Министерство науки и высшего образования Российской Федерации
Министерство внутренних дел Российской Федерации

Московский университет Министерства внутренних дел
Российской Федерации имени В.Я. Кикотя

МАТЕМАТИКА ДЛЯ ПСИХОЛОГОВ: ЭЛЕМЕНТЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Учебное пособие

Москва
Московский университет
МВД России имени В.Я. Кикотя

2021

УДК 519.2
ББК 22.17
М34

Рецензенты:
начальник кафедры информационно-компьютерных технологий в деятельности ОВД Белгородского юридического института МВД России имени И.Д. Путилина кандидат технических наук, доцент **А. Н. Прокопенко**; доцент кафедры трасологии и баллистики УНК ЭКД Волгоградской академии МВД России кандидат юридических наук **А. А. Нурушев**

Коллектив авторов:
Н. В. Задохина, Н. М. Дубинина, А. А. Страхов, Я. В. Цыганкова

М34 **Математика для психологов: элементы теории вероятностей и математической статистики** : учебное пособие / [Н. В. Задохина и др.]. – М. : Московский университет МВД России имени В.Я. Кикотя, 2021. – 89 с.
ISBN 978-5-9694-1012-1

Учебное пособие нацелено на повышение качества освоения обучающимися методов теории вероятностей и математической статистики, применяемых при решении прикладных психологических задач.

Пособие может быть использовано в процессе обучения курсантов, слушателей и студентов, а также для подготовки научно-педагогических и научных кадров по направлению «Психология служебной деятельности» в образовательных организациях МВД России.

УДК 519.2
ББК 22.17

ISBN 978-5-9694-1012-1

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ

Широкое применение методов теории вероятностей и математической статистики в различных областях познания обусловлено тем, что данные методы располагают понятийным аппаратом, который по своей широте и универсальности приближается к философскому, отражая общие количественные характеристики качественно различных явлений и предметов[1].

Очевидно, что методы теории вероятностей и математической статистики позволяют изучать не только количественные характеристики объектов и процессов, а также выявлять различные закономерности и взаимосвязи.

Кроме того, реализация статистических методов непосредственно связана с процессом использования современных информационных технологий. Именно такой синтез лежит в основе решения прикладных психологических задач.

Среди программных продуктов, реализующих методы статистического анализа, особой популярностью пользуются программы Statistica и SPSS Statistics. Однако данные программы, обладая мощными вычислительными возможностями, стоят весьма недешево. Поэтому, если в рамках психологического исследования необходимо автоматизировать процесс статистического наблюдения, реализацию некоторых статистических методов, а также визуализировать результаты исследований, целесообразно использовать табличный процессор MS Excel. Средствами MS Excel можно производить весь цикл обработки данных психологических исследований, начиная от сбора, регистрации, обработки и заканчивая их эффектной визуализацией[2].

---

[1] Задохина Н. В. Статистические методы в деятельности эксперта-криминалиста // Образование. Наука. Научные кадры. 2020. № 3. С. 195.

[2] Терентьева А. Н., Задохина Н. В. Автоматизация сбора и обработки данных психологических исследований // Совершенствование профессиональной подготовки психологов для подразделений органов внутренних дел : Межведомственная научно-практическая конференция : сборник материалов. М., 2020. С. 419.

ГЛАВА 1. Теория вероятностей

1.1. Основные понятия и определения

Теория вероятностей – это математическая дисциплина, изучающая закономерности в случайных массовых процессах.

Зарождение теории вероятностей относится к середине XVII в. и основано на работах Б. Паскаля, П. Ферма, Х. Гюйгенса в области теории азартных игр. Дальнейшее развитие теории вероятностей связано с именами Я. Бернулли, П. Лапласа, К. Гаусса, П. Чебышева, А. Ляпунова, А. Маркова, А. Колмогорова и многих других ученых.

Исходными понятиями теории вероятности являются следующие: испытание, событие, вероятность события.

Испытание – осуществление некоторого комплекса условий, в результате которого происходят или не происходят определенные события.

Примеры испытаний: подбрасывание монеты, выбор игральной карты, сдача экзамена и т. д.

Событие – всякий возможный исход испытания. Обычно события обозначают прописными буквами латинского алфавита A, B, C и т. д., снабжая их при необходимости индексами ($A_1, A_2 \ldots$).

Примеры событий: A – появление орла при бросании монеты, B – наступление четверга после пятницы, C – кипение воды при $t = 100\ ^{o}C$ и т. д.

Очевидно, что все эти события отличаются возможностью их появления. Так, событие C обязательно происходит, событие B никогда не происходит, а событие A может произойти или не произойти в результате испытания. В этой связи различают три типа событий: достоверное, невозможное, случайное.

Достоверное событие – это событие, которое в результате испытания наступит обязательно (событие C).

Невозможное событие – это событие, которое в результате испытания никогда не наступит (событие B).

Случайное событие – это событие, которое в результате испытания может либо наступить, либо не наступить (событие A).

По своему логическому содержанию события могут быть *элементарными* (*простыми*) и *составными* (*сложными*). Составные события всегда можно разложить на элементарные события. Элементарные события, в свою очередь, являются простейшими неразложимыми *исходами* испытания.

Элементарные события всегда взаимно исключают друг друга, и в результате испытания обязательно происходит одно и только одно из этих элементарных событий.

Элементарные события обозначают строчной буквой греческого алфавита *омега* – ω_i. Всю совокупность элементарных событий (исходов испытания) обозначают прописной буквой греческого алфавита *омега* – $\Omega = \{\omega_i\}$ и называют *пространством элементарных событий*. Очевидно, что с каждым испытанием связано свое пространство элементарных событий.

Пример 1. Пусть испытание заключается в том, что из четырех карточек, пронумерованных цифрами 1, 2, 3, 4, наугад выбирается одна карточка. Пространство элементарных событий состоит из четырех элементарных событий: вынули карточку с номером 1 (ω_1), вынули карточку с номером 2 (ω_2), вынули карточку с номером 3 (ω_3), вынули карточку с номером 4 (ω_4), поэтому $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$.Элементарные события ω_1, ω_2, ω_3, ω_4 полностью описывают возможные исходы испытания и взаимно исключают друг друга. На элементарных событиях ω_1, ω_2, ω_3, ω_4 строятся более сложные (составные) события.

Пусть A – событие, состоящее в том, что наугад выбранная карточка имеет четный номер. Очевидно, что событие A осуществляется тогда, когда осуществляется одно из элементарных событий ω_2, ω_4, т. е. $A = \{\omega_2, \omega_4\}$. Элементарные исходы ω_2, ω_4 называют исходами, *благоприятствующими* событию A.

Таким образом, любое случайное событие представляет собой подмножество множества Ω.

Для понимания введенных понятий полезна геометрическая интерпретация с помощью диаграмм Эйлера-Венна. Представим пространство элементарных Ω событий в виде прямоугольника. Тогда каждому из элементарных событий ω_1, ω_2, ω_3, ω_4 соответствует точка внутри прямоугольника, а случайному событию A соответствует некоторая область, являющаяся подмножеством множества Ω (рис. 1.1).

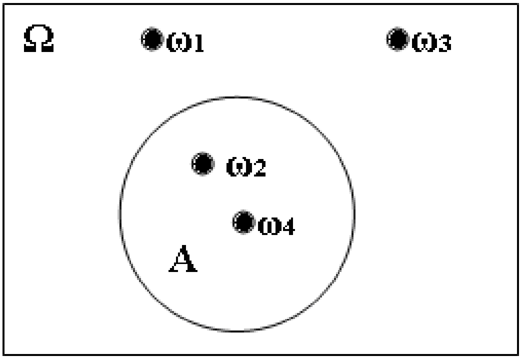


Рис. 1.1. Событие A

Следует заметить, что все пространство элементарных исходов $\Omega = \{\omega_i\}$ естественно назвать *достоверным событием*, так как оно наступает *при любом исходе* испытания. *Невозможным событием* в используемой терминологии можно считать событие, противоположное достоверному событию $\overline{\Omega} = \varnothing$[1], так как оно не наступает *ни при каком исходе* испытания.

По характеру появления в испытаниях можно выделить *совместные* и *несовместные* события.

События A и B называются **несовместными**, если в результате испытания они не могут появиться одновременно, т. е. наступление одного из них исключает наступление другого. Например, элементарные исходы при подбрасывании монеты

[1] Символ Ø означает пустое множество.

несовместны, так как выпадение орла исключает выпадение решки.

События A и B называются **совместными**, если в результате испытания они могут появиться одновременно, т. е. наступление одного из них не исключает возможность наступления другого.

Например, два стрелка стреляют в одну цель, попадание одного из них не исключает попадания другого.

Несколько событий $A_1, A_2, \dots A_n$ называются **попарно несовместными**, если любые два из них несовместны.

Например, при подбрасывании игральной кости никакие два элементарных исхода (появление на верхней грани 1, 2, 3, 4, 5, 6 очков) не могут произойти одновременно.

Событие $\overline{A}$ называется **противоположным** событию A и означает, что событие A не происходит. Например, в опыте с приобретением лотерейных билетов событию A – «ни один из билетов не выиграл» противоположно событие $\overline{A}$ – «хотя бы один из билетов выиграл».

События $A_1, A_2, \dots A_n$ образуют *полную группу* несовместных событий, если они попарно несовместны и в результате данного испытания происходит одно и только одно из этих событий. Противоположные события также образуют *полную группу*.

Из событий с помощью определенных операций можно образовывать новые события. Поскольку математической моделью случайных событий являются множества, то операции над событиями будут аналогичны операциям над множествами. Ниже рассмотрим каждую операцию подробно.

Суммой (или объединением) событий A и B называется такое событие $A + B$ (или $A \cup B$), которое состоит в том, что наступит хотя бы одно из событий: или A, или B, или оба события одновременно, если это возможно (рис. 1.2).

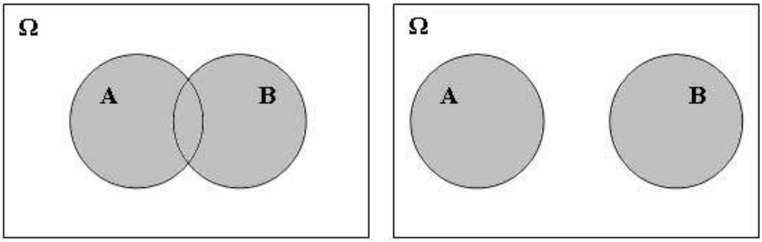


Рис. 1.2. Сумма совместных и несовместных событий

Пример 2. Пусть событие A – «идет дождь», событие B – «дует ветер». События A и B совместны, поэтому $A + B$ означает, что или «идет дождь, нет ветра», или «дует ветер, нет дождя», или «и дождь идет, и дует ветер».

Пример 3. Бросают игральную кость. Пусть событие A – «выпадет три очка», событие B – «выпадет шесть очков». События A и B несовместны, тогда событие $A + B$ означает, что или «выпадет три очка», или «выпадет шесть очков».

Произведением (или пересечением) событий A и B называется такое событие $A{\cdot}B$ (или $A{\cap}B$), которое состоит в том, что события A и B наступят одновременно (рис. 1.3).

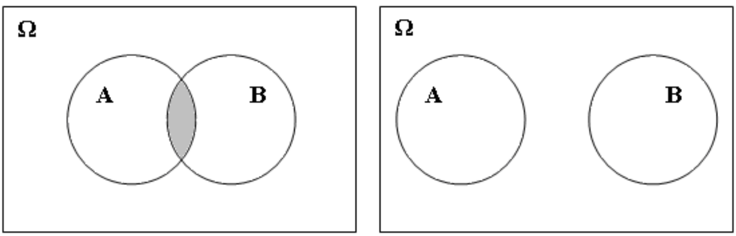


Рис. 1.3. Произведение совместных и несовместных событий

Очевидно, что если события A и B несовместны, то $A{\cdot}B = \varnothing$.

Пример 4. Пусть событие A – «идет дождь», событие B – «дует ветер». События A и B совместны, поэтому $A{\cdot}B$ означает, что «и дождь идет, и дует ветер».

Пример 5. Бросают игральную кость. Пусть событие A – «выпадет три очка», событие B – «выпадет шесть очков». События A и B несовместны, тогда событие $A \cdot B = \varnothing$.

Следует заметить, что для указания операций умножения и сложения используются также логические связки «и» и «или» соответственно. Поэтому, если говорят: «A и B», то это равносильно $A \cdot B$, если говорят: «A или B», то это равносильно $A + B$.

Разностью событий A и B называется такое событие $A - B$ ($A \backslash B$), которое состоит в том, что событие A происходит, а событие B не происходит $A \backslash B$ (рис. 1.4).

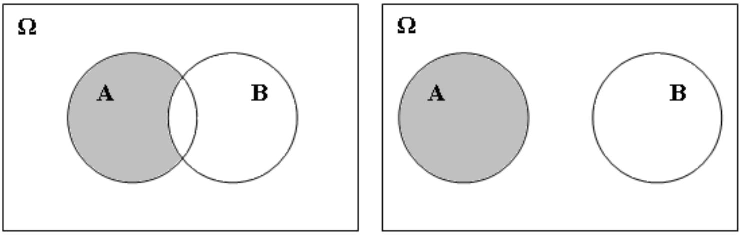


Рис. 1.4. Разность совместных и несовместных событий

Пример 6. Пусть событие A – «идет дождь», событие B – «дует ветер». События A и B совместны, поэтому $A - B$ означает, что «идет дождь, нет ветра».

Пример 7. Бросают игральную кость. Пусть событие A – «выпадет три очка», событие B – «выпадет шесть очков». События A и B несовместны, тогда событие $A - B$ означает, что «выпадет три очка».

Отрицанием события A называется событие $\overline{A}$ (не A), которое означает, что событие A не происходит ($A + \overline{A} = \Omega$). Причем, если в результате испытания может произойти событие A, то может произойти и противоположное ему событие $\overline{A}$ (рис. 1.5). Событие $\overline{A}$ следует рассматривать как *дополнение* A до Ω.

Пример 8. Пусть событие A – «идет дождь», тогда событие $\overline{A}$ – «не идет дождь».

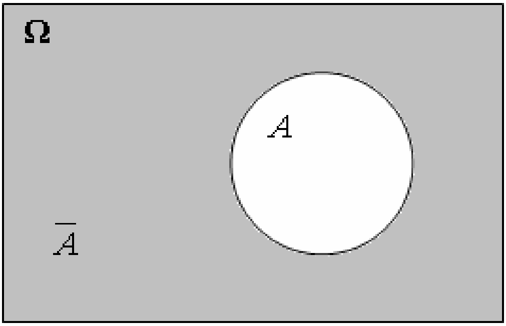


Рис. 1.5. Отрицание события

Пример 9. Пусть событие A – «студент сдал зачет по истории», событие B – «студент сдал зачет по философии». Используя соответствующие операции, запишите выражения для следующих событий: 1) «студент сдал оба зачета»; 2) «студент сдал хотя бы один зачет»; 3) «студент не сдал оба зачета»; 4) «студент сдал только зачет по истории»; 5) «студент сдал только зачет по философии»; 6) «студент сдал только один зачет»; 7) «студент не сдал хотя бы один зачет»; 8) «студент сдал не более одного зачета».

Решение:

1) событие «студент сдал оба зачета» означает, что «студент сдал зачет по истории и сдал зачет по философии», поэтому данное событие можно записать в виде: $A \cdot B$;

2) событие «студент сдал хотя бы один зачет» означает, что «студент сдал зачет по истории или сдал зачет по философии или сдал оба зачета», поэтому данное событие можно записать в виде: $A + B$;

3) событие «студент не сдал оба зачета» означает, что «студент не сдал зачет по истории и не сдал зачет по философии», поэтому данное событие можно записать в виде: $\overline{A} \cdot \overline{B}$;

4) событие «студент сдал только зачет по истории» означает, что «студент сдал зачет по истории и не сдал зачет по философии», поэтому данное событие можно записать в виде: $A \cdot \overline{B}$ или $A - B$;

5) событие «студент сдал только зачет по философии» означает, что «студент не сдал зачет по истории и сдал зачет по философии», поэтому данное событие можно записать в виде: $\overline{A}\cdot B$ или $B-A$;

6) событие «студент сдал только один зачет» означает, что «студент сдал только зачет по истории или сдал только зачет по философии», поэтому данное событие можно записать в виде: $A\cdot\overline{B}+\overline{A}\cdot B$;

7) событие «студент не сдал хотя бы один зачет» означает, что «студент не сдал зачет по истории или не сдал зачет по философии или не сдал оба зачета», поэтому данное событие можно записать в виде: $\overline{A}+\overline{B}$;

8) событие «студент сдал не более одного зачета» означает, что «студент сдал только один зачет или оба зачета не сдал», поэтому данное событие можно записать в виде: $A\cdot\overline{B}+\overline{A}\cdot B+\overline{A}\cdot\overline{B}$.

Следует заметить, что событие «студент сдал не более одного зачета» является противоположным событию «студент сдал оба зачета», поэтому данное событие можно записать также в виде: $\overline{A\cdot B}$.

$\overline{A\cdot B}$ означает: «неверно, что студент сдал оба зачета».

Очевидно, что различные события можно сравнивать по степени их возможного появления.

Вероятность – это численная мера объективной возможности наступления некоторого события. Вероятность события A обозначается $P(A)$ (от англ. probability – вероятность).

Классическое определение вероятности

Классический метод подсчета вероятностей применяется в том случае, если по условию задачи:

1) проводится одно испытание;
2) число исходов испытания конечно;
3) все исходы испытания равновозможны[1];

[1] События называются равновозможными, если есть основания считать, что ни одно из этих событий не является более возможным, чем другие. Например, элементарные исходы (орел и решка) при подбрасывании монеты являются равновозможными.

4) исходы образуют полную группу несовместных событий (взаимно исключают друг друга, и в результате испытания обязательно наступает один из исходов).

В этом случае вероятность случайного события A равна:

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

где m – число исходов, благоприятствующих наступлению события A, а n – число всех возможных исходов испытания.

Основные свойства вероятности:

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

2. $P(\Omega) = \frac{n}{n} = 1$ – вероятность достоверного события, так как $m = n$.

3. $P(\varnothing) = P(\overline{\Omega}) = \frac{0}{n} = 0$ – вероятность невозможного события, так как $m = 0$.

Пример 10. На 100 карточках написаны числа от 1 до 100. Определить вероятность того, что на случайно взятой карточке содержится цифра 5.

Решение. Испытание – случайным образом выбирают карточку. Событие A – на карточке содержится цифра 5. Число всех возможных исходов испытания конечно и равно n = 100 (карточки с числами от 1 до 100). Число исходов, благоприятствующих событию A, равно m = 19 (карточки с числами 5, 15, 25, 35, 45, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 65, 75, 85, 95).

Исходы испытания образуют полную группу несовместных событий и равновозможны. Отсюда следует:

$$P(A)=\frac{m}{n}=\frac{19}{100}=0{,}19.$$

Пример 11. При наборе телефонного номера абонент забыл последнюю цифру и набрал ее наугад, помня только, что эта цифра нечетная. Найти вероятность того, что номер набран правильно.

Решение. Испытание – абонент наугад набирает одну нечетную цифру. Событие A – цифра набрана правильно. Число всех возможных исходов испытания конечно и равно $n = 5$ (все нечетные цифры 1, 3, 5, 7, 9). Число исходов, благоприятствующих событию A, равно 1, т. е. $m = 1$ (правильна только одна нечетная цифра).

Исходы испытания образуют полную группу несовместных событий и равновозможны. Отсюда следует:

$$P(A) = \frac{m}{n} = \frac{1}{5} = 0{,}2.$$

Нахождение вероятности по формуле $P(A) = \frac{m}{n}$ называют также *непосредственным подсчетом вероятности*. При подсчете числа возможных исходов n и числа благоприятных исходов m часто используются некоторые правила и формулы комбинаторики.

Пример 12. Замок имеет четырехзначный цифровой код. Наугад набираются четыре цифры. Какова вероятность открыть при этом замок, если известно, что в коде все цифры различны?

Решение. Испытание – наугад набирается четырехзначный цифровой код. Событие A – код набран правильно.

Число всех возможных исходов испытания n конечно и равно числу всех возможных комбинаций из четырех цифр. Поскольку в коде замка важен не только набор цифр, но и их порядок, то для подсчета всех возможных комбинаций используем комбинаторную формулу *размещений без повторений*:

$$n = A_{10}^{4} = \frac{10!}{(10-4)!} = \frac{6!\cdot 7\cdot 8\cdot 9\cdot 10}{6!} = 5\,040.$$

Число исходов, благоприятствующих событию A, равно 1, т. е. m = 1 (правильна только одна комбинация из четырех цифр).

Исходы испытания образуют полную группу несовместных событий и равновозможны. Отсюда следует:

$$P(A) = \frac{m}{n} = \frac{1}{5040} \approx 0{,}0002.$$

Следует заметить, что $n!$ – факториал числа n (от лат. factor – множитель). $n! = 1\cdot 2\cdot 3\ldots\cdot n$ (произведение натуральных чисел до n включительно). Факториал определен только для целых неотрицательных чисел. Полагают: $0! = 1$; $1! = 1$; $2! = 1 \cdot 2 = 2$; $3! = 1 \cdot 2 \cdot 3 = 6$ и т. д.

Пример 13. Для участия в лотерее «Спортлото» на карточке, содержащей 49 чисел, нужно отметить 6 чисел. Какова вероятность выиграть?

Решение. Испытание – наугад выбирают 6 чисел из 49. Событие A – все 6 чисел выбраны правильно. Число всех возможных исходов испытания n конечно и равно числу всех возможных сочетаний из 49 чисел по 6. Поскольку не важен порядок заполнения карточки, то для подсчета всех возможных комбинаций используем комбинаторную формулу *сочетаний без повторений*:

$$n = C_{49}^{6} = \frac{49!}{(49-6)! \cdot 6!} = \frac{43! \cdot 44 \cdot 45 \cdot 46 \cdot 47 \cdot 48 \cdot 49}{43! \cdot 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6} = 13\,983\,816.$$

Число исходов, благоприятствующих событию A, равно 1, т. е. $m = 1$ (правильна только одна комбинация из шести цифр).

Исходы испытания образуют полную группу несовместных событий и равновозможны. Отсюда следует:

$$P(A) = \frac{m}{n} = \frac{1}{13983816}.$$

Пример 14. Из 15 студентов учебной группы 8 занимаются научной работой на кафедре экономического анализа, остальные на кафедре статистики. Какова вероятность, что из 9 случайно отобранных студентов 5 занимаются научной работой на кафедