интенсивность отказов (среднее число отказов за единицу времени).

Поэтому функция надежности принимает вид: R (t) = 1 – F(t)= $\ell^{-\lambda t}$

Функция надежности, заданную зависимостью R (t) = $\ell^{-\lambda t}$, называют ***показательным законом надежности*** и используют для вычисления вероятности безотказной работы элемента за время t.

Основное свойство показательного закона надежности заключается в том, что вероятность безотказной работы элемента за время t не зависит от времени его работы до рассматриваемого временного интервала, а зависит лишь длительности интервала t при известной интенсивности отказов λ.

Таким образом, условная вероятность безотказной работы элемента на интервале времени T равно безусловной вероятности, т.е. прошлый опыт работы не учитывается.

Пример 2.16

Длительность безотказной работы элемента распределена по показательному закону при $t>0$: $F(t)= 1 - e^{-0,02t}$. Найти вероятность того, что за время $t=40$ *ч:*

а) элемент не откажет;

б) элемент выйдет из строя.

Решение.

Условие задачи указывает на показательный закон надежности для времени $t=40$ *ч*.

а) Вероятность безотказной работы за 40 ч. равна: $R\ (40) = e^{-0,02\cdot 40} = e^{-0,8}$ = 0,45;

б) Вероятность противоположного события – отказа элемента за 40 ч. равна:

$F\ (40)$ = 1 – 0,45=0,55.

Ответ. а) 0,45; б) 0,55.

Пример 2.17

Время t – случайная величина, за которую в железнодорожном депо расформировывают состав. За один час может быть расформировано шесть поездов. Определить вероятность того, что время, за которое состав будет расформирован:

а) меньше 0,5 ч.;

б) меньше 24 мин., но больше 6 мин.

Решение.

Так как вероятность расформировать состав подчиняется показательному закону, обозначим среднее число поездов в час как интенсивность отказов: $\lambda = 6$ ч$^{-1}$.

а) За время t =0, 5 ч вероятность расформировать состав равна:

$F\ (0,5) = P(T<0,5) = 1 - e^{-6\cdot 0,5} = 1 - e^{-3} = 1 - 0,05 = 0,95$.

б) Вероятность того, что время, за которое состав будет расформирован, при t= 6 мин. = 0,1 ч.,

при t= 24 мин. = 0,4 ч. составляет:

$P\ (0,1 < T < 0,4) = e^{-6\cdot 0,4} - e^{-6\cdot 0,1}$ = 0,55 – 0,09 = 0,46.

Ответ. а) 0,95; б) 0,46.

Примеры для самостоятельного решения.

1. В магазине имеется 15 автомобилей определенной марки. Среди них 7 черного цвета, 6 серого и 2 белого. Представители фирмы обратились в магазин с предложением о продаже им 3 автомобилей этой марки, безразлично какого цвета. Составьте ряд распределения числа проданных автомобилей черного цвета при условии, что автомобили отбирались случайно.

2. В городе 4 коммерческих банка. У каждого риск банкротства в течение года составляет 20%. Составьте ряд распределения числа банков, которые могут обанкротиться в течение следующего года.

3. Охотник стреляет по дичи до первого попадания, но успевает сделать не более четырех выстрелов. Составить закон распределения числа промахов, если вероятность попадания в цель при одном выстреле равна 0,7. Найти дисперсию этой случайной величины.

4.Книга издана тиражом 20000 экземпляров. Вероятность того, что в книге имеется дефект брошюровки, равна 0,0002. Найти вероятность того, что тираж содержит 5 неправильно сброшюрованных книг.

5.Задана плотность распределения вероятностей непрерывной случайной величины X:

$$f(x) = \begin{cases} Ax, & если\ x \in [1;5) \\ 0, & если\ x \in [1;5) \end{cases}$$

Найти:

1) коэффициент *A;*

2) функцию распределения *F(X)*;

3) построить графики *f(x), F(x)*;

4) вероятность того, что X примет значение из промежутка [1;3).

6. Автобусы некоторого маршрута идут строго по расписанию и интервалом 7 минут. Составить функцию плотности случайной величины T – времени ожидании очередного автобуса пассажиром, который наудачу подошёл к остановке. Найти вероятность того, что он будет ждать автобус не более трёх минут. Найти функцию распределения *F(t)*.

7. Время в годах безотказной работы прибора подчинено показательному закону, т.е. плотность

распределения этой случайной величины такова: $f(t)=2e^{-2t}$ при $t \geq 0$ и $f(t)=0$ при

$t < 0$.

а) Найти формулу функции распределения этой случайной величины.

б) Определить вероятность того, что прибор проработает не более года.

в) Определить вероятность того, что прибор безотказно проработает 3 года.

г) Определить среднее ожидаемое время безотказной работы прибора.

8. Диаметр подшипников, изготовленные на заводе, представляет собой случайную величину, распределенную нормально с математическим ожиданием 1,5 см и средним квадратическим отклонением 0,04 см. Найти вероятность того, что размер наугад взятого подшипника колеблется от 1,4 до 1,6 см.

Тема 3. Система случайных величин

3.1. Дискретная двумерная случайная величина

Определение. Упорядоченная пара случайных величин (X; Y) называется двумерной случайной величиной.

Определение. Возможным значением двумерной случайной величины (X; Y) называется упорядоченная пара чисел вида ($X = x_i$; $Y = y_j$), а ее вероятностью – вероятность события ($X = x_i$; $Y = y_j$): $p_{ij} = p(X = x_i;\ Y = y_j)$.

Определение. Законом распределения дискретной двумерной случайной величины называется перечень возможных значений (x_i; y_j) этой величины и их вероятностей p_{ij} ($i = 1,2, \dots, m$; $j = 1,2, \dots, n$).

Обычно двумерное распределение задается в виде таблицы:

X	$Y=y_1$	$Y=y_2$	$\dots$	$Y=y_n$
$X=x_1$	p_{11}	p_{12}	$\dots$	p_{1n}
$X=x_2$	p_{21}	p_{22}	$\dots$	p_{2n}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$X=x_m$	p_{m1}	p_{m2}	$\dots$	p_{mn}

Так как события ($X = x_i$; $Y = y_j$), где $i = 1, 2, \dots, m$ и $j = 1, 2, \dots, n$, образуют полную группу, то сумма вероятностей p_{ij} в данной таблице равна 1.

Функцией распределения системы случайных величин (X,Y) называется функция F(x,y) , которая для любых действительных чисел x и y равна вероятности совместного выполнения двух событий {X<x} и {Y<y} , т.е.

$$F(x, y) = P(X<x, Y<y) \qquad (3.1)$$

Свойства двумерной функции распределения:

1. $0 \le P(x, y) \le 1$;
2. *функция $F(x,y)$ неубывающая: $F(x_2,y) \le F(x_1,y)$ при $x_2 \le x_1$; $F(x,y_2) \le F(x,y_1)$ при $y_2 \le y_1$;*
3. *функция $F(x, y)$ непрерывна слева по каждому из своих аргументов;*
4. $F(x, -\infty) - F(-\infty, y) = F(-\infty, -\infty) = 0$;
5. $F(+\infty, +\infty) = 1$;
6. $F(x, +\infty) = F_X(x)$, $F(+\infty, y) = F_Y(y)$;
7. $F(x, y) = \sum\sum p_{ij}$.

Вероятность попадания случайной точки (X, Y) в прямоугольник D:

$$P\{x_1 \le X \le x_2,\ y_1 \le Y \le y_2\} = F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) + F(x_1,y_1) \qquad (3.2)$$

Независимые случайные величины

Случайные величины называются независимыми, если независимыми являются события: {X<x} и {Y<y} для любых двух действительных чисел x и y. В противном случае случайные величины называются зависимыми.

Для двух дискретных случайных величин X и Y необходимым и достаточным условием их **независимости** является равенство:

$$P\{X=x_i,\ Y=y_j\} = P\{X=x_i\}\cdot P\{Y=y_j\} \qquad (3.3)$$

Условные законы распределения

Условным законом распределения одной из случайных величин, входящих в систему (X,Y), называется закон ее распределения , найденный при условии, что другая случайная величина приняла определенное значение.

В системе двух дискретных случайных величин (X, Y), условным законом распределения СВ Y при условии, что X=x_i называется совокупность вероятностей:

$$P\{Y \mid X = x_i\} = \frac{P\{X = x_i\ Y = y_j\}}{P\{X = x_i\}} \qquad (3.4)$$

В системе двух дискретных случайных величин (X,Y), условным законом распределения СВ X при условии, что Y=y_j называется совокупность вероятностей:

$$P\{X \mid Y = y_j\} = \frac{P\{X = x_i\ Y = y_j\}}{P\{Y = y_j\}} \qquad (3.5)$$

Безусловные вероятности дискретных компонент *X* и *Y* находятся по формулам

$$p(X = x_i) = \sum_{j=1}^{m} p_{ij}, \quad p(Y = y_j) = \sum_{i=1}^{n} p_{ij}. \qquad (3.6)$$

Математическое ожидание и дисперсия системы случайных величин

Математическим ожидание системы случайных величин (X, Y) называется совокупность двух математических ожиданий *M(X) и M(Y)* определяемых равенством:

$$\text{M(X)} = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i p_{ij} \qquad \text{M(Y)}= \sum_{i=1}^{n}\sum_{j=1}^{m} y_j p_{ij} \qquad (3.7)$$

Условным математическим ожиданием дискретной случайной величиной Y при условии, что X=x_i определяется равенством:

$$M(Y \mid X = x_i) = \sum_{j=1}^{m} y_j \cdot p(y_j \mid x_i) \qquad (3.8)$$

Условным математическим ожиданием дискретной случайной величиной X при условии, что $Y=y_j$ определяется равенством:

$$M(X \mid Y = y_j) = \sum_{i=1}^{n} x_i \cdot p(x_i \mid y_j) \qquad (3.9)$$

Дисперсией системы случайных величин (X, Y) называется величина, которая вычисляется по формуле:

$$D(X) = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i - M(x))^2 p_{ij} \qquad D(Y) = \sum_{i=1}^{n}\sum_{j=1}^{m}(y_j - M(y))^2 p_{ij} \qquad (3.10)$$

Корреляционный момент и коэффициент корреляции

Для характеристики связи между величинами X и Y случит корреляционный момент K_{xy} или *ковариация – cov (X, Y),* которая вычисляется по формуле:

Определение. Ковариацией (корреляционным моментом) компонент *X* и *Y* называется математическое ожидание произведения отклонений этих величин от своих математических ожиданий:

$$cov(X,Y) = M\left[(X - M(X)) \cdot (Y - M(Y))\right] = M(XY) - M(X) \cdot M(Y), \qquad (3.11)$$

Ковариацию удобно вычислять по формуле:

$$cov\,(X, Y) = M\,(X, Y) - M(X)\,M(Y) \qquad (3.12)$$

Если X и Y независимые величины, то ковариация равна 0.

Если *cov* (X, Y) не равна 0, то X и Y зависимые, в этом случае величины называют коррелированными.

Если случайные величины коррелированные, то определяется коэффициент корреляции по формуле:

$$r_{XY} = \frac{\mathrm{cov}(X,Y)}{\sigma_x \cdot \sigma_y} \qquad (3.13)$$

Коэффициент корреляции характеризует степень линейной зависимости случайных величин X и Y.

Свойства коэффициента корреляции:

1.$r_{xy} \in [-1;1]$*;*

2. Если X и Y независимые величины, то $r_{xy}=0$*;*

3. Если X и Y связаны линейной зависимостью Y=aX+b, то $|r_{xy}|=1$*;*

4. Если $|r_{xy}|=1$*, то X и Y связаны линейно функциональной зависимостью.*

Пример 3.1

Задана таблица распределения дискретной двумерной случайной величины:

Y \ X	1	2
1	0,16	0,28
2	0,12	0,11
3	0,08	0,25

Найти:

1. Функцию распределения системы (X, Y).
2. Безусловные законы распределения.
3. Вероятность события P {X$\le$Y}.
4. Условные законы распределения: X при y= 3, Y при x= 1.
5. Найти числовые характеристики компонент системы.
6. Проверить зависимость компонент X, Y.
7. Найти ковариацию X, Y.
8. Найти коэффициент корреляции.

Решение.

1. В соответствии с формулой $F(x, y) = \sum\sum p_{ij}$ получаем:

при x$\le$1 и y$\le$1 F (X, Y) = 0;

при 1$\le$x <2 и y$\le$1 F (X, Y) = 0;

при x >2 и y$\le$1 F (X, Y) = 0;

при 1< x$\le$2 и 1< y$\le$2 F (X, Y) = P{X=1, Y=1} = 0,16;

при 1< x$\le$2 и 2< y$\le$3 F (X, Y) = P{X=1, Y=1} + P{X=1, Y=2} = 0,16 + 0,12 = 0,28;

при 1< x$\le$2 и при x >2 F(X,Y) = P{X=1, Y=1} + P{X=1, Y=2} + P {X=1, Y=3} = 0,16 + 0,12 + 0,08 = 0,36;

при x >2 и 1 < y$\le$2 F(X,Y) = P{X=1, Y=1}+ P{X=2, Y=1} = 0,16 + 0,28=0,44;

при x >2 и 2 < y$\le$3 F(X,Y) = P{X=1, Y=1}+ P{X=2, Y=1}+ P{X=2, Y=1}+ P{X=2, Y=2} = 0,16+ 0,28+0,12+0,11=0,67;

при x >2 и при y >3 F(X,Y) = P{X=1, Y=1}+ P{X=2, Y=1} + P{X=2, Y=1} + P{X=2, Y=2} + P{X=1, Y=3}+ P{X=2, Y=3}= = 0,16+ 0,28+0,12+0,11+0,08+0,25 = 1

Таким образом, функция распределения будет иметь вид:

$$F(X,Y)=\begin{cases} \begin{array}{cccc} x/y & x\le 1 & 1<x\le & x>2 \\ y\le 1 & 0 & 0 & 0 \\ 1<y\le 2 & 0 & 0{,}16 & 0{,}44 \\ 2<y\le 3 & 0 & 0{,}28 & 0{,}67 \\ y>3 & 0 & 0{,}36 & 1 \end{array} \end{cases}$$

2. Случайная величина X принимает два значения x_1 =1 и x_2 =2. Вероятность этих значений соответственно равны:

$p_1=\{p_{11}+p_{21}+p_{31}\}$ = 0,16+0,12+0,08=0,36;

$p_2=\{p_{12}+p_{22}+p_{32}\}$ = 0,28+0,11+0,25=0,64.

Следовательно, безусловный закон распределения случайной величины X имеет вид:

X	1	2
p	0,36	0,64

Случайная величина Y принимает три значения x_1 =1 и x_2 =2, x_3 =3. Вероятность этих значений соответственно равны:

$p_1 = \{p_{11}+p_{12}\}$ = 0,16 + 0,28 = 0,44;

$p_2 = \{p_{21}+p_{22}\}$ = 0,12 + 0,11 = 0,23;

$p_3 = \{p_{31}+p_{32}\}$ = 0,08 + 0,25 = 0,33.

Следовательно, безусловный закон распределения случайной величины Y имеет вид:

Y	1	2	3
p	0,44	0,23	0,33

3. Вероятность события $\{X\le Y\}$ равна:

$P\,\{X\le Y\}$ = 0,16 + 0,12 + 0,08 + 0,11 + 0,25 = 0,72.

4. Условные вероятности найдем по формулам:

$$P\{X\mid Y=y_j\}=\frac{P\{X=x_i\ Y=y_j\}}{P\{Y=y_j\}}\ \text{и}\ P\{Y\mid X=x_i\}=\frac{P\{X=x_i\ Y=y_j\}}{P\{X=x_i\}}.$$

Так как P{Y=3} =0,33, то

$$p_1\,(x_1)=\frac{0{,}08}{0{,}33}=\frac{8}{33}$$

$$p_2\,(x_2)=\frac{0{,}25}{0{,}33}=\frac{25}{33}$$

Следовательно, условный закон распределения случайной величины Y при условии, что x=1 имеет вид:

X	1	2
p{Y=3}	$\frac{8}{33}$	$\frac{25}{33}$

Так как P{X=1} =0,36, то

$p_1\ (y_1) = \frac{0{,}16}{0{,}36} = \frac{4}{9}$;

$p_2(y_2) = \frac{0{,}12}{0{,}36} = \frac{1}{3}$;

$p_3(y_3) = \frac{0{,}08}{0{,}36} = \frac{2}{9}$.

Следовательно, условный закон распределения случайной величины Y при условии, что x=1 имеет вид:

Y	1	2	3
p{X=1}	$\frac{4}{9}$	$\frac{1}{3}$	$\frac{2}{9}$

5. Используя формулы для нахождения числовых характеристик дискретных случайных величин, найдем: M(X), M(Y), D(X), D(Y), σ(X), σ(Y).

$$M(X) = 1\cdot 0{,}36 + 2\cdot 0{,}64 = 0{,}36 + 1{,}28 = 1{,}64$$

$$M(Y) = 1\cdot 0{,}44 + 2\cdot 0{,}23 + 3\cdot 0{,}33 = 0{,}44 + 0{,}46 + 0{,}99 = 1{,}89$$

$$D(X) = (1\cdot 0{,}36 + 2^2\cdot 0{,}64) - (1{,}64)^2 = (0{,}36 + 2{,}56) - 2{,}6896 = 2{,}92 - 2{,}6896 = 0{,}2304$$

$$D(Y) = (1\cdot 0{,}44 + 2^2\cdot 0{,}23 + 3^2\cdot 0{,}33) - (1{,}89)^2 = (0{,}44 + 0{,}92 + 2{,}97) - 3{,}5721 = 0{,}7579$$

$$\sigma(X) = \sqrt{0{,}2304} \approx 0{,}48$$

$$\sigma(Y) = \sqrt{0{,}7579} \approx 0{,}87$$

6. Условие зависимости дискретных случайных величин X и Y имеет вид:

$$P\{X=x_i,\ Y=y_j\} = P\{X=x_i\}\cdot P\{Y=y_j\}$$

По условию P {X=1; Y=1} = 0,16

P{X=1} = 0,36

P{Y=1} = 0,44 отсюда 0,16 $\neq$ 0,36 $\cdot$0,44.

Вывод: компоненты системы {X, Y} – зависимы.

7. Определим ковариацию по формуле *cov (X, Y) = M (X, Y) – M(X) M(Y)*:

$cov(X,Y) = (1\cdot 1\cdot 0{,}16 + 1\cdot 2\cdot 0{,}12 + 1\cdot 3\cdot 0{,}08 + 2\cdot 1\cdot 0{,}28 + 2\cdot 2\cdot 0{,}11 + 2\cdot 3\cdot 0{,}25) - 1{,}64\cdot 1{,}89 = 3{,}14 - 3{,}0996 = 0{,}0404$

8. Найдем коэффициент корреляции:

$$r_{XY} = \frac{\mathrm{cov}(X,Y)}{\sigma_x\cdot\sigma_y} = \frac{0{,}0404}{0{,}48\cdot 0{,}87} \approx 0{,}0967$$

3.2. Непрерывная двумерная случайная величина

Плотность распределения вероятностей системы случайных величин.

Определение. Плотностью распределения вероятностей системы двух непрерывных случайных величин (X, Y) называется вторая смешанная производная ее функции распределения, определяемая формулой:

$$f(x,y) = \frac{d^2F(x,y)}{dx\cdot dy} = F''_{xy}(x,y) \qquad (3.14)$$

Свойства двумерной плотности распределения вероятностей:

1. $f(x,y)\ge 0;$
2. $\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f(x,y)dxdy = 1$
3. $P\{(X,Y)\in D\} = \iint\limits_D f(x,y)dxdy$; *где D – произвольная область;*
4. $\int\limits_{-\infty}^{+\infty} f(x,y)dy = f_X(x)$ $\qquad$ $\int\limits_{-\infty}^{+\infty} f(x,y)dx = f_Y(y)$

Независимые случайные величины

Для двух непрерывных случайных величин X и Y необходимым и достаточным условием их **независимости** является равенство:

$$f(x,y) = f_X(x)\cdot f_Y(y) \qquad (3.15)$$

Условные законы распределения

В системе двух непрерывных случайных величин (X, Y), условным законом распределения СВ Y при условии, что X=x называется совокупность вероятностей:

$$f\{y\,|\,x\} = \frac{f\{x\ y\}}{f_X\{x\}} \qquad (3.16)$$

В системе двух непрерывных случайных величин (X,Y), условным законом распределения СВ X при условии, что Y=y называется совокупность вероятностей:

$$f\{x \mid y\} = \frac{f\{x\ y\}}{f_Y\{y\}} \tag{3.17}$$

Математическое ожидание и дисперсия системы случайных величин

Математическим ожидание системы случайных величин (X, Y) называется совокупность двух математических ожиданий *M(X) и M(Y)* определяемых равенством:

$$M(X) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x \cdot f(x,y)dxdy \quad M(Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y \cdot f(x,y)dxdy \tag{3.18}$$

Условным математическим ожиданием непрерывной случайной величиной Y при условии, что X=x определяется равенством:

$$M(Y \mid x) = \int_{-\infty}^{+\infty} y \cdot f(y \mid x)dy \tag{3.19}$$

Условным математическим ожиданием непрерывной случайной величиной X при условии, что Y=y определяется равенством:

$$M(X \mid y) = \int_{-\infty}^{+\infty} x \cdot f(x \mid y)dx \tag{3.20}$$

Дисперсией системы непрерывных случайных величин (X, Y) называется величина, которая вычисляется по формуле:

$$D(X) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^2 \cdot f(x,y)dxdy - M^2(X)$$

(3.21)

$$D(Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y^2 \cdot f(x,y)dxdy - M^2(Y)$$

Ковариацией системы двух непрерывных случайных величин называется величина, которая определяется формулой:

$$K_{XY} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x - M(X)) \cdot (y - M(Y)) \cdot f(x,y)dxdy$$

(3.22)

$$K_{XY} = \iint xy \cdot f(x,y)dxdy - M(X)M(Y)$$

Коэффициент корреляции: $r_{XY} = \dfrac{K_{XY}}{\sigma_x \cdot \sigma_y}$ (3.23)

Пример 3.2

Совместное распределение случайных величин X и Y задано плотностью распределения вероятностей:

$$f(x,y) = \begin{cases} C\cdot(y-yx), & (x,y)\in D, \\ 0, & (x,y)\notin D. \end{cases}$$

$$\text{где } D = \{0 \le x \le 1; \quad 0 \le y \le 1\}$$

Найти:

1. Коэффициент C.
2. Плотность распределения отдельных компонент X и Y.
3. Вероятность попадания случайной точки (x, y) в область $D_1=\{0{,}7 \le x \le 3;\ 0 \le y \le 0{,}3\}$.
4. Условные плотности отдельных компонент X и Y.
5. Найти числовые характеристики компонент X и Y.
6. Проверить зависимость компонент X, Y.
7. Найти ковариацию X, Y.
8. Найти коэффициент корреляции.

Решение.

1. Коэффициент C найдем из условия нормировки: $\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f(x,y)dxdy = 1$

$$\int_0^1\int_0^1 \mathrm{C}\cdot(y-xy)dxdy = \int_0^1 Cdx\int_0^1 (y-xy)dy = C\int dx\left(\frac{y^2}{2}-\frac{xy^2}{2}\right)\Big|_0^1 =$$

$$= C\int_0^1\left(\frac{1}{2}-\frac{x}{2}\right)dx = C\left(\frac{x}{2}-\frac{x^2}{4}\right)\Big|_0^1 = \frac{C}{4}$$

$$\frac{C}{4} = 1 \Rightarrow C = 4$$

Отсюда.

$$f(x,y) = \begin{cases} 4\cdot(y-yx), & (x,y)\in D, \\ 0, & (x,y)\notin D. \end{cases}$$

$$\text{где } D = \{0 \le x \le 1; \quad 0 \le y \le 1\}$$

2. Найдем плотности распределения отдельных компонент X и Y.

$$f_X(x) = \int\limits_0^1 4(y-xy)\cdot dy = \frac{4y^2}{2}\Bigg|_0^1 - \frac{4xy^2}{2}\Bigg|_0^1 = 2-2x$$

Плотность распределения компоненты X имеет вид: $f_X(x)=\begin{cases}2-2x, & x\in[0;1],\\ 0, & x\notin[0;1].\end{cases}$

$$f_Y(y)=\int_0^1 4(y-xy)\cdot dx=4yx\Big|_0^1-\frac{4x^2y}{2}\Big|_0^1=4y-2y=2y$$

Плотность распределения компоненты Y имеет вид: $f_Y(y)=\begin{cases}2y, & y\in[0;1],\\ 0, & y\notin[0;1].\end{cases}$

3.Для нахождения попадания случайной точки (x, y) в область D_1 воспользуемся формулой:

$$P\{(X,Y)\in D\}=\iint_D f(x,y)dxdy$$

$$P\{(X,Y)\in D\}=\iint_{D_1} f(x,y)dxdy=\int_{0,7}^{1}dx\int_0^{0,3}4(y-xy)dy+\int_1^3 dx\int_0^{0,3}0dy=\int_{0,7}^{1}dx\left(2y^2-2xy^2\right)\Big|_0^{0,3}=$$

$$=\int_{0,7}^{1}\left(0{,}18-0{,}18x\right)dx=\left(0{,}18x-0{,}09x^2\right)\Big|_{0,7}^{1}=[\left(0{,}18\cdot1-0{,}09\cdot1\right)-\left(0{,}18\cdot0{,}7-0{,}09\cdot0{,}49\right)]=$$

$$=0{,}09-(0{,}126-0{,}0441)=0{,}09-0{,}0819=0{,}0081.$$

4. Для нахождения условной плотности распределения *f (x | y)* воспользуемся формулой: $f\{x\,|\,y\}=\dfrac{f\{x\ y\}}{f_Y\{y\}}$

$$f\{x\,|\,y\}=\frac{f\{x\ y\}}{f_Y\{y\}}=\frac{4(y-xy)}{2y}=\frac{4y(1-x)}{2y}=2(1-x)$$

Условная плотность распределения *f (y| x)* имеет вид:

$$f\ (y\,|\,x)=\begin{cases}2y, & x\in[0;1]; \quad y\in[0;1],\\ 0, & x\notin[0;1]; \quad y\notin[0;1].\end{cases}$$

Для нахождения условной плотности распределения *f(y|x)* воспользуемся формулой: $f\{y\,|\,x\}=\dfrac{f\{x\ y\}}{f_X\{x\}}$

$$f\{y\,|\,x\}=\frac{f\{x\ y\}}{f_X\{x\}}=\frac{4(y-xy)}{2(1-x)}=\frac{4y(1-x)}{2(1-x)}=2y$$

Условная плотность распределения *f (x | y)* имеет вид:

$$f\ (x\,|\,y)=\begin{cases}2-2x, & x\in[0;1]; \quad y\in[0;1],\\ 0, & x\notin[0;1]; \quad y\notin[0;1].\end{cases}$$

5. Используя формулы для нахождения числовых характеристик непрерывных случайных величин, найдем: M(X), M(Y), D(X), D(Y), σ(X), σ(Y).

$$M(X)=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} x\cdot f(x,y)dxdy$$

$$M(X)=\int\limits_0^1\int\limits_0^1 4x\cdot(y-xy)dxdy=\int\limits_0^1 4x\cdot dx\int\limits_0^1(y-xy)dy=\int\limits_0^1 4xdx\left(\frac{y^2}{2}-\frac{xy^2}{2}\right)\Bigg|_0^1=4\int\limits_0^1\left(\frac{x}{2}-\frac{x^2}{2}\right)dx=$$

$$=\left(x^2-\frac{2x^3}{3}\right)\Bigg|_0^1=1-\frac{2}{3}=\frac{1}{3}$$

$$\mathrm{M(Y)}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{y\cdot f(x,y)dxdy}$$

$$\mathrm{M(Y)}=\int_0^1\int_0^1 \mathrm{4y\cdot(y-xy)dxdy}=\int_0^1 4\cdot\mathrm{dx}\int_0^1(\mathrm{y^2-xy^2})\mathrm{dy}=\int_0^1 4\mathrm{dx}\left(\frac{\mathrm{y^3}}{3}-\frac{\mathrm{xy^3}}{3}\right)\Big|_0^1=$$

$$=4\int_0^1\left(\frac{1}{3}-\frac{\mathrm{x}}{3}\right)\mathrm{dx}=4\cdot\left(\frac{\mathrm{x}}{3}-\frac{\mathrm{x^2}}{6}\right)\Big|_0^1=\frac{4}{3}-\frac{2}{3}=\frac{2}{3}$$

$$D(X)=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} x^2\cdot f(x,y)dxdy-M^2(X)$$

$$D(X)=\int\limits_0^1\int\limits_0^1 4x^2\cdot(y-xy)dxdy-\left(\frac{1}{3}\right)^2=\int\limits_0^1 4x^2\cdot dx\int\limits_0^1(y-xy)dy-\frac{1}{9}=\int\limits_0^1 4x^2dx\left(\frac{y^2}{2}-\frac{xy^2}{2}\right)\Bigg|_0^1-\frac{1}{9}=$$

$$=4\int\limits_0^1\left(\frac{x^2}{2}-\frac{x^3}{2}\right)dx-\frac{1}{9}=4\left(\frac{x^3}{6}-\frac{x^4}{8}\right)\Bigg|_0^1-\frac{1}{9}=\left(\frac{2}{3}-\frac{1}{2}\right)-\frac{1}{9}=\frac{1}{18}$$

$$D(Y)=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} y\cdot f(x,y)dxdy-M^2(Y)$$

$$D(Y)=\int\limits_0^1\int\limits_0^1 4y^2\cdot(y-xy)dxdy-\left(\frac{2}{3}\right)^2=\int\limits_0^1 4\cdot dx\int\limits_0^1(y^3-xy^2)dy-\left(\frac{4}{9}\right)=\int\limits_0^1 4\cdot dx\left(\frac{y^4}{4}-\frac{xy^4}{4}\right)\Bigg|_0^1-\left(\frac{4}{9}\right)=$$

$$=\int\limits_0^1(1-x)dx-\left(\frac{4}{9}\right)=\left(x-\frac{x^2}{2}\right)\Bigg|_0^1-\left(\frac{4}{9}\right)=\left(1-\frac{1}{2}\right)-\frac{4}{9}=\frac{1}{2}-\frac{4}{9}=\frac{1}{18}$$

$$\sigma(X)=\sqrt{\frac{1}{18}}=\frac{1}{3\sqrt{2}}\qquad\qquad \sigma(Y)=\sqrt{\frac{1}{18}}=\frac{1}{3\sqrt{2}}.$$

6. Для проверки зависимости случайных величин X и Y воспользуемся формулой:

$$f(x,y)=f_X(x)\cdot f_Y(y)$$

$$4(y-xy)=2y\cdot(2-2x)\Rightarrow 4y-4xy=4y-4xy$$

Вывод: компоненты системы {X, Y} – независимы.

7. Определим ковариацию по формуле:

$$K_{XY} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x - M(X)) \cdot (y - M(Y)) \cdot f(x,y)dxdy$$

$$K_{XY} = \int_0^1\int_0^1 xy \cdot 4(y - xy)dxdy - \frac{1}{3}\cdot\frac{2}{3} = \int_0^1 4xdx\int_0^1 (y^2 - xy^2)dy - \frac{2}{9} =$$

$$= \int_0^1 4xdx\left(\frac{y^3}{3} - \frac{xy^3}{3}\right)\Big|_0^1 - \frac{2}{9} = 4\int_0^1\left(\frac{x}{3} - \frac{x^2}{3}\right)dx - \frac{2}{9} =$$

$$= 4\left(\frac{x^2}{6} - \frac{x^3}{9}\right)\Big|_0^1 - \frac{2}{9} = 4\left(\frac{1}{6} - \frac{1}{9}\right) - \frac{2}{9} = \frac{2}{9} - \frac{2}{9} = 0$$

8. Коэффициент корреляции найдем по формуле: $r_{XY} = \dfrac{K_{XY}}{\sigma_x \cdot \sigma_y}$

$$r_{XY} = \frac{0}{\frac{1}{3\sqrt{2}}\cdot\frac{1}{3\sqrt{2}}} = 0$$

Примеры для самостоятельного решения

1. Дана дискретная двумерная случайная величина (*X*, *Y*). Найти:

а) безусловные законы распределения компонент *X* и *Y*;

б) математические ожидания составляющих компонент *M* (*X*), *M*(*Y*), ковариацию компонент *cov* (*X*, *Y*);

в) условный закон распределения *X* при *Y* = *C* и найти *M* (*X* / *Y* = *C*);

г) закон распределения случайной величины *T* = *aX* + *b*, математическое ожидание *M*(*T*) и дисперсию *D*(*T*);

д) закон распределения случайной величины *Z* = *X* + *Y*; математическое ожидание *M*(*Z*) и дисперсию *D*(*Z*);

$X \backslash Y$	–7	–3	0	1	2	4
–2	0,08	0,1	0,05	0,02	0,03	0,13
–1	0,01	0,05	0,06	0,03	0,05	0,04
0	0	0,1	0,07	0,03	0,01	0,02
3	0,01	0,05	0,02	0,02	0,01	0,01

a = 2, *b* = –1, *C* = 0

2. Совместное распределение случайных величин X и Y задано плотностью распределения вероятностей:

$$f(X,Y) = \begin{cases} C(1 - xy^3), & (x,y) \in D, \\ 0, & (x,y) \notin D \end{cases}$$

$$D = \{-1 \le x \le 1; \quad -1 \le y \le 1\}$$

Найти:

1. Коэффициент С.
2. Плотность распределения отдельных компонент X и Y.
3. Вероятность попадания случайной точки (x, y) в область

$D_1 = \{0,7 \leq x \leq 3;\ 0 \leq y \leq 0,3\}$.

4. Условные плотности отдельных компонент X и Y.
5. Найти числовые характеристики компонент X и Y.
6. Проверить зависимость компонент X, Y.
7. Найти ковариацию X, Y.
8. Найти коэффициент корреляции.

Тема 4. Закон больших чисел

4.1. Закон больших чисел

Практика изучения случайных явлений показывает, что хотя результаты отдельных наблюдений, даже проведенных в одинаковых условиях, могут сильно отличаться, в то же время средние результаты для достаточно большого числа наблюдений устойчивы и слабо зависят от результатов отдельных наблюдений.

Теоретическим обоснованием этого замечательного свойства случайных явлений является *закон больших чисел*. Названием "закон больших чисел" объединена группа теорем, устанавливающих устойчивость средних результатов большого количества случайных явлений и объясняющих причину этой устойчивости.

Простейшая форма закона больших чисел, и исторически первая теорема этого раздела – *теорема Бернулли*, утверждающая, что если вероятность события одинакова во всех испытаниях, то с увеличением числа испытаний частота события стремится к вероятности события и перестает быть случайной.

Теорема Пуассона утверждает, что частота события в серии независимых испытаний стремится к среднему арифметическому его вероятностей и перестает быть случайной.

Предельные теоремы теории вероятностей, теоремы *Муавра-Лапласа* объясняют природу устойчивости частоты появлений события. Природа эта состоит в том, что предельным распределением числа появлений события при неограниченном возрастании числа испытаний (если вероятность события во всех испытаниях одинакова) является *нормальное распределение.*

Центральная предельная теорема объясняет широкое распространение *нормального закона* распределения. Теорема утверждает, что всегда, когда случайная величина образуется в результате сложения большого числа независимых случайных величин с конечными дисперсиями, закон распределения этой случайной величины оказывается практически *нормальным* законом.

Теорема, приведенная ниже под названием "***Закон больших чисел***" утверждает, что при определенных, достаточно общих, условиях, с увеличением числа случайных величин их среднее арифметическое стремится к среднему арифметическому математических ожиданий и перестает быть случайным.

Теорема Ляпунова объясняет широкое распространение *нормального закона* распределения и поясняет механизм его образования. Теорема позволяет утверждать, что всегда, когда случайная величина образуется в результате сложения большого числа независимых случайных величин, дисперсии которых малы по сравнению с дисперсией суммы, закон распределения этой случайной величины оказывается практически *нормальным* законом. А, поскольку, случайные величины всегда порождаются бесконечным количеством причин, и чаще всего ни одна из них не имеет дисперсии, сравнимой с дисперсией самой случайной величины, то большинство встречающихся в практике случайных величин подчинено нормальному закону распределения.

Закон больших чисел. Если случайные величины x_1, x_2, …, x_n, … попарно независимы и $\lim\limits_{n\to\infty}\frac{1}{n^2}\sum\limits_{i=1}^{n}D(X_n)=0$, то для любого ε > 0 выполняется условие:

$$\lim_{n\to\infty}P\left(\left|\frac{X_1+X_2+...+X_n}{n}-\frac{M(X_1)+M(X_2)+...+M(X_n)}{n}\right|<\varepsilon\right)=1 \qquad (4.1)$$

Теорема Ляпунова.

Пусть x_1, x_2, …, x_n, … – неограниченная последовательность независимых случайных величин с математическими ожиданиями m_1, m_2, …, m_n, … и дисперсиями s_1^2, s_2^2, …, s_n^2 ….

Обозначим $Y_n=\sum\limits_{i=1}^{n}X_i,\quad M_n=\sum\limits_{i=1}^{n}m_i,\quad S_n^2=\sum\limits_{i=1}^{n}\sigma_i^2\quad Z_n=\frac{Y_n-M_n}{S_n^2}$, тогда

$$\lim_{n\to\infty}P(\alpha<Z_n<\beta)=\Phi(\beta)-\Phi(\alpha) \qquad (4.2)$$

для любых действительных чисел α и β, где Ф(x) - функция распределения нормального закона.

В основе качественных и количественных утверждений закона больших чисел лежит *неравенство Чебышева*. Оно определяет верхнюю границу вероятности того, что отклонение значения случайной величины от ее математического ожидания больше некоторого заданного числа. Замечательно, что неравенство Чебышева дает оценку вероятности события $|X-M(X)|>\varepsilon$ для случайной величины, распределение которой неизвестно, известны лишь ее математическое ожидание и дисперсия.

Неравенство Чебышева устанавливает нижнюю границу вероятности и показывает, что эта граница зависит лишь от дисперсии.

Неравенство Чебышева. Если случайная величина x имеет дисперсию, то для любого e> 0 справедливо неравенство

$$P\{|X-M(X)|>\varepsilon\}\le\frac{D(X)}{\varepsilon^2}$$
$$P\{|X-M(X)|\le\varepsilon\}>1-\frac{D(X)}{\varepsilon^2} \qquad (4.3)$$

где Mx и Dx – математическое ожидание и дисперсия случайной величины x.

Пример 4.1

Устройство состоит из 10 независимо работающих элементов. Вероятность отказа каждого элемента за время Т равна 0,05. С помощью неравенства Чебышева оценить вероятность того, что абсолютная величина разности между числом отказавших элементов и средним числом (математическим ожиданием) отказов за время Т окажется: а) меньше двух; б) не меньше двух.

Решение.

а) Обозначим через X дискретную случайную величину – число отказавших элементов за время Т. Найдем математическое ожидание и дисперсию, если

n = 10, p=0,05, q=0,95.

$M(X) = n \cdot p = 10 \cdot 0{,}05 = 0{,}5$

$D(X) = n \cdot p \cdot q = 10 \cdot 0{,}05 \cdot 0{,}95 = 0{,}475$

Воспользуемся неравенством Чебышева: $P\{|X - M(X)| \le \varepsilon\} > 1 - \dfrac{D(X)}{\varepsilon^2}$

Подставив сюда M(X) = 0,5; D(X) = 0,475, ε = 2, получим:

$$P\{|X - 0{,}5| \le 2\} > 1 - \frac{0{,}475}{4} = 0{,}88$$

б) События $|X - 0{,}5| < 2$ и $|X - 0{,}5| \ge 2$ противоположны, поэтому сумма их вероятностей равна единице. Следовательно,

$P\{|X - 0{,}5| > 2\} \le 1 - 0{,}88 = 0{,}12$.

Ответ. а) 0,88 б) 0,12.

Пример 4.2

Дискретная случайная величина X задана законом распределения:

X	0,3	0,6
p	0,2	0,8

Используя неравенство Чебышева, оценить вероятность того, что

$|X - M(X)| < 0{,}2$.

Решение.

Найдем математическое ожидание и дисперсию величины X:

$$M(X) = 0{,}3 \cdot 0{,}2 + 0{,}6 \cdot 0{,}8 = 0{,}54$$

$$D(X) = (0,3^2 \cdot 0{,}2 + 0,6^2 \cdot 0{,}8) - 0{,}54^2 = (0{,}018 + 0{,}288) - 0{,}2916 =$$
$$= 0{,}0144$$

Воспользуемся неравенством Чебышева в форме $P\{|X-M(X)|\le\varepsilon\}>1-\frac{D(X)}{\varepsilon^2}$

Подставляя M(X) = 0,54, D(X) = 0,0144, ε = 0,2, окончательно получим:

$$P\{|X-0{,}54|\le 0{,}2\}>1-\frac{0{,}0144}{0{,}04}=0{,}64.$$

Ответ. 0,64

Неравенство Маркова. Для любой неотрицательной случайной величины X, имеющей конечное математическое ожидание, и $\forall\varepsilon>0$, вероятность того, что она примет значение, превосходящее ε, не больше чем отношения математического ожидания этой случайной величины M(X) к значению ε:

$$P\{X>\varepsilon\}\le\frac{M(X)}{\varepsilon}\quad или\quad P\{X\le\varepsilon\}>1-\frac{M(X)}{\varepsilon} \tag{4.4}$$

В отличие от неравенства Чебышева, неравенство Маркова применяется только для оценки вероятности неотрицательных случайных величин, для которых неизвестен закон распределения.

Пример 4.3

Среднее число молодых специалистов ежегодно направляемых в аспирантуру экономических вузов, составляет 200 человек. Пользуясь неравенством Маркова, оценить вероятность того, что в данном году будет направлено в аспирантуру не более 220 человек.

Решение.

По условию задачи M(X) = 200, а ε=220. Применяя неравенство Маркова

$P\{X\le\varepsilon\}>1-\frac{M(X)}{\varepsilon}$, получим, что

$$P\{X\le 220\}>1-\frac{200}{220}\Rightarrow P\{X\le 220\}>1-0{,}091\ \Rightarrow P\{X\le 220\}>0{,}909$$

Итак, вероятность того, что в этом году в аспирантуру экономического вуза направят не белее 220 человек больше 0,909.

Ответ. $P\{X\le 220\}>0{,}909$

Пример 4.4

Средний срок службы холодильника пять лет. Оценить вероятность того. что данный холодильник не прослужит более 20 лет.

Решение.

Пусть X – срок службы холодильника. Используя неравенство Маркова, оценим вероятность:

$$P\{X\le 20\}>1-\frac{5}{20}\Rightarrow\quad P\{X\le 20\}>1-0{,}25\quad\Rightarrow\quad P\{X\le 20\}>0{,}75.$$

Ответ. P > 0,75.

Пример 4.5

Сумма вкладов в сбербанке составляет $4 \cdot 10^6$ рублей. Вероятность того, что случайно выбранный вклад не превышает 1000 рублей, равна 0,8. Оценить число вкладчиков в этом банке.

Решение.

Пусть в банке n вкладчиков. Обозначим через X величину случайно снятого вклада. Найдем математическое ожидание вкладов отдельных вкладчиков $M(X) = \frac{4 \cdot 10^6}{n}$. По условию задачи $\varepsilon =$ 1000=10^3. Следовательно, P |x <1000| = 0,8. Используя неравенство Маркова, эту вероятность запишем в виде:

$$0{,}8 = P\{X \le 10^3\} > 1 - \frac{4 \cdot 10^6}{10^3 \cdot n} \Rightarrow \quad 0{,}2 < \frac{4000}{n} \quad \Rightarrow \quad 0{,}2n < 4000 \quad \Rightarrow \quad n < 20\,000.$$

Таким образом, в этом банке могут содержать вклады не более 20 000 вкладчиков.

Ответ. n <20 000.

Теорема Чебышева. Если X_1, X_2, …, X_n – последовательность попарно независимых случайных величин, имеющих конечные математическое ожидание и дисперсию, ограниченную некоторой константой C: $D(X_1) \le C$,

$D(X_2) \le C$, … , $D(X_n) \le C$, …, то среднее арифметическое этих случайных величин сходится по вероятности к среднему арифметическому их математических ожиданий:

$$\frac{1}{n}\sum_{i=1}^{n} X_i \qquad\qquad \frac{1}{n}\sum_{i-1}^{n} M(X_i) \text{ для } \forall \varepsilon > 0$$

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} M(X_i)\right| < \varepsilon\right) = 1 \qquad (4.5)$$

Формула (3) означает, что среднее арифметическое случайных величин X_1, X_2, …, X_n при $n \to \infty$ мало отличается от среднего арифметического их математических ожиданий $M(X_1)$, $M(X_2)$, …, $M(X_n)$, т.е. обладает свойством устойчивости.

Пример 4.6

Для определения средней продолжительности горения лампы в партии из 200 одинаковых ящиков взяли на выбор по одной лампочке из каждого ящика. Оценить вероятность того, что продолжительность горения выборки (200 шт.) отличается от средней продолжительности горения всех ламп (партии) по модулю меньше, чем на 5 часов, если по ГОСТу стандартное отклонения не превышает 7 часов.

Решение.

Пусть X_i – продолжительность горения электролампы взятой из i – го ящика. По условию задачи $\sigma \le$7ч., тогда $D_i = \sigma_i^2 \le 7^2$ ч^2.

Общая средняя продолжительность горения в выборке $\frac{x_1+x_2+..+x_{200}}{200}$, средняя продолжительность горения одной лампы в партии $m=\frac{Mx_1+Mx_2+Mx_{200}}{200}$.

Оценим вероятность снизу: $P\left(\left|\frac{1}{200}\sum_{i=1}^{200}X_i-m\right|<5\right)$.

Поскольку лампы стандартные, то по теореме Чебышева с параметрами задачи: $\sigma_i^2=$ 49,

ε =5, n=200 получаем: $P\geq 1-\frac{49}{200\cdot 5^2}=1-0{,}0098=0{,}9902$.

Ответ. 0,9902

Пример 4. 7

Последовательность независимых случайных величин X_1, X_2, ..., X_n, … задана законом распределения:

X_n	$-n\alpha$	0	$n\alpha$
p	$\frac{1}{2^n}$	$1-\frac{1}{2^{n-1}}$	$\frac{1}{2^n}$

Применима ли к заданной последовательности теорема Чебышева?

Решение.

1. Поскольку случайные величины X_n независимы, то они подавно и попарно независимы, т.е. первое требование теоремы Чебышева выполняется.

2. Найдем $M(X_n)$ $=-n\alpha\cdot\frac{1}{2^n}+0\cdot 1-\frac{1}{2^{n-1}}+n\alpha\cdot\frac{1}{2^n}=$ 0 т.е. требование конечности математических ожиданий выполняется.

Остается проверить выполнимость требования равномерной ограниченности дисперсий.

По формуле $D(X_n) = M(X_n^2) - M^2(X_n)$ находим дисперсию. Так как $M(X_n) =0$, то получаем:

$$D(X)=(-n\alpha)^2\cdot\frac{1}{2^n}+0\cdot\left(1-\frac{1}{2^{n-1}}\right)+(n\alpha)^2\cdot\frac{1}{2^n}=\frac{(n\alpha)^2+(n\alpha)^2}{2^n}=\frac{2(n\alpha)^2}{2^n}=\frac{n^2}{2^{n-1}}\cdot\alpha^2$$

$$D(X)=\frac{n^2}{2^{n-1}}\cdot\alpha^2$$

Предположим, что n изменяется непрерывно. Чтобы подчеркнуть это допущение, обозначим n через (x), и исследуем на экстремум функцию: $\varphi(x) = \dfrac{x^2}{2^{x-1}}$.

$$\varphi(x) = \frac{x^2}{2^{x-1}} \Rightarrow \varphi'(x) = \frac{2x(2^{x-1}) - x^2 \cdot 2^{x-1} \cdot \ln 2}{(2^{x-1})^2} = \frac{x \cdot (2 - x\ln 2)}{2^{x-1}}$$

$$\frac{x \cdot (2 - x\ln 2)}{2^{x-1}} = 0$$

$$x_1 = 0 \quad 2 - x\ln 2 = 0$$

$$x\ln 2 = 2$$

$$x_2 = \frac{2}{\ln 2}$$

Отбросим первую точку как не представляющую интереса (n не принимает значения, равного нулю, что видно из условия задачи).

В точке $x_2 = \dfrac{2}{\ln 2}$ функция φ(x) имеет максимум.

Учитывая, что $\dfrac{2}{\ln 2} \approx 2{,}9$ и что n – целое положительное число, вычислим дисперсию $D(X) = \dfrac{n^2}{2^{n-1}} \cdot \alpha^2$ для ближайших к числу 2,9 (слева и справа) целых чисел, т.е. для n = 2 и

n = 3.

При n. = 2 дисперсия $D(X) = 2\alpha^2$, при n = 3 дисперсия $D(X) = \dfrac{9}{4}\alpha^2$.

Очевидно, $\dfrac{9}{4}\alpha^2 > 2\alpha^2$.

Таким образом, наибольшая возможная дисперсия равна $\dfrac{9}{4}\alpha^2$, т.е. дисперсия случайных величин X_n равномерно ограничена числом $\dfrac{9}{4}\alpha^2$.

Итак, все требования теоремы Чебышева выполняются, следовательно, к рассматриваемой последовательности эта теорема применима.

Теорема Бернулли. Пусть m_n – число успехов в n испытаниях Бернулли и p – вероятность успеха в отдельном испытании. Тогда при любом $\varepsilon > 0$ справедливо равенство:

$$\lim_{n\to\infty} P\left(\left|\frac{m_n}{n} - p\right| < \varepsilon\right) = 1 \qquad (4.6).$$

Из равенства 4.4 следует неравенство: $P\left(\left|\frac{m_n}{n}-p\right|<\varepsilon\right)\geq 1-\frac{np}{n\varepsilon^2}$ (4.7)

Пример 4.8

При контрольной проверке изготовляемых приборов было установлено, что в среднем 15 штук из 100 оказываются с теми или иными дефектами. Оценить вероятность того, что доля приборов с дефектами среди 400 изготовленных будет по абсолютной величине отличаться от математического ожидания этой доли не более, чем на 0,05.

Решение.

Для оценки вероятности используем формулу: $P\left(\left|\frac{m_n}{n}-p\right|<\varepsilon\right)\geq 1-\frac{pq}{n\varepsilon^2}$.

По условию n = 400, ε = 0,05, p= 0,15, q = 1 – 0,15=0,85

$$P\left(\left|\frac{m_n}{n}-0{,}15\right|<0{,}05\right)\geq 1-\frac{0{,}15\cdot 0{,}85}{400\cdot(0{,}05)^2}$$

$$P\left(\left|\frac{m_n}{n}-0{,}15\right|<0{,}05\right)\geq 1-\frac{0{,}1275}{400\cdot(0{,}05)^2}$$

$$P\left(\left|\frac{m_n}{n}-0{,}15\right|<0{,}05\right)\geq 1-0{,}1275$$

$$P\left(\left|\frac{m_n}{n}-0{,}15\right|<0{,}05\right)\geq 0{,}8725$$

Ответ. P$\geq$0, 8725

Пример 4.9

Вероятность того, что изделие является качественным, равна 0,9. Сколько следует проверить изделий, чтобы с вероятностью не меньшей 0,95 можно было утверждать, что абсолютная величина отклонения доли качественных изделий от 0,9 не превысит 0,01?

Решение.

По условию задачи p=0,9, q= 1 – 0,9 = 0,1, ε=0,01. Подставим в правую часть неравенства $P\left(\left|\frac{m_n}{n}-p\right|<\varepsilon\right)\geq 1-\frac{pq}{n\varepsilon^2}$ данные и найдем n.

$$1-\frac{0{,}9\cdot 0{,}1}{n\cdot 0{,}0001}\geq 0{,}95 \;\Rightarrow\; \frac{0{,}09}{0{,}0001n}\leq 0{,}05 \;\Rightarrow\; 0{,}0001n\geq\frac{9}{5} \;\Rightarrow\; n\geq\frac{1{,}8}{0{,}0001} \;\Rightarrow\; n\geq 18000$$

Ответ. n $\geq$81 000.

Центральная предельная теорема. Если случайные величины X_1, X_2, …, X_n, … попарно независимы, одинаково распределены и имеют математическое ожидание M(X) и конечную дисперсию σ^2, то в пределе при $n\to\infty$ функция распределения случайной величины $X\,\frac{\sum_{i=1}^{n}(X_i - M(X))}{\sigma\sqrt{n}}$ неограниченно приближается к нормальной функции распределения:

$$P\left(\left|\frac{\sum_{i=1}^{n}(X_i - M(X))}{\sigma\sqrt{n}}\right| < x\right) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \ell^{-\frac{t^2}{2}}\,dt \qquad (4.8)$$

$$P(k_1 \le x \le k_2) = \Phi\left(\frac{k_2 - na}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{k_1 - na}{\sigma\sqrt{n}}\right) \qquad (4.9)$$

Пример 4.10

Доходы жителей города имеют математическое ожидание 10 тыс. руб. и среднее квадратическое отклонение 2 тыс. руб. (в месяц). Найти вероятность того, что средний доход 100 случайно выбранных жителей составит от 9,5 до 10,5 тыс. руб.

Решение.

Переформулируем условие задачи для суммарного дохода: он должен составлять от 950 до 1050 тыс. руб. Используя центральную предельную теорему, получаем:

$$P(950 < S_{100} < 1050) = \Phi_0\left(\frac{1050 - 100\cdot 10}{2\sqrt{100}}\right) - \Phi_0\left(\frac{950 - 100\cdot 10}{2\sqrt{100}}\right) = 2\Phi_0(2{,}5) = 0{,}9876.$$

где Φ_0 (2,5) =0,4938 (Приложение 2)

Ответ. 0,9876

Пример 4.11

Вероятность появления нестандартной детали равна 0,05. Найти вероятность того, что среди 500 отобранных случайно деталей относительная частота появления нестандартной детали отклоняется от вероятности не более чем на 0,03.

Решение.

По условию задачи: p= 0,05, n=500, ε =0,03, q = 1 – 0,05= 0,95.

Вероятность оценим по формуле: $P\left(\left|\frac{m}{n} - p\right| \le \varepsilon\right) \approx 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right)$.

Отсюда: $P\left(\left|\frac{m}{500} - 0{,}05\right| \le \varepsilon\right) \approx 2\Phi\left(0{,}03\sqrt{\frac{500}{0{,}05\cdot 0{,}95}}\right) \approx 2\Phi(3)$

По таблице Лапласа Ф (3) = 0, 49865 (Приложение 2), тогда

2 Ф (3) = 2 · 0,49865= 0,9973.

Ответ. 0,9973.

Пример 4.12

Срок службы электрической лампы имеет показательное распределение с математическим ожиданием 1000 часов. Найти вероятность того, что средний срок службы для 100 ламп составит не менее 900 часов.

Решение.

Примем для простоты 1000 часов за единицу времени. Вспомним числовые характеристики показательного распределения: $M\xi=\frac{1}{\lambda}$, $D\xi=\frac{1}{\lambda^2}$. Отсюда следует, что среднее квадратическое отклонение совпадает с математическим ожиданием (и оба они здесь равны единице). Для суммарного срока службы, используя центральную предельную теорему, получаем:

$$P(S_{100} \geq 90) = 1 - \Phi\left(\frac{90-100}{\sqrt{100}}\right) = \frac{1}{2} + \Phi_0(1) = 0{,}8413.$$

Ответ. 0, 8413

Примеры для самостоятельного решения

1. В осветительную сеть параллельно включено 20 ламп. Вероятность того, что за время T лампа будет включена, равна 0,8. Пользуясь неравенством Чебышева, оценить вероятность того, что абсолютная величина разности между числом включенных ламп и средним числом (математическим ожиданием) включенных ламп за время T окажется:

а) меньше трех; б) не меньше трех.

2. Вероятность появления события в каждом испытании равна 1/4. Используя неравенство Чебышева, оценить вероятность того, что число X появлений события заключено в пределах от 150 до 250, если будет произведено 800 испытаний.

4. Средний вес клубня картофеля равен 120 *г*. Какова вероятность того, что наугад взятый клубень картофеля весит не более 360 *г*?

5. Средний срок службы мотора 4 года. Оценить вероятность того, что данный мотор не прослужит более 20 лет.

6. Последовательность независимых случайных величин, X_1, X_2, ..., X_n, … задана законом распределения:

X_n	a	$-a$
p	$\frac{n}{2n+1}$	$\frac{n+1}{2n+1}$

Применима ли к заданной последовательности теорема Чебышева?

7. Вероятность наличия зазубрины на металлических брусках, заготовленных для обточки, равна 0,2. Оценить вероятность того, что в партии из 1000 брусков отклонение числа пригодных брусков от800 не превышает 5%.

8. Пусть вероятность того, что покупателю обувного магазина необходимы туфли 41-го размера, равна 0,15. Оценить границы процента покупателей среди 2000 побывавших в магазине, которым нужны такие туфли, если эти границы надо гарантировать с вероятностью 0,98.

Тема 5. Элементы математической статистики

5.1. Статистическое распределение

Различают два вида совокупностей однородных объектов:

1. Генеральная совокупность – исходное множество объектов с соответствующим признаком, о котором необходимо составить представление;
2. Выборочная (выборка) совокупность – сравнительно небольшая часть генеральной совокупности, которая подлежит непосредственному исследованию.

В силу случайного попадания объектов в выборку числовые данные в выборке также случайны.

Задачей математической статистики является отбор выборочной информации, ее обработка по результатам, на основе которой делают выводы о параметрах генеральной совокупности, рассчитываются оценки числовых характеристик распределения генеральной совокупности и устанавливается степень достоверности этих оценок.

Пусть для изучения количественного признака X из генеральной совокупности извлечена выборка x_1, $x_2,\ldots$, x_k объема n. Наблюдавшиеся значения x_i ($i = \overline{1,k}$) признака X называют вариантами, а последовательность вариант, записанных в возрастающем порядке, – вариационным рядом.

Определение. Статистическим распределением выборки называют перечень вариант x_i вариационного ряда и соответствующих им частот n_i (сумма всех частот равна объему выборки n) или относительных частот ω_i (сумма всех относительных частот равна единице).

Определение. Полигоном частот называют ломаную, отрезки которой соединяют точки (x_1, n_1), (x_2, n_2), …, (x_k, n_k).

Определение. Полигоном относительных частот называют ломаную, отрезки которой, соединяют точки:

$$\left(x_1,\frac{n_1}{n}\right),\left(x_2,\frac{n_2}{n}\right),\ldots,\left(x_k,\frac{n_k}{n}\right).$$

При непрерывном распределении признака весь интервал, в котором заключены все наблюдаемые значения признака, разбивают на ряд частичных интервалов длины h и находят n_i сумму частот вариант, попавших в i – й интервал.

Определение. Гистограммой частот называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длины h, а высоты равны отношению n_i/h (плотность частоты).

Площадь частичного *i* – го прямоугольника равна n_i – сумме частот вариант, попавших в *i* – й интервал.

Площадь гистограммы частот равна сумме всех частот (объем выборки).

Определение. Гистограммой относительных частот называют ступенчатую фигуру, состоящую из прямоугольников, основаниями, которых служат частичные интервалы длины h, а высоты равны отношению w_i/h (плотность относительной частоты).

Площадь частичного i –го прямоугольника равна w_i – относительной частоте вариант, попавших i –й интервал.

Площадь гистограммы частот равна сумме всех относительных частот (единице).

Пример 5.1

Построить полигон частот по данному распределению выборки:

x_i	1	4	5	7
n_i	20	10	14	6

Решение.

Отложим на оси абсцисс варианты x_i, а на оси ординат, соответствующие им частоты n_i.

Отметим в полученной системе координат точки: (1;20), (4; 10), (5; 14), (7;6). Соединив полученные точки отрезками, получим искомый полигон

(Рис. 5.1).

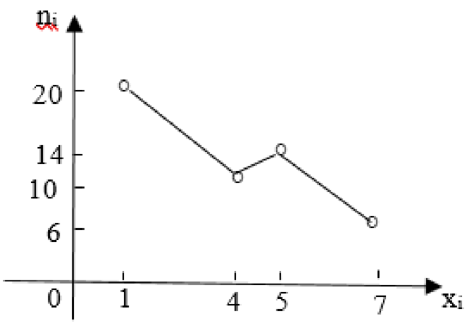


Рис. 5.1

Пример 5.2

Выборка дана в иле распределения частот:

x_i	2	4	5	7	10
n_i	15	20	10	10	45

а) найти распределение относительных частот;

б) построить полигон относительных частот.

Решение.

а) Найдем объем выборки: 15+20+10+10+45=100.

Найдем относительные частоты: w_1 = 15:100=0,15; w_2 = 20:100= 0,2; w_3 = 10:100= 0,1; w_4 =45:100=0,45.

Искомое распределение относительных частот имеет вид:

x_i	2	4	5	7	10
n_i	0,15	0,2	0,1	0,1	0,45

б) Построим полигон относительных частот:

Отложим на оси абсцисс варианты x_i, а на оси ординат, соответствующие им частоты w_i.

Отметим в полученной системе координат точки: (2; 0,15), (4; 0,2), (5; 0,1), (7; 0,1), (10; 0,45). Соединив полученные точки отрезками, получим искомый полигон (Рис. 5.2).

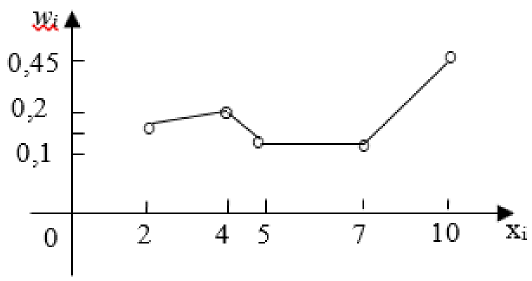


Рис. 5.2

Пример 5.3

Построить гистограмму частот по данному распределению выборки объемом n=100:

Номер интервала i	Частичный интервал x_i-x_{i-1}	Сумма частот вариант интервала n_i	Плотность частоты n_i/h
1	1–5	10	2,5
2	5–9	20	5
3	9–13	50	12,5
4	13–17	12	3
5	17–21	8	2

Решение.

Построим на оси абсцисс заданные интервалы длины *h=4*. На каждом интервале построим прямоугольник высотой равный плотности частоты n_i/h. Рис. 5.3.

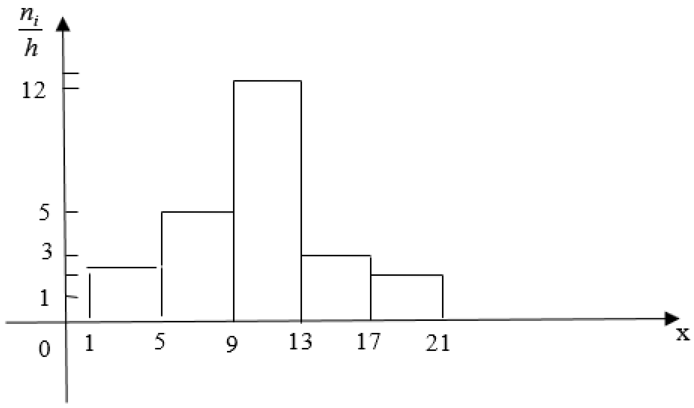


Рис. 5.3

Пример 5.4

Построить гистограмму относительных частот по данному распределению выборки:

Номер интервала i	Частичный интервал x_i-x_{i-1}	Сумма частот вариант интервала n_i
1	0 – 2	20
2	2 – 4	30
3	4 – 6	50

Решение.

Найдем относительные частоты: w_1 = 20:100=0,2; w_2 = 30:100= 0,3; w_3 = 50:100= 0,5.

Найдем плотность относительных частот: w_1 = 0,2:2=0,1; w_2 = 0,3:2= 0,15; w_3 = 0,5:2= 0,25.

Построим на оси абсцисс заданные интервалы длины $h=2$. На каждом интервале построим прямоугольник высотой равный плотности частоты w_i/h. Рис. 5.4.

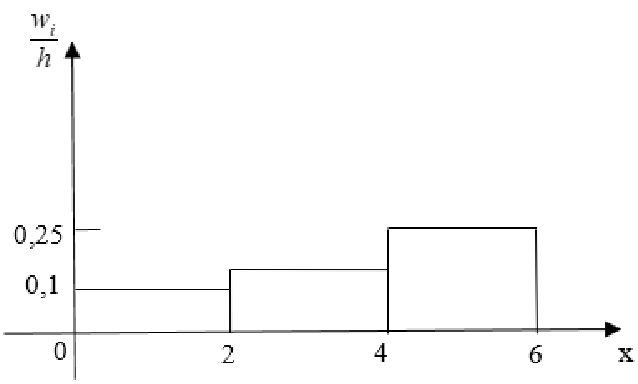


Рис. 5.4

5.2. Числовые характеристики статистического распределения выборки

Определение. Выборочное среднее $\bar{x}$ – среднее арифметическое значение признака выборочной совокупности

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{k} x_i \cdot n_i \qquad (5.1)$$

Определение. Выборочной дисперсией *D* называют среднее арифметическое квадратов отклонения наблюдаемых значений признака от их среднего значения $\bar{x}$:

$$D = \sigma^2 = \frac{1}{n} \sum_{i=1}^{k} (x_i - \bar{x})^2 \cdot n_i. \qquad (5.2)$$

Дисперсию можно рассчитать по формуле:

$$D = \overline{x^2} - (\bar{x})^2. \qquad (5.3)$$

Определение. Выборочным средним квадратическим отклонением называется величина $\sigma = \sqrt{D},$ характеризующая отклонение, разброс в линейных размерах данных выборки относительно выборочного среднего.

5.3 Точечные оценки

Определение. Статистической оценкой неизвестного параметра генеральной совокупности называется приближенное значение, полученное по данным выборки.

Определение. Точечная оценка – оценка, которая определяется одним числом θ. Это точка на числовой оси, около которой находится оцениваемый параметр генеральной совокупности θ_0.

Определение. Оценка θ параметра θ_0 называется несмещенной, если $M(\theta) = \theta_0$; в противном случае – смещенной.

Определение. Оценка θ параметра θ_0 называется состоятельной, если для любого положительного δ $\lim\limits_{n \to \infty} p(|\theta - \theta_0| < \delta) = 1$, то есть θ стремится к θ_0 по вероятности и означает неограниченное увеличение точности с ростом объема выборки.

Определение. Оценка θ параметра θ_0 называется эффективной, если она является несмещенной и имеет наименьшую дисперсию при заданном объеме выборки.

Теорема 1. Выборочное среднее – несмещенная, состоятельная и эффективная оценка математического ожидания признака генеральной совокупности.

Теорема 2. Дисперсия выборочного среднего в *n* раз меньше дисперсии генеральной совокупности:

$$D(\bar{x}) = \frac{\sigma_0^2}{n}. \qquad (5.4)$$

Теорема 3. Математическое ожидание выборочной дисперсии рассчитывается по формуле

$$M(\sigma^2) = \frac{n-1}{n}\sigma_0^2. \qquad (5.5)$$

Следовательно, дисперсия выборочного среднего является смещенной оценкой генеральной дисперсии. Чтобы получить несмещенную оценку вводится исправленная дисперсия:

$$s^2 = \frac{n}{n-1}\sigma^2. \qquad (5.6)$$

Пример 5.5

Выручка в магазине от продажи обуви составила соответственно по месяцам следующие значения (млн. руб.):

Месяц	1	2	3	4	5	6	7	8	9	10	11	12
P	0,2	0,5	0,4	0,2	0,4	0,5	0,2	0,2	0,4	0,5	0,4	0,2

Найти выборочную среднюю, выборочную дисперсию, исправленную дисперсию.

Решение.

По условию задачи в первом месяце магазин выручил 2 млн. рублей из 10 млн. руб., во втором месяце – 5 млн. руб. в третьем – 4 млн. руб. и т. д., а каждый месяц выручка должна составлять 10 млн. руб., поэтому выборочную среднюю найдем по формуле:

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

$$\bar{x}_в = \frac{0{,}2+0{,}5+0{,}4+0{,}2+0{,}4+0{,}5+0{,}2+0{,}2+0{,}4+0{,}5+0{,}4+0{,}2}{12} = \frac{4{,}1}{12} = 0{,}34$$

Выборочную дисперсию найдем по формуле: $D = \overline{x^2} - (\bar{x})^2$.

$$D_в = \frac{0{,}04+0{,}25+0{,}16+0{,}04+0{,}16+0{,}25+0{,}04+0{,}04+0{,}16+0{,}25+0{,}16+0{,}04}{12} - (0{,}34)^2 =$$

$$= 0{,}1325 - 0{,}1156 = 0{,}0169$$

Находим несмещенную оценку дисперсии по формуле: $s^2 = \frac{n}{n-1} \cdot D_в$

$$s^2 = \frac{12}{12-1} \cdot 0{,}0169 = 0{,}018.$$

Ответ. $\overline{x}_в = 0{,}34$; $D_в = 0{,}0169$; . $s^2 = 0{,}018$

5.4. Метод произведений вычисления выборочных средней и дисперсии

А. Метод произведения

– равноотстоящие варианты.

Рассмотрим выборку в виде распределения с равноотстоящими вариантами и соответствующими им частотами,

Выборочная средняя находится по формуле: $\overline{x}_в = M_1^* h + C$ (5.7)

Выборочная дисперсия по формуле: $D_в = [M_2^* - (M_1^*)^2] \cdot h^2$ (5.8)

где h – шаг (разность между двумя вариантами);

C – ложный нуль (варианта, которая расположена в середине вариационного ряда);

$u_i = \frac{x_i - C}{h}$ – условная варианта;

$M_1^* = \frac{\sum n_i u_i}{n}$ – условный момент первого порядка;

$M_2^* = \frac{(\sum n_i u_i)^2}{n}$ – условный момент второго порядка.

Пример 5.6

Найти методом произведения выборочную среднюю и выборочную дисперсию по заданному распределению выборки объемом n =100 (Таблица 5.1).

Таблица 5.1

Варианта x_i	12	14	16	18	20	22
Частота n	5	15	50	16	10	4

Решение.

Найдем условные варианты, приняв в качестве ложного нуля, варианту 16 (C=16).

$$u_1 = \frac{12-16}{2} = -2; \quad u_2 = \frac{14-16}{2} = -1; \; u_3 = \frac{16-16}{2} = 0;$$

$$u_4 = \frac{18-16}{2} = 1; \quad u_5 = \frac{20-16}{2} = 2; \quad u_6 = \frac{22-16}{2} = 3$$

Составим расчетную таблицу 5.2:

Таблица 5.2

x_i	n_i	u_i	$n_i u_i$	$n_i u_i^2$
12	5	-2	-10	20
14	15	-1	-15	15
16	50	0	0	0
18	16	1	16	16
20	10	2	20	40
22	4	3	12	36
$\sum$	100		23	127

Для нахождения средней выборочной используем формулу: $\overline{x}_в = M_1^* h + C$

$$M_1^* = \frac{\sum n_i u_i}{n} = \frac{23}{100} = 0{,}23$$

$$\overline{x}_в = 0{,}23 \cdot 2 + 16 = 16{,}46$$

Для нахождения выборочной дисперсии используем формулу: $D_в = [M_2^* - (M_1^*)^2] \cdot h^2$

$$M_2^* = \frac{\sum n_i u^2{}_i}{n} = \frac{127}{100} = 1{,}27$$

$$D_в = [1{,}27 - (0{,}23)^2] \cdot 2^2 = [1{,}27 - 0{,}0529] \cdot 4 = 4{,}87$$

Ответ. $\overline{x}_в = 16{,}46$; $D_в = 4{,}87$.

– не равноотстоящие варианты.

Если первоначальные варианты не являются равноотстоящими, то интервал, в котором заключены все варианты выборки, делят на несколько равных частичных интервалов длины h.

Затем находят середины частичных интервалов, которые и образуют последовательность равноотстоящих вариант. В качестве частоты каждой середины интервала принимают сумму частот вариант, которые попали в соответствующий частичный интервал.

При вычислении выборочной дисперсии для уменьшения ошибки, вызванной группировкой, делают поправку Шеппарда, а именно, вычитают из полученной дисперсии 1/12 квадрата длины частичного интервала.

$$D^* = D_в - (1/12) \cdot h^2 \qquad (5.9)$$

Пример 5.7

Найти методом произведения выборочную среднюю и выборочную дисперсию по заданному распределению выборки объемом n =100 (Таблица 5.3)

Таблица 5.3

Варианта x_i	2	3	7	9	11	12,5	16	18	23	25	26
Частота n	3	5	10	6	10	4	12	13	8	20	9

Решение.

Разобьем данное распределение на частичные интервалы. Для начала найдем размах: $x_{max} - x_{min} = 26 - 2 = 24$.

Длина частичного интервала сделаем равным 6 (h=6). Таким образом, получим следующие интервалы, представленные в таблице 5.4:

Таблица 5.4

Интервалы	2–8	8–14	14–20	20–26
Частота n	3+5+10=18	6+10+4=20	12+13= 25	8+20+9=37

Из интервального распределения делаем распределение с равноотстоящими вариантами.

$$x_1^* = \frac{2+8}{2} = 5; \quad x_2^* = \frac{14+8}{2} = 11; \quad x_3^* = \frac{20+14}{2} = 17; \quad x_4^* = \frac{26+20}{2} = 23$$

Таким образом, распределение с равноотстоящими вариантами имеет вид (таблица 5.5):

Таблица 5.5

Варианта x_i^*	5	11	17	23
Частота n	18	20	25	37

Найдем условные варианты, приняв в качестве ложного нуля, варианту 17 (С=17).

$$u_1 = \frac{5-17}{6} = -2 \quad u_2 = \frac{11-17}{6} = -1; \quad u_3 = \frac{17-17}{6} = 0; \quad u_4 = \frac{23-17}{6} = 1.$$

Составим расчетную таблицу 5.6:

Таблица 5.6

x_i^*	n_i	u_i	$n_i u_i$	$n_i u_i^2$
5	18	-2	-36	72
11	20	-1	-20	20
17	25	0	0	0
23	37	1	37	37
$\sum$	100		-19	129

Для нахождения средней выборочной используем формулу:
$\bar{x}_в = M_1^* h + C$.

$$M_1^* = \frac{\sum n_i u_i}{n} = \frac{-19}{100} = -0{,}19$$

$$\bar{x}_в = -0{,}19 \cdot 6 + 17 = -1{,}14 + 17 = 15{,}86$$

Для нахождения выборочной дисперсии используем формулу: $D_в = [M_2^* - (M_1^*)^2] \cdot h^2$

$$M_2^* = \frac{\sum n_i u^2{}_i}{n} = \frac{129}{100} = 1{,}29$$

$$D_в^* = [1{,}29 - (-0{,}19)^2] \cdot 6^2 = [1{,}29 - 0{,}0361] \cdot 36 = 45{,}14$$

$$D_в = 45{,}14 - \frac{6^2}{12} = 45{,}14 - 3 = 42{,}14$$

Ответ. $\bar{x}_в = 15{,}86$; $D_в = 42{,}14$

5.5. Метод сумм вычисления выборочных средней и дисперсии

Рассмотрим выборку в виде распределения с равноотстоящими вариантами и соответствующими им частотами,

Выборочная средняя находится по формуле: $\bar{x}_в = M_1^* h + C$

Выборочная дисперсия по формуле: $D_в = [M_2^* - (M_1^*)^2] \cdot h^2$

где h – шаг (разность между двумя вариантами);

С – ложный нуль (варианта, которая расположена в середине вариационного ряда);

$u_i = \dfrac{x_i - C}{h}$ – условная варианта;

$M_1^* = \dfrac{d_1}{n}$ – условный момент первого порядка;

$M_2^* = \dfrac{s_1 + 2s_2}{n}$ – условный момент второго порядка.

где $d_1 = a_1 - b_1$, $s_1 = a_1 + b_1$, $s_2 = a_2 + b_2$.

Пример 5.8

Найти методом сумм выборочную среднюю и выборочную дисперсию по заданному распределению выборки объемом n =100 (Таблица 5.7).

Таблица 5.7

Варианта x_i	48	52	56	60	64	68	72	76	80	84
Частота n	2	4	6	8	12	30	18	8	7	5

Решение.

В качестве ложного нуля, варианту 68 (С=68). Для вычисления a_1, b_1, a_2, b_2. составим расчетную таблицу 5.8:

В первом столбце запишем варианты, во втором – соответствующие частоты, в третьем столбце выбираем ложный нуль – это варианта с наибольшей частотой (68). Сверху вниз до ложного нуля и снизу вверх для каждой варианты находим накопительные частоты: для 48 частота 2 переписывается; для варианты 52 находим накопительную частоту, для этого 2 + 4 = 6, для варианты 56 накопительная частота равна 2+4+6 = 12, для варианты 60 накопительная частота 2+4+6+8 =20, для варианты 64 накопительная частота равна 2+4+6+8+12 = 32.

Аналогично находим накопительные частоты и снизу вверх. В четвертом столбце сверху и снизу ложного нуля запишем также нули. И повторим алгоритм нахождения накопительных частот в четвертом столбце.

Для определения a_1 находим сумму накопительных частот в третьем столбе снизу ложного нуля: 38+20+12+5 = 75; для определения b_1 находим сумму накопительных частот в третьем столбе сверху ложного нуля: 2+6+12+20+32 =72.

Для определения a_2 находим сумму накопительных частот в четвертом столбе снизу ложного нуля: 5+17+37= 59; для определения b_2 находим сумму накопительных частот в четвертом столбе сверху ложного нуля: 2+8+20+40= 70.

Таблица 5.8

x_i	n_i	$b_1 = 72$	$b_2 = 70$
48	2	2	2
52	4	6	8
56	6	12	20
60	8	20	40
64	12	32	0
68	30	0	0
72	18	38	0
76	8	20	37
80	7	12	17
84	5	5	5
$\sum$	$n=100$	$a_1 = 75$	$a_2 = 59$

Для нахождения средней выборочной используем формулу: $\overline{x}_в = M_1^* h + C$.

$$M_1^* = \frac{d_1}{n} = \frac{a_1 - b_1}{100} = \frac{75-72}{100} = 0{,}03$$

$$\overline{x}_в = 0{,}03 \cdot 4 + 68 = 0{,}12 + 68 = 68{,}12$$

Для нахождения выборочной дисперсии используем формулу: $D_в = [M_2^* - (M_1^*)^2] \cdot h^2$

$$M_2^* = \frac{s_1 + 2s_2}{n} = \frac{(a_1+b_1) + 2(a_2+b_2)}{100} = \frac{(75+72)+2(59+70)}{100} = \frac{405}{100} = 4{,}05$$

$$D_в{}^* = [4{,}05 - (-0{,}03)^2] \cdot 4^2 = [4{,}05 - 0{,}0009] \cdot 16 = 64{,}79$$

Ответ. $\overline{x}_в = 68{,}12$; $D_в = 64{,}79$

5.6. Асимметрия и эксцесс эмпирического распределения

Асимметрия эмпирического распределения определяется равенством:

$$A_s = \frac{m_3}{\sigma^3} \qquad (5.10)$$

где σ – средне квадратическое отклонение; m_3 – центральный эмпирический момент третьего порядка, который удобно вычислять по формуле:

$$m_3 = [M_3^* - 3M_1^* M_2^* + 2(M_1^*)^3] \cdot h^3 \qquad (5.11)$$

Эксцесс эмпирического распределения определяется равенством:

$$e_к = \frac{m_4}{\sigma^4} - 3 \qquad (5.12)$$

где σ – среднее квадратическое отклонение;

m4 – центральный эмпирический момент четвертого порядка, который удобно вычислять по формуле:

$$m_4 = [M_4^* - 4M_1^* M_3^* + 6(M_1^*)^2 \cdot M_2^* - 3(M_1^*)^4] \cdot h^4 \qquad (5.13)$$

где $M_k^* = \frac{\sum n_i u_i^k}{n}$ – условный момент k – го порядка;

$u_i = \frac{x_i - C}{h}$ – условная варианта.

Вычисление асимметрии и эксцесса через условные моменты удобно осуществлять методом произведения или методом сумм.

– метод произведений

Пример 5.9

Найти методом произведения асимметрию и эксцесс по заданному распределению выборки объемом n =100 (Таблица 5.9)

Таблица 5.9

Варианта x_i	12	14	16	18	20	22
Частота n	5	15	50	16	10	4

Решение.

Составим расчетную таблицу (таблица 5.10) для вычисления асимметрии и эксцесса.

Таблица 5.10

x_i	n_i	u_i	$n_i u_i$	$n_i u_i^2$	$n_i u_i^3$	$n_i u_i^4$
12	5	-2	-10	20	-40	80
14	15	-1	-15	15	-15	15
16	50	0	0	0	0	0
18	16	1	16	16	16	256
20	10	2	20	40	160	810
22	4	3	12	36	324	1024
$\sum$	100		23	127	149	595

Найдем условные моменты третьего и четвертого порядка:

$$M_3^* = \frac{\sum n_i u_i^3}{n} = \frac{149}{100} = 1{,}49$$

$$M_4^* = \frac{\sum n_i u_i^4}{n} = \frac{595}{100} = 5{,}95.$$

Так как моменты первого и второго порядка соответственно равны:

$$M_1^* = \frac{\sum n_i u_i}{n} = \frac{23}{100} = 0{,}23$$

$$M_2^* = \frac{\sum n_i u^2{}_i}{n} = \frac{127}{100} = 1{,}27,$$

то центральные эмпирические моменты третьего и четвертого порядка будут равны:

$$m_3 = [M_3^* - 3M_1^* M_2^* + 2(M_1^*)^3] \cdot h^3 =$$
$$= [1{,}49 - 3 \cdot 0{,}23 \cdot 1{,}27 + 2 \cdot (0{,}23)^3] \cdot 2^3 =$$
$$= [1{,}49 - 0{,}8763 + 0{,}024334] \cdot 8 = 0{,}638034 \cdot 8 = 5{,}104$$

$$m_4 = [M_4^* - 4M_1^* M_3^* + 6(M_1^*)^2 \cdot M_2^* - 3(M_1^*)^4] \cdot h^4 =$$
$$= [5{,}95 - 4 \cdot 0{,}23 \cdot 1{,}49 + 6 \cdot (0{,}23)^2 \cdot (1{,}27) - 3(0{,}23)^4] \cdot 2^4 =$$
$$= [5{,}95 - 1{,}3708 + 0{,}403098 - 0{,}00839523] \cdot 16 = 4{,}97390277 \cdot 16 =$$
$$= 79{,}582$$

В примере 5.6 была найдена дисперсия $D_в$ = 4,87 из которой найдем среднее квадратическое отклонение: $\sigma = \sqrt{D_в} = \sqrt{4{,}87} = 2{,}21$.

Отсюда имеем:

$$A_s = \frac{5{,}104}{(2{,}21)^3} = \frac{5{,}104}{10{,}793861} \approx 0{,}47 \quad e_k = \frac{79{,}582}{(2{,}21)^4} - 3 = \frac{79{,}582}{23{,}85443281} - 3 \approx 0{,}34$$

Ответ. A_s = 0,47; e_k = 0,34.

– метод сумм

Пример 5.10

Найти методом сумм асимметрию и эксцесс по заданному распределению выборки объемом n =100 (Таблица 5.11).

Таблица 5.11

Варианта x_i	48	52	56	60	64	68	72	76	80	84
Частота n	2	4	6	8	12	30	18	8	7	5

Решение.

Найдем условные моменты первого, второго третьего, четвертого порядка по формулам:

$$M_1^* = \frac{d_1}{n}; \qquad M_2^* = \frac{s_1 + 2s_2}{n};$$

(5.14)

$$M_3^* = \frac{d_1 + 6d_2 + 6d_3}{n}; \qquad M_2^* = \frac{s_1 + 4s_2 + 36s_3 + 24s_4}{n}.$$

Для нахождения d_1, d_2, d_3, s_1, s_2, s_3, s_4 составим расчетную таблицу 5.12.

В качестве ложного нуля, варианту 68 (С=68).

В первом столбце запишем варианты, во втором – соответствующие частоты, в третьем столбце выбираем ложный нуль – это варианта с наибольшей частотой (68). Сверху вниз до ложного нуля и снизу вверх для каждой варианты находим накопительные частоты: для 48 частота 2 переписывается; для варианты 52 находим накопительную частоту, для этого

2+4 = 6, для варианты 56 накопительная частота равна 2+4+6 = 12, для варианты 60 накопительная частота 2+4+6+8 =20, для варианты 64 накопительная частота равна 2+4+6+8+12 = 32.

Аналогично находим накопительные частоты и снизу вверх.

В четвертом, пятом, шестом столбцах сверху и снизу ложного нуля запишем также нули. И повторим алгоритм нахождения накопительных частот в каждом столбце.

Для определения a_1 находим сумму накопительных частот в третьем столбе снизу ложного нуля: 38+20+12+5 = 75; для определения b_1 находим сумму накопительных частот в третьем столбе сверху ложного нуля: 2+6+12+20+32 = 72.

Для определения a_2 находим сумму накопительных частот в четвертом столбе снизу ложного нуля: 5+ 17+37 = 59; для определения b_2 находим сумму накопительных частот в четвертом столбе сверху 2+8+20+40 = 70.

Для определения a_3 находим сумму накопительных частот в пятом столбе снизу ложного нуля: 5+22 = 27; для определения b_3 находим сумму накопительных частот в пятом столбе сверху 2+10+30 = 42.

Для определения a_4 находим сумму накопительных частот в четвертом столбе снизу ложного нуля: 5; для определения b_4 находим сумму накопительных частот в четвертом столбе сверху 2+12 =14.

Таблица 5.12

x_i	n_i	$b_1 = 72$	$b_2 = 70$	$b_3 = 42$	$b_4 = 14$
48	2	2	2	2	2
52	4	6	8	10	12
56	6	12	20	30	0
60	8	20	40	0	0
64	12	32	0	0	0
68	30	0	0	0	0
72	18	38	0	0	0
76	8	20	37	0	0
80	7	12	17	22	0
84	5	5	5	5	5
$\sum$	$n = 100$	$a_1 = 75$	$a_2 = 59$	$a_3 = 27$	$a_4 = 5$

Откуда найдем:

$d_1 = a_1 - b_1 = 75 - 72 = 3$; $d_2 = a_2 - b_2 = 59 - 70 = -11$;

$d_3 = a_3 - b_3 = 27 - 42 = -15$

$s_1 = a_1 + b_1 = 75+72 = 147$; $s_2 = a_2 + b_3 = 59+70 = 129$;

$s_3 = a_3 + b_3 = 27+42 = 69$; $s_4 = 5+14 = 19$.

$$M_1^* = \frac{d_1}{n} = \frac{3}{100} = 0{,}03; \qquad M_2^* = \frac{s_1 + 2s_2}{n} = \frac{147 + 2 \cdot 129}{100} = 4{,}05;$$

$$M_3^* = \frac{d_1 + 6d_2 + 6d_3}{n} = \frac{3 + 6 \cdot (-11) + 6 \cdot (-15)}{100} = -1{,}53;$$

$$M_2^* = \frac{s_1 + 4s_2 + 36s_3 + 24s_4}{n} = \frac{147 + 4 \cdot 129 + 36 \cdot 69 + 24 \cdot 19}{100} = 48{,}93$$

Таким образом, центральные эмпирические моменты третьего и четвертого порядка соответственно равны:

$$m_3 = [M_3^* - 3M_1^* M_2^* + 2(M_1^*)^3] \cdot h^3 =$$
$$= [-1{,}53 - 3 \cdot 0{,}03 \cdot 4{,}05 + 2 \cdot (0{,}03)^3] \cdot 4^3 =$$
$$= [-1{,}53 - 0{,}3645 + 0{,}000054] \cdot 64 = -1{,}894446 \cdot 64 = -121{,}245$$

$$m_4 = [M_4^* - 4M_1^* M_3^* + 6(M_1^*)^2 \cdot M_2^* - 3(M_1^*)^4] \cdot h^4 =$$
$$= [48{,}93 - 4 \cdot 0{,}03 \cdot (-1{,}53) + 6 \cdot (0{,}03)^2 \cdot (4{,}05) - 3(0{,}03)^4] \cdot 2^4 =$$
$$= [48{,}93 + 0{,}1836 + 0{,}02187 - 0{,}00000243] \cdot 16 = 12578{,}679$$

В примере 5.8 была найдена дисперсия $D_в$ = 64,78 из которой найдем среднее квадратическое отклонение: $\sigma = \sqrt{D_в} = \sqrt{64{,}78} = 8{,}05$.

Отсюда имеем:

$$A_s = \frac{-121{,}245}{(8{,}05)^3} = \frac{-121{,}245}{521{,}660125} \approx -0{,}25$$

$$e_k = \frac{12578{,}679}{(8{,}05)^4} - 3 = \frac{12578{,}679}{4199{,}36400626} - 3 \approx 26{,}97$$

Ответ. $A_s = -0{,}25$; $e_k = 26{,}97$.

5.7. Интервальные оценки

Определение. Интервал ($\theta - \delta$, $\theta + \delta$), в пределах которого с вероятностью γ находится оцениваемый параметр генеральной совокупности θ_0, называется доверительным интервалом. Значение γ называется доверительной вероятностью или надежностью оценки; предельная погрешность δ – точность оценки.

Доверительный интервал для оценки математического ожидания нормального распределения определяется следующим образом:

$$\bar{x} - \delta < a < \bar{x} + \delta, \qquad (5.15)$$

причем, если стандартное отклонение этого распределения известно, то

$$\delta = \frac{t \cdot \sigma_0}{\sqrt{n}}; \qquad (5.16)$$

если стандартное отклонение неизвестно, то

$$\delta = \frac{t_\gamma \cdot s}{\sqrt{n}}. \qquad (5.17)$$

Здесь число t определяется из равенства $\Phi(t) = \gamma/2$; t_γ находится по таблице коэффициентов Стьюдента при заданных n и γ; s – исправленное выборочное среднее квадратическое отклонение.

Пример 5.11

Заданы среднее квадратическое отклонение σ=8 нормально распределенной случайной величины X, выборочная средняя $\bar{x}$=18,91 и объем выборки n=49.

Найти доверительные интервалы для оценки неизвестного математического ожидания *a* с заданной надежностью γ=0,95.

Решение.

По надежности γ=0,95 из соотношения Φ(z) = γ/2 находим значение функции Лапласа (Приложение 2): Φ(z) =0,475.

По таблице значений функции Лапласа (Приложение 2) находим z=1,96.

Используя неравенство для интервальной оценки математического ожидания:

$$\bar{x}_в - \frac{z \cdot \sigma_x}{\sqrt{n}} < a < \bar{x}_в + \frac{z \cdot \sigma_x}{\sqrt{n}} \tag{5.18}$$

$$18{,}91 - \frac{1{,}96 \cdot 8}{\sqrt{49}} < a < 18{,}91 + \frac{1{,}96 \cdot 8}{\sqrt{49}}$$

Получаем: $$18{,}91 - \frac{15{,}68}{7} < a < 18{,}91 + \frac{15{,}68}{7}$$

$$18{,}91 - 2{,}24 < a < 18{,}91 + 2{,}24$$

$$16{,}67 < a < 21{,}15$$

Ответ. 16,67 < a < 21,15

Определение. Интервальной оценкой (с надежностью γ) стандартного отклонения σ_x нормально распределенного количественного признака *X* по исправленному выборочному стандартному отклонению *s* служит доверительный интервал:

$$\frac{s}{1+q} < \sigma_x < \frac{s}{1-q} \tag{5.19}$$

Пример 5.12

На нескольких предприятиях проведена проверка качества 100 изделий, после чего осуществлена обработка полученных данных. В результате получено несмещенное значение выборочного среднего квадратического отклонения s = 4. Считая распределения качественных изделий нормальным, найти с надежностью 0,95 доверительный интервал для оценки среднего квадратического отклонения.

Решение.

Для нахождения доверительного интервала для оценки среднего квадратического отклонения используем формулу: $\frac{s}{1+q} < \sigma_x < \frac{s}{1-q}$

На основании данных значений γ =0,95, n=100 по таблице найдем значение q (Приложение 4): q (100; 0,95) =0,143, q <1

Подставляем в неравенство:

$$\frac{4}{1+0{,}143} < \sigma_x < \frac{4}{1-0{,}143}$$

3,5 < σ_x <4,67

Ответ. 3,5 < σ_x <4,67

Примеры для самостоятельного решения

1. По заданному распределению выборки найти эмпирическую функцию распределения.

x_i	2	5	7	8
n_i	1	3	2	4

2. В результате эксперимента получены данные, записанные в виде статистического ряда.

а) Постройте интервальный ряд, содержащий 8 интервалов;

б) Построить гистограмму абсолютных и относительных и частот;

в) Найти числовые характеристики выборки: $\overline{x}$, D_B, σ_B и несмещённые оценки: $\overline{D}_B$, $\overline{\sigma}_B$

г) Найти доверительный интервал для математического ожидания а и среднего квадратического отклонения σ при надёжности $\gamma = 0{,}99$

193	203	224	202	216	195	202	192	207	212
207	192	213	216	193	193	207	213	212	213
202	203	216	193	202	212	192	193	195	192

3. Найти методом произведения и сумм выборочную среднюю, выборочную дисперсию, асимметрию и эксцесс по заданному распределению с равноотстоящими вариантами и не равноотстоящими вариантами. (Распределения подбираются самостоятельно)

4. На овцеводческой ферме из стада произведена выборка для взвешивания 36 овец. Их средний вес оказался равным 50 кг. Предположив распределение веса нормальным и определив несмещенную оценку выборочной дисперсии s^2 = 16, найти доверительный интервал для оценки математического ожидания с надежностью а) 0,8; б) 0,9; в) 0,95.

5. По данным выборки объема n = 20 найдено несмещенное значение выборочного среднего квадратического отклонения s = 2 нормально распределенной случайной величины X. Найти с надежностью 0,95 доверительный интервал для оценки среднего квадратического отклонения.

Тема 6. Элементы корреляционного анализа

6.1. Коэффициент корреляции

Рассмотрим совокупность двух случайных величин X и Y, которые назовем признаками

Для описания связей между двумя случайными величинами применяется математическое понятие функция *f* , в которой каждому значению переменной X (независимой переменной) ставится в соответствие единственного значения Y: *f(X)= Y,* такого рода однозначные зависимости между X и Y называется *функциональным.*

Если в совокупности значений определенному значению, рассматриваемого в качестве аргумента, ставится в соответствие несколько значений из другой совокупности, то такая зависимость называется *статистической,* а зависимость между переменными величинами называется *корреляционной.*

Функциональные связи можно обнаружить и измерить лишь на единичных или групповых объектах, однако этого нельзя сделать с корреляционными связями, которые можно изучить только на групповых объектах, методами математической статистики.

Корреляционная связь между признаками бывает линейной и нелинейной. Положительной и отрицательной.

Задачей корреляционного анализа является установления направления и формы связи между варьирующими признаками, измерением ее тесноты, проверки достоверности выборочных показателей корреляции.

Сопряженность между переменными величинами X и Y можно установить, сопоставляя числовые значения одного из них с соответствующими значениями другого.

Если при увеличении одной величины увеличивается другая, то это указывает на *положительную* связь между ними.

Если при увеличении одной величины другая уменьшается, то говорят об отрицательной связи.

Подобную связь устанавливают при наличии однозначных отношений между X и Y, когда идет речь о приращении или уменьшении значении функции по заданным значениям аргумента.

Теснота связи определяется с помощью коэффициента корреляции по формуле:

$$r_в = \frac{\sum xy - n\cdot\bar{x}\cdot\bar{y}}{n\cdot\sigma_x\cdot\sigma_y} \qquad (6.1)$$

Коэффициент корреляции – это число, которое лежит в пределах [-1;1]

Если связь между признаками отсутствует, то $r_{XY}= 0$.

Чем сильнее сопряженность между признаками X и Y, тем выше значение коэффициента корреляции. Таким образом, при $|r| >0$ этот показатель характеризует не только наличие связи между признаками, но и степень сопряженности между ними.

При положительной (прямой) связи, когда большему значению одного признака соответствует большее значение другого признака, коэффициент корреляции имеет положительный знак и находится в пределах (0; 1].

При отрицательной (обратной) связи, когда большему значению одного признака соответствует меньшее значение другого признака, коэффициент корреляции имеет положительный знак и находится в пределах [–1;0).

Коэффициент корреляции нашел широкое применение в практике, но не является универсальным показателем корреляционных связей, так как способен характеризовать только линейные связи.

Пример 6.1

На основании полученных измерений найти выборочный коэффициент корреляции.

X	4	6	8	10	12
Y	5	8	7	9	14

Решение.

Коэффициент корреляции найдем по формуле: $r_в = \dfrac{\sum xy - n\cdot\overline{x}\cdot\overline{y}}{n\cdot\sigma_x\cdot\sigma_y}$, для этого найдем: $\overline{x}$, $\overline{y}$, $\sum xy$.

$$\overline{x} = \frac{4+6+8+10+12}{5} = 8 \qquad \overline{y} = \frac{5+8+7+9+14}{5} = 8{,}6$$

$$\sigma_x = \sqrt{\frac{(4-8)^2+(6-8)^2+(8-8)^2+(10-8)^2+(12-8)^2}{5} - 8^2} = \sqrt{72-64} = \sqrt{8} = 2{,}83$$

$$\sigma_y = \sqrt{\frac{(5-8{,}6)^2+(8-8{,}6)^2+(7-8{,}6)^2+(9-8{,}6)^2+(14-8{,}6)^2}{5} - (8{,}6)^2} = \sqrt{83-73{,}96} = \sqrt{9{,}04} = 3{,}01$$

$$\sum xy = 4\cdot5+6\cdot8+8\cdot7+10\cdot9+12\cdot14 = 20+48+56+90+168 = 382$$

Следовательно, коэффициент корреляции равен:

$$r_в = \frac{382 - 5\cdot8\cdot8{,}6}{5\cdot2{,}83\cdot3{,}01} = \frac{382-344}{42{,}5913} \approx 0{,}89$$

Вывод: связь между признаками прямая и сильная.

Ответ. 0,89

Пример 6.2

На основании полученных измерений найти выборочный коэффициент корреляции.

X	8	2,4	7,2	8	1,6
Y	2,6	2,4	1,6	5,4	4

Решение.

Коэффициент корреляции найдем по формуле: $r_в = \dfrac{\sum xy - n \cdot \bar{x} \cdot \bar{y}}{n \cdot \sigma_x \cdot \sigma_y}$, для этого найдем: $\bar{x}$, $\bar{y}$, $\sum xy$.

$$\bar{x} = \frac{8 + 2,4 + 7,2 + 8 + 1,6}{5} = 5,44 \qquad \bar{y} = \frac{2,6 + 2,4 + 1,6 + 5,4 + 4}{5} = 3,2$$

$$\sigma_x = \sqrt{\frac{(8 - 5,44)^2 + (2,4 - 5,44)^2 + (7,2 - 5,44)^2 + (8 - 5,44)^2 + (1,6 - 5,44)^2}{5} - (5,44)^2} =$$

$$= \sqrt{37,632 - 29,594} = \sqrt{8,04} = 2,835$$

$$\sigma_y = \sqrt{\frac{(2,6 - 3,2)^2 + (2,4 - 3,2)^2 + (1,6 - 3,2)^2 + (5,4 - 3,2)^2 + (4 - 3,2)^2}{5} - (3,2)^2} =$$

$$= \sqrt{12,05 - 10,24} = \sqrt{1,81} = 1,345$$

$$\sum xy = 8 \cdot 2,6 + 2,4 \cdot 2,4 + 7,2 \cdot 1,6 + 8 \cdot 5,4 + 1,6 \cdot 4 = 20,8 + 5,76 + 11,52 + 43,2 + 6,4 =$$

$$= 87,68$$

Следовательно, коэффициент корреляции равен:

$$r_в = \frac{87,68 - 5 \cdot 5,44 \cdot 3,2}{5 \cdot 2,835 \cdot 1,345} = \frac{87,68 - 87,04}{19,065} \approx 0,0336$$

Ответ: 0,0336

Вывод: связь между признаками прямая и слабая.

Пример 6.3

На основании полученных измерений найти выборочный коэффициент корреляции.

X	3	5	7	9	10
Y	14	10	9	9	6

Решение.

Коэффициент корреляции найдем по формуле: $r_в = \dfrac{\sum xy - n \cdot \bar{x} \cdot \bar{y}}{n \cdot \sigma_x \cdot \sigma_y}$, для этого найдем: $\bar{x}$, $\bar{y}$, $\sum xy$.

$$\overline{x} = \frac{3+5+7+9+10}{5} = 6{,}8 \qquad \overline{y} = \frac{14+10+9+9+6}{5} = 9{,}6$$

$$\sigma_x = \sqrt{\frac{(3-6{,}8)^2+(5-6{,}8)^2+(7-6{,}8)^2+(9-6{,}8)^2+(10-6{,}8)^2}{5} - (6{,}8)^2} =$$

$$= \sqrt{52{,}8-46{,}24} = \sqrt{6{,}56} = 2{,}56$$

$$\sigma_y = \sqrt{\frac{(14-9{,}6)^2+(10-9{,}6)^2+(9-9{,}6)^2+(9-9{,}6)^2+(6-9{,}6)^2}{5} - (9{,}6)^2} =$$

$$= \sqrt{98{,}8-92{,}16} = \sqrt{6{,}64} = 2{,}58$$

$$\sum xy = 3\cdot 14 + 5\cdot 10 + 9\cdot 7 + 9\cdot 9 + 10\cdot 6 = 42+50+63+81+60 = 296$$

Следовательно, коэффициент корреляции равен:

$$r_в = \frac{296-5\cdot 6{,}8\cdot 9{,}6}{5\cdot 2{,}56\cdot 2{,}58} = \frac{296-326{,}4}{33{,}024} \approx -0{,}93$$

Ответ: – 0,93

Вывод: связь между признаками обратная и сильная.

6.2. Регрессионный анализ

Зависимость между признаками можно представить аналитически (с помощью формул и уравнений) и графически. График корреляционной зависимости строят по уравнению функции: $\overline{y}_x = f(x)$ или $\overline{x}_y = f(y)$, которое получило название уравнения *регрессии*.

Изменение функции в зависимости от одной или нескольких аргументов называется *регрессией.*

Напомним, что отличие статистической зависимости от функциональной заключается в следующем:

– в функциональной связи – между аргументом и функцией существует однозначное соответствие (каждому определенному значению аргумента х соответствует определенное значение у),

– в статистической связи – разным значениям одной переменной соответствуют различные распределения другой переменной, в которой можно указать частные средние $\overline{y}_x$ или $\overline{x}_y$. Поэтому форма статистической связи может быть описана не как зависимость отдельных значений у от х, а как зависимость частных средний $\overline{y}_x$ от значений х ($\overline{x}_y$ от значений у).

Для выражения регрессии служат корреляционные уравнения или уравнения регрессии, эмпирические и теоретические вычисленные ряды регрессии, их графики – линии регрессии и коэффициенты линейной и нелинейной регрессии.

Показатели регрессии выражают корреляционную связь двусторонне, учитывая изменение усредненных значений $\overline{y}_x$ признака Y при изменении значений x_i признака X и наоборот, изменение усредненных значений $\overline{x}_y$ признака X при изменении значений y_i признака Y.

Различных форм и видов корреляционной связи много. Задача сводится к тому, чтобы в каждом конкретном случае выявить форму связи и выразить ее соответствующим корреляционным уравнением, что позволяет предвидеть возможные изменения одного признака Y на основании изменений другого признака X, связанного с первым корреляционно.

Линейная зависимость между переменными X и Y описывается уравнением общего вида:

$$\overline{y}_x = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + \ldots, \tag{6.2}$$

где a_i – параметры уравнения, определяющие соотношения между аргументами x_i и функцией $\overline{y}_x$.

В практике учитываются не все возможные аргументы, а лишь некоторые, в простейшем случае всего один:

$$\overline{y}_x = a + bx \tag{6.3}$$

где a – свободный коэффициент

b – коэффициент регрессии

Поскольку показатели регрессии выражают корреляционную связь двусторонне, уравнение регрессии (3) записывается в виде:

$$\overline{y}_x - \overline{y} = r_в \cdot \frac{\sigma_y}{\sigma_x} \cdot (x - \overline{x}) \qquad \overline{x}_y - \overline{x} = r_в \cdot \frac{\sigma_x}{\sigma_y} \cdot (y - \overline{y}) \tag{6.4}$$

Пример 6.4

На основании полученных измерений найти линейную регрессию Y на X и X на Y.

X	4	6	8	10	12
Y	5	8	7	9	14

Решение.

Подставляя полученные данные в уравнения линии регрессии (4) получим:

Y на X

$$\overline{y}_x - 8{,}6 = 0{,}89 \cdot \frac{3{,}01}{2{,}83} \cdot (x - 8)$$

$$\overline{y}_x - 8{,}6 = 0{,}95 \cdot (x - 8)$$

$$\overline{y}_x = 0{,}95x - 7{,}6 + 8{,}6$$

$$\overline{y}_x = 0{,}95x + 1$$

X на Y

$$\overline{x}_y - 8 = 0{,}89 \cdot \frac{2{,}83}{3{,}01} \cdot (y - 8{,}6)$$

$$\overline{x}_y - 8 = 0{,}84(y - 8{,}6)$$

$$\overline{x}_y = 0{,}84y - 7{,}22 + 8{,}6$$

$$\overline{x}_y = 0{,}84y + 0{,}78$$

Примеры для самостоятельного решения

На основании полученных измерений найти линейную регрессию Y на X и X на Y и коэффициент корреляции:

1.

X	22	27	25	26	29	45
Y	53	29	39	37	37	38

2.

X	28	32	36	38	40	42
Y	18	12	20	34	30	42

3.

X	9	8,8	8	7,2	6,4	6
Y	6	5,4	4,2	3,2	2,8	2,6

4.

X	18,5	17,4	17	16,5	16,2	15,7
Y	8,8	9,2	9,4	10,6	14,4	14,6

5.

X	23,7	21,4	19,4	19,2	18,6	16,4
Y	2,9	3,7	3,7	3,8	3,9	5,3

Контрольный тест

1. Случайная величина – это

а) величина, которая принимает то или иное числовое значение, но заранее неизвестно какое именно;
б) величина, которая происходит только в различных условиях;
в) величина, которая при совокупности одних и тех же условий может произойти, а может не произойти;
г) величина причины, появления которой неизвестны.

2. Если событие невозможное, то вероятность

а) лежит между 0 и 1; б) равна 0;
в) близка к 1; г) равна 1.

3. Вероятность попадания в цель равна 0,3, а вероятность ее уничтожения 0,05. Найти вероятность того, что при попадании в цель она не будет уничтожена.

а) 0,1924; б) 0,015; в) 0,285; г) 0,17.

4. Игральный кубик подбрасывается один раз. Тогда вероятность того, что на верхней грани выпадет число очков больше трех, равно:

а) 1/3 б) 1/6 в) 2/3 г) 1/2

5. Гипотезами называют события, которые

а) являются независимыми и образуют полную группу;
б) являются несовместными;
в) являются несовместными и образуют полную группу;
г) образуют полную группу.

6. Имеется 8 синих и 5 красных шаров. Вытаскиваем наугад 2 шара. Какова вероятность того, что вытащенные шары красные?

а) $\frac{20}{169}$ б) $\frac{2}{5}$ в) $\frac{5}{39}$ г) $\frac{5}{8}$

7. Случайная величина называется дискретной, если она

а) принимает значения в некотором промежутке;

б) принимает конечное множество значений;

в) принимает конечное или бесконечное множество значений;

г) принимает конечное или бесконечное, но обязательно счетное множество значений.

8. Укажите условие нормировки дискретной случайной величины

а) $\sum_{i=1}^{n} p_i = 0$ б) $\int_{-\infty}^{+\infty} f(x)dx = 0$ в) $\sum_{i=1}^{n} p_i = 1$ г) $\int_{-\infty}^{+\infty} f(x)dx = 1$

9. Какая из формул определяет функцию распределения, если случайная величина дискретная?

а) $F(x) = \sum_{i=1}^{n} p_i$ б) $F(x) = \int_{-\infty}^{+\infty} f(x)dx = 0$

в) $F(x) = f\,'(x)$ г) нет верного ответа

10. Какая из формул определяет функцию распределения, если случайная величина непрерывная?

а) $F(x) = \sum_{i=1}^{n} p_i$ б) $F(x) = \int_{-\infty}^{+\infty} f(x)dx = 0$

в) $F(x) = f\,'(x)$ г) нет верного ответа

11. Функция распределения дискретной случайной величины представляет собой

а) сумму вероятностей;

б) сумму произведений всех возможных значений на соответствующие им вероятности;

в) производную функции плотности вероятности;

г) интеграл на всей числовой оси от функции плотности вероятности.

12. Графиком функции распределения непрерывной случайной величины является …

а) кривая распределения;

б) разрывная ступенчатая фигура, состоящая из отрезков, параллельных оси абсцисс;

в) многоугольник распределения;

г) неубывающая кривая в интервале от 0 до 1.

13. Определите математическое ожидание случайной величины X.

X	– 3	0	5
p_i	0,4	0,1	0,5

а) 3,7; б) 1,3 в) 2; г) 3,8.

14. Дан закон распределения дискретной случайной величины

x_i	1	2	3	4	5
p_i	0,14	0,28	0,17	0,32	?

Чему равно значение вероятности p_5?

а) 0,1; б) 0; в) 0,09; г) 0,2

15. Плотность вероятности случайной величины X, распределенной по экспоненциальному закону с параметром $\lambda = 2$, имеет вид:

а) $f(x) = \begin{cases} 0, & x < 0 \\ \ell^{-2x}, & x \geq 0 \end{cases}$

б) $f(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}\ell^{-2x}, & x \geq 0 \end{cases}$

в) $f(x) = \begin{cases} 0, & x < 0 \\ 2\ell^{-2x}, & x \geq 0 \end{cases}$

г) $f(x) = \begin{cases} 0, & x < 0 \\ -2\ell^{x}, & x \geq 0 \end{cases}$

16. Среди выражений:

1) центр распределения;
2) среднее значение;
3) плотность вероятности;
4) математическое ожидание

Синонимами являются

а) 1 и 4) б) все, кроме 3 в) все, кроме 1 г) 2 и 4

17. Площадь фигуры, ограниченной графиком плотности распределения и осью абсцисс, приближенно равна

а) 1; б) 0,5 в) 0,1 г) 100

18. Нормальный закон распределения (закон Гаусса) представлен

1) только для дискретной величины;
2) и для дискретной, и для непрерывной случайной величины;
3) только для непрерывной величины;
4) верного ответа нет.

а) 2 б) 3 в)1 г) 4

19. Какой символ пропущен в формуле закона Гаусса: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \ell^{-\frac{(x-*)^2}{2\sigma^2}}$

а) 1 б) m в) σ г) D(x)

20. Какая характеристика характеризует положение случайной величины?

а) M(X) б) D (x) в) σ г) p

21. Дисперсия независимых случайных величин X и Y соответственно равны D (X) = 2, D (Y) = 1. Найти дисперсию D(Z) случайной величины Z = X+2Y – 3.

а) D = 2 б) D = 5 в) D = 4 г) D = 6

22. Коэффициент корреляции двух независимых величин X и Y равен

а) 0 б) – 1 в) 1 г) ∞

23. К числовым характеристикам системы случайных величин не относятся:

а) условные математические ожидания
б) коэффициент корреляции
в) дисперсия
г) размещение

24. Указать ***неверное*** утверждение:

а) По многомерной функции распределения всегда можно найти одномерные (маргинальные) распределения отдельных компонент.

б) По одномерным (маргинальным) распределениям отдельных компонент всегда можно найти многомерную функцию распределения.

в) По многомерной функции плотности всегда можно найти одномерные (маргинальные) плотности распределения отдельных компонент.

25. К простейшим предельным теоремам теории вероятности не относятся:

а) неравенство Маркова
б) неравенство Чебышева
в) алгебра вероятностей событий
г) теорема Бернулли

26. Часть исследуемых объектов, выбранных случайным образом, называется …

а) генеральной совокупностью;
б) выборочной совокупностью;
в) вариационным рядом.
г) другой ответ

27. Вся совокупность исследуемых объектов, объединенных по определенному признаку, называется…

а) генеральной совокупностью;
б) выборочной совокупностью;
в) вариационным рядом.
г) другой ответ

28. Ранжированный ряд значений, выписанных в порядке возрастания или убывания, называют …
а) функцией распределения;
б) статистическим рядом распределения;
в) выборкой;
г) интервальный ряд распределения

29. «Варианты» означают …
а) относительные частоты;
б) значения случайной величины;
в) вероятности;
г) абсолютные частоты.

30. Полигоном частот называют…
а) ступенчатую фигуру;
б) ломанную фигуру;
в) график в отрезках;
г) плавную линию.

31. Число вариант в ряду распределения называют:
а) относительной частотой;
б) интервалом данной варианты;
в) весом;
г) не правильного ответа.

32. Для правильной организации сборки узла необходимо оценить вероятность, с которой размеры деталей отклоняются от середины поля допуска не более чем на 2 мм. Известно, что середина поля допуска совпадает с математическим ожиданием размеров обрабатываемых деталей, а среднеквадратическое отклонение равно 0,25 мм.

а) 0,94 б) 0,89 в) 0,98 г) 0,88

33. Допишите формулу нахождения среднего значения:

а) n б) n – 1 в) σ г) $\sqrt{2\pi}$

34. Уравнение регрессии имеет вид: Y = 5,1 – 1,7х оно показывает, что при увеличении Х на 1 единицу своего измерения Y в среднем:

а) увеличится на 1,7 единиц своего измерения
б) увеличится на 3,4 единиц своего измерения
в) уменьшится на 1,7 единиц своего измерения
г) уменьшится на 3,4 единиц своего измерения

35.Термин регрессия в статистике понимают как:

а) функцию связи, зависимости;
б) направление развития явления вспять;
в) функцию анализа случайных событий во времени;
г) уравнение линии связи.

а) а, б б) в, г в) а, г г) б, в

36. Установите соответствие:

a. r = – 0,3
b. r = 0,6
c. r = – 0,8
d. r = 0,8
e. r = 0,3

1) зависимость между X и Y сильная, возрастающая
2) зависимость между X и Y слабая, возрастающая
3) зависимость между X и Y сильная, убывающая
4) зависимость между X и Y слабая, убывающая
5) зависимость между X и Y средняя, возрастающая

37. Вставьте пропущенное слово:

Если случайная величина X принимает только ______________значения и имеет математическое ожидание *M(X),* то для любого положительного числа ε верно неравенство: $P(X > \varepsilon) \le \frac{M(X)}{\varepsilon}$.

а) неотрицательное;
б) положительное;
в) неположительное;
г) отрицательное.

38.Неравенство Чебышева

$$P(\left|X - M(x)\right| > \varepsilon) \le \frac{D(x)}{\varepsilon^2}$$

устанавливает:

а) нижнюю границу вероятности рассматриваемого случайного события;
б) верхнюю границу вероятности рассматриваемого случайного события;
в) левую границу вероятности рассматриваемого случайного события;
г) правую границу вероятности рассматриваемого случайного события.

39. Центральная предельная теорема устанавливает тот факт, что совокупность действие большого числа случайных величин приводит к определенному закону распределения, а именно…

а) биноминальному закону распределения;
б) показательному закону распределения;
в) равномерному закону распределения;
г) нормальному закону распределения.

40. В формуле Бернулли $\lim_{n\to\infty} P\left(\left|\frac{m_n}{n} - p\right| < \varepsilon\right) = 1$ m_n обозначает:

а) общее число испытаний;
б) вероятность успеха;
в) вероятность неуспеха;
г) число успехов в n испытаниях.

Приложение

Приложение 1

ЗНАЧЕНИЯ ФУНКЦИИ СТАНДАРТНОГО РАСПРЕДЕЛЕНИЯ $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot exp\left(-\frac{x^2}{2}\right)$

x	сотые доли x									
	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	0,3970	0,3965	0,3961	0,3956	0,3951	0,3945	0,3939	0,3932	0,3925	0,3918
0,2	0,3910	0,3902	0,3894	0,3885	0,3876	0,3867	0,3857	0,3847	0,3836	0,3825
0,3	0,3814	0,3802	0,3790	0,3778	0,3765	0,3752	0,3739	0,3726	0,3712	0,3697
0,4	0,3683	0,3668	0,3653	0,3637	0,3621	0,3605	0,3589	0,3572	0,3555	0,3538
0,5	0,3521	0,3508	0,3485	0,3467	0,3448	0,3429	0,3410	0,3391	0,3372	0,3352
0,6	0,3332	0,3312	0,3292	0,3271	0,3251	0,3230	0,3209	0,3187	0,3166	0,3144
0,7	0,3123	0,3101	0,3079	0,3056	0,3034	0,3011	0,2989	0,2966	0,2943	0,2920
0,8	0,2897	0,2874	0,2850	0,2827	0,2803	0,2780	0,2756	0,2732	0,2709	0,2685
0,9	0,2661	0,2637	0,2613	0,2589	0,2565	0,2541	0,2516	0,2492	0,2468	0,2444
1,0	0,2420	0,2396	0,2371	0,2347	0,2323	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	0,2179	0,2155	0,2131	0,2107	0,2083	0,2059	0,2036	0,2012	0,1989	0,1965
1,2	0,1942	0,1919	0,1895	0,1872	0,1849	0,1826	0,1804	0,1781	0,1758	0,1736
1,3	0,1714	0,1691	0,1669	0,1647	0,1626	0,1604	0,1582	0,1561	0,1539	0,1518
1,4	0,1497	0,1476	0,1456	0,1435	0,1415	0,1394	0,1374	0,1354	0,1334	0,1315
1,5	0,1295	0,1276	0,1257	0,1238	0,1219	0,1200	0,1182	0,1163	0,1145	0,1127
1,6	0,1109	0,1092	0,1074	0,1057	0,1040	0,1023	0,1006	0,0989	0,0973	0,0957
1,7	0,0940	0,0925	0,0909	0,0893	0,0878	0,0863	0,0848	0,0833	0,0818	0,0804
1,8	0,0790	0,0775	0,0761	0,0748	0,0734	0,0721	0,0707	0,0694	0,0681	0,0669
1,9	0,0656	0,0644	0,0632	0,0620	0,0608	0,0596	0,0584	0,0573	0,0562	0,0551
2,0	0,0540	0,0529	0,0519	0,0508	0,0498	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0,0440	0,0431	0,0422	0,0413	0,0404	0,0396	0,0387	0,0379	0,0371	0,0363
2,2	0,0355	0,0347	0,0339	0,0332	0,0325	0,0317	0,0310	0,0303	0,0297	0,0290
2,3	0,0283	0,0277	0,0270	0,0264	0,0258	0,0252	0,0246	0,0241	0,0235	0,0229
2,4	0,0224	0,0219	0,0213	0,0208	0,0203	0,0198	0,0194	0,0189	0,0184	0,0180
2,5	0,0175	0,0171	0,0167	0,0163	0,0158	0,0154	0,0151	0,0147	0,0143	0,0139
2,6	0,0136	0,0132	0,0129	0,0126	0,0122	0,0119	0,0116	0,0113	0,0110	0,0107
2,7	0,0104	0,0101	0,0099	0,0096	0,0093	0,0091	0,0088	0,0086	0,0084	0,0081
2,8	0,0079	0,0077	0,0075	0,0073	0,0071	0,0069	0,0067	0,0065	0,0063	0,0061
2,9	0,0060	0,0058	0,0056	0,0055	0,0053	0,0051	0,0050	0,0048	0,0047	0,0046
3,0	0,0044	0,0043	0,0042	0,0040	0,0039	0,0038	0,0037	0,0036	0,0035	0,0034
3,1	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026	0,0025	0,0025
3,2	0,0024	0,0023	0,0022	0,0022	0,0021	0,0020	0,0020	0,0019	0,0018	0,0018
3,3	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014	0,0013	0,0013
3,4	0,0012	0,0012	0,0012	0,0011	0,0011	0,0010	0,0010	0,0010	0,0009	0,0009
3,5	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007	0,0007	0,0007	0,0006
3,6	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005	0,0005	0,0005	0,0005	0,0004
3,7	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003	0,0003	0,0003	0,0003
3,8	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002	0,0002	0,0002	0,0002	0,0002
3,9	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0001	0,0001

Для **4,00** $\leq$ ***x*** $\leq$ **4,23** значение функции с точностью до четырех знаков после запятой равно 0,0001

Для ***x*** $\geq$ **4,24** значение функции с точностью до четырех знаков после запятой равно 0,0000

Приложение 2

ЗНАЧЕНИЯ НОРМИРОВАННОЙ ФУНКЦИИ ЛАПЛАСА $\Phi(x) = \int\limits_0^x \varphi(t)dt$

x	с о т ы е д о л и *x*									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,9778	0,4783	0,9788	0,4793	0,4798	0,4803	0,4808	0,4812	0,9817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4998	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999

Для $x \geq 3{,}9$ значение функции с точностью до четырех знаков после запятой равно 0,5000

Приложение 3

ЗНАЧЕНИЯ ФУНКЦИИ e^{-x} ДЛЯ $0 \le x < 1$, $e \approx 2{,}7183$

x	сотые доли x									
	0	1	2	3	4	5	6	7	8	9
0,0	1,0000	0,9900	0,9802	0,9704	0,9608	0,9512	0,9418	0,9324	0,9231	0,9139
0,1	0,9048	0,8958	0,8869	0,8781	0,8694	0,8607	0,8521	0,8437	0,8353	0,8270
0,2	0,8187	0,8106	0,8025	0,7945	0,7866	0,7788	0,7711	0,7634	0,7558	0,7483
0,3	0,7408	0,7334	0,7261	0,7189	0,7118	0,7047	0,6977	0,6907	0,6839	0,6771
0,4	0,6703	0,6637	0,6570	0,6505	0,6440	0,6376	0,6313	0,6250	0,6188	0,6126
0,5	0,6065	0,6005	0,5945	0,5886	0,5827	0,5769	0,5712	0,5655	0,5599	0,5543
0,6	0,5488	0,5434	0,5379	0,5326	0,5273	0,5220	0,5169	0,5117	0,5066	0,5016
0,7	0,4966	0,4916	0,4868	0,4819	0,4771	0,4724	0,4677	0,4630	0,4584	0,4538
0,8	0,4493	0,4449	0,4404	0,4360	0,4317	0,4274	0,4232	0,4190	0,4148	0,4107
0,9	0,4066	0,4025	0,3985	0,3946	0,3906	0,3867	0,3829	0,3791	0,3753	0,3716

Примечание

1. Значения функции e^{-x}, содержащей только тысячные доли в показателе, приведены в таблице:

$e^{-0{,}001} = 0{,}9990$	$e^{-0{,}004} = 0{,}9960$	$e^{-0{,}007} = 0{,}9930$
$e^{-0{,}002} = 0{,}9980$	$e^{-0{,}005} = 0{,}9950$	$e^{-0{,}008} = 0{,}9920$
$e^{-0{,}003} = 0{,}9970$	$e^{-0{,}006} = 0{,}9940$	$e^{-0{,}009} = 0{,}9910$

2. При расчете значений функций с показателем степени, содержащим десятые, сотые и тысячные доли, можно использовать обе вышеприведенные таблицы. Например,

$e^{-0{,}825} = e^{-0{,}82} \cdot e^{-0{,}005} \approx 0{,}4404 \cdot 0{,}9950 \approx 0{,}4382$.

3. При расчете значений функции e^{-x} при $x \ge 1$ можно использовать алгебраические преобразования. Например,

$e^{-1} \approx 1/2{,}7183 \approx 0{,}3679$;

$e^{-1{,}5} = e^{-1} \cdot e^{-0{,}5} \approx 0{,}3679 \cdot 0{,}6065 \approx 0{,}2231$ или $e^{-1{,}5} = e^{-0{,}75} \cdot e^{-0{,}75} \approx (0{,}4724)^2 \approx 0{,}2231$;

$e^{-3{,}5} = e^{-3} \cdot e^{-0{,}5} \approx (0{,}3679)^3 \cdot 0{,}6065 \approx 0{,}0302$.

4. Для получения значения функции с любой требуемой точностью можно использовать формулу разложения этой функции в ряд Маклорена, сходящийся для всех x:

$e^{-x} = \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^n}{n!} = 1 - \frac{x}{1!} + \frac{x^2}{2!} - \frac{x^3}{3!} + \ldots$ (при этом погрешность получаемого значения функции определяется абсолютной величиной первого отброшенного члена ряда).

5. При расчете значений функции с положительным показателем можно воспользоваться соотношением $e^{a} \approx 1/e^{-a}$. Например, $e^{0{,}825} = 1/e^{-0{,}825} \approx 1/0{,}4382 \approx 2{,}282$.

Приложение 4

ЗНАЧЕНИЯ ПАРАМЕТРА ТОЧНОСТИ ОЦЕНКИ СТАНДАРТНОГО ОТКЛОНЕНИЯ НОРМАЛЬНОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ

$$q = q(\gamma, n)$$

(n – объем выборки, γ – доверительная вероятность)

n \ γ	$\gamma = 0{,}95$	$\gamma = 0{,}99$	$\gamma = 0{,}999$	n \ γ	$\gamma = 0{,}95$	$\gamma = 0{,}99$	$\gamma = 0{,}999$
5	1,37	2,67	5,64	**20**	0,37	0,58	0,88
6	1,09	2,01	3,88	**25**	0,32	0,49	0,73
7	0,92	1,62	2,98	**30**	0,28	0,43	0,63
8	0,80	1,38	2,42	**35**	0,26	0,38	0,56
9	0,71	1,20	2,06	**40**	0,24	0,35	0,50
10	0,65	1,08	1,80	**45**	0,22	0,32	0,46
11	0,59	0,98	1,60	**50**	0,21	0,30	0,43
12	0,55	0,90	1,45	**60**	0,188	0,269	0,38
13	0,52	0,83	1,33	**70**	0,174	0,245	0,34
14	0,48	0,78	1,23	**80**	0,161	0,226	0,31
15	0,46	0,73	1,15	**90**	0,151	0,211	0,29
16	0,44	0,70	1,07	**100**	0,143	0,198	0,27
17	0,42	0,66	1,014	**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

Примечание

Точность оценки стандартного отклонения нормальной случайной величины генеральной совокупности определяется значением $s \cdot q$, то есть, интервальная оценка (доверительный интервал) для стандартного отклонения определяется как $s \cdot (1 - q) < \sigma_0 < s \cdot (1 + q)$, где s – исправленное выборочное стандартное отклонение. Поскольку по определению σ_0 неотрицательная величина, то в случае $q > 1$ интервальную оценку для стандартного отклонения σ_0 нормальной случайной величины генеральной совокупности следует определять как $0 < \sigma_0 < s \cdot (1 + q)$.

Приложение 5

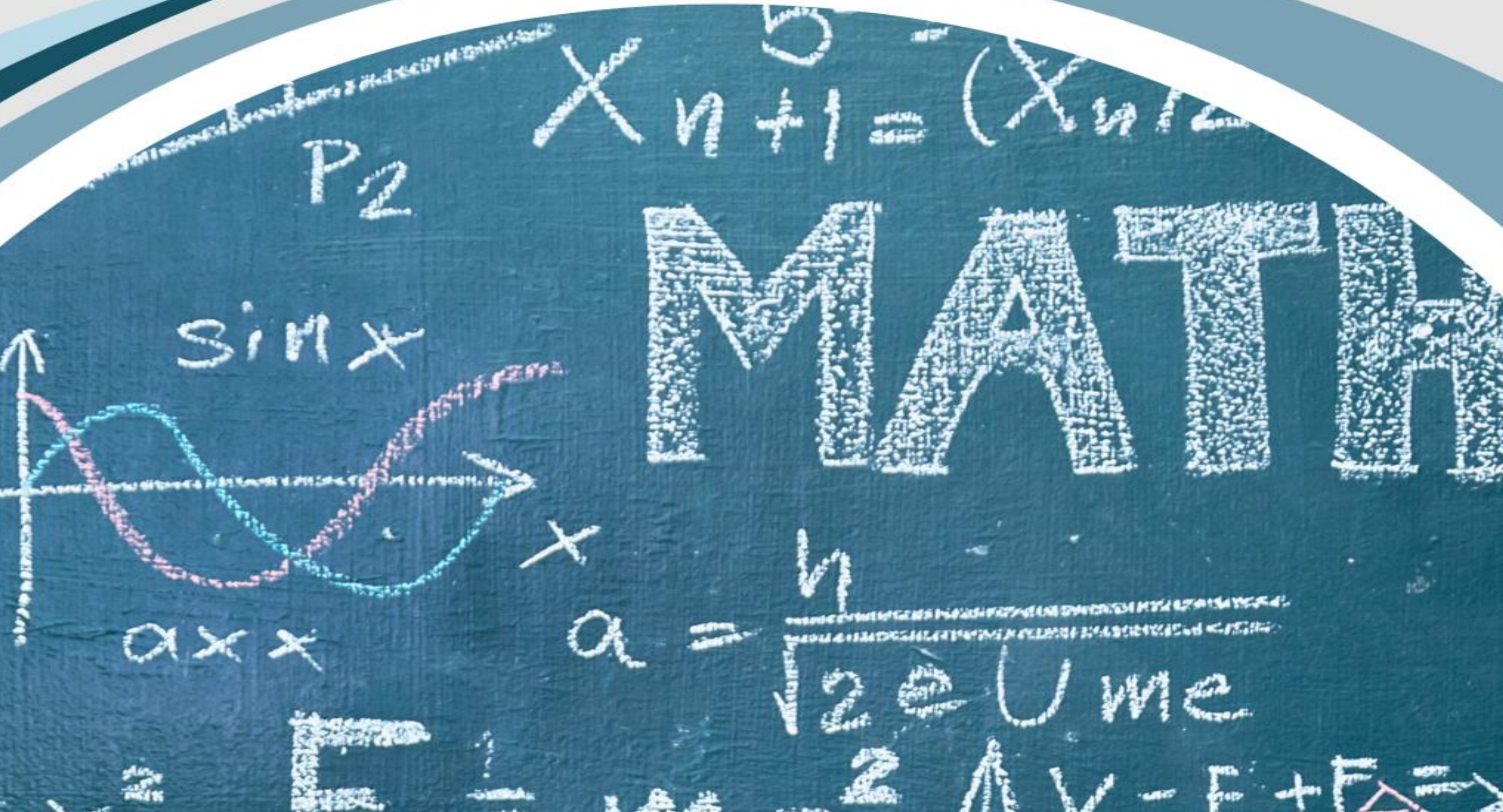
ЗНАЧЕНИЯ КОЭФФИЦИЕНТОВ СТЬЮДЕНТА $t_\gamma = t(\gamma, n)$

(n – объем выборки, γ – доверительная вероятность)

n \ γ	$\gamma = 0{,}8$	$\gamma = 0{,}9$	$\gamma = 0{,}95$	$\gamma = 0{,}98$	$\gamma = 0{,}99$	$\gamma = 0{,}999$
3	1,886	2,920	4,303	6,965	9,925	31,599
4	1,638	2,353	3,182	4,541	5,841	12,924
5	1,533	2,132	2,776	3,747	4,604	8,610
6	1,476	2,015	2,571	3,365	5,032	6,859
7	1,440	1,943	2,447	3,143	3,707	5,959
8	1,415	1,895	2,365	2,998	3,499	5,405
9	1,397	1,860	2,306	2,896	3,355	5,401
10	1,383	1,833	2,262	2,821	3,250	4,781
11	1,372	1,812	2,228	2,764	3,169	4,587
12	1,363	1,796	2,201	2,718	3,106	4,437
13	1,356	1,782	2,179	2,681	3,055	4,318
14	1,350	1,771	2,160	2,650	3,012	4,221
15	1,345	1,761	2,145	2,624	3,977	4,140
16	1,341	1,753	2,131	2,602	2,947	4,073
17	1,337	1,746	2,120	2,583	2,921	4,015
18	1,333	1,740	2,110	2,567	2,898	3,965
19	1,330	1,734	2,101	2,552	2,878	3,922
20	1,328	1,729	2,093	2,539	2,861	3,883
21	1,325	1,725	2,086	2,528	2,845	3,850
22	1,323	1,721	2,080	2,518	2,831	3,819
23	1,321	1,717	2,074	2,508	2,819	3,792
24	1,319	1,714	2,069	2,500	2,807	3,767
25	1,318	1,711	2,064	2,492	2,797	3,745
26	1,316	1,708	2,060	2,485	2,787	3,725
27	1,315	1,706	2,056	2,479	2,779	3,707
28	1,314	1,703	2,052	2,473	2,771	3,690
29	1,313	1,701	2,048	2,467	2,763	3,674
30	1,311	1,699	2,045	2,462	2,756	3,659
35	1,307	1,692	2,032	2,443	2,720	3,600
40	1,304	1,685	2,023	2,426	2,708	3,558
45	1,301	1,681	2,016	4,415	2,692	3,527
50	1,299	1,677	2,009	4,405	2,679	3,502
60	1,296	1,672	2,001	2,391	2,662	3,464
70	1,294	1,668	1,996	2,383	2,649	3,439
80	1,292	1,664	1,991	2,376	2,640	3,418
90	1,291	1,662	1,987	2,370	2,633	3,403
100	1,290	1,660	1,984	2,365	2,627	3,392
120	1,289	1,658	1,980	2,358	2,617	3,374
150	1,288	1,656	1,976	2,353	2,609	3,357
200	1,286	1,653	1,972	2,345	2,601	3,340
500	1,283	1,648	1,965	2,334	2,586	3,310
∞	1,282	1,645	1,960	2,326	2,576	3,291

Литература

1. Васильев, А. А. Теория вероятностей и математическая статистика: учебник и практикум для вузов / А. А. Васильев. — 2-е изд., испр. и доп. — Москва: Издательство Юрайт, 2025. — 224 с.
2. Волощук, В. А. Теория вероятностей и математическая статистика: шпаргалка: учебное пособие: [16+] / В. А. Волощук ; Научная книга. – 2-е изд. – Саратов: Научная книга, 2020. – 48 с.
3. Гмурман, В. Е. Теория вероятностей и математическая статистика: учебник для вузов / В. Е. Гмурман. — 12-е изд. — Москва: Издательство Юрайт, 2025. — 479 с.
4. Гусева, Е. Н. Теория вероятностей и математическая статистика: учебное пособие: [16+] / Е. Н. Гусева. – 7-е изд., стер. – Москва: ФЛИНТА, 2021. – 220 с.
5. Калинина, В. Н. Теория вероятностей и математическая статистика: учебник для вузов / В. Н. Калинина. — 2-е изд., перераб. и доп. — Москва: Издательство Юрайт, 2023. — 472 с.
6. Ковалев, Е. А. Теория вероятностей и математическая статистика для экономистов: учебник и практикум для вузов / Е. А. Ковалев, Г. А. Медведев; под общей редакцией Г. А. Медведева. — 2-е изд., испр. и доп. — Москв: Издательство Юрайт, 2025. — 284 с.
7. Коган, Е. А. Теория вероятностей и математическая статистика: учебник / Е. А. Коган, А. А. Юрченко. — Москва: ИНФРА-М, 2023. — 250 с.
8. Колданов, А. П. Теория вероятностей и математическая статистика: учебник: [16+] / А. П. Колданов, П. А. Колданов. – Москва: Издательский дом Высшей школы экономики, 2023. – 249 с.
9. Кремер, Н. Ш. Теория вероятностей и математическая статистика: учебник и практикум для вузов / Н. Ш. Кремер. — 5-е изд., перераб. и доп. — Москва: Издательство Юрайт, 2025. — 538 с.
10. Лунгу, К. Н. Высшая математика. Руководство к решению задач. Ч. 2: Учебное пособие / Лунгу К.Н., Макаров Е.В., - 2-е изд. - Москва: ФИЗМАТЛИТ, 2015. - 384 с.
11. Наливайко, Л. В. Комбинаторика, теория вероятностей и математическая статистика: учебное пособие / Л.В. Наливайко, Д.С. Шунскайте. - Москва: ИНФРА-М, 2024. - 296 с.
12. Сапожников, П. Н. Теория вероятностей, математическая статистика в примерах, задачах и тестах: учебное пособие / П.Н. Сапожников, А.А. Макаров, М.В. Радионова. — Москва: КУРС: ИНФРА-М, 2022. — 496 с.



ISBN 978-5-908003-08-7

Усл. печ. л 3,0
Объем издания 19,0 MB
Оформление электронного издания:
НОО Профессиональная наука, mail@scipro.ru
Дата размещения: 15.08.2025 г.
URL: http://scipro.ru/conf/mathematical_statistics08_25.pdf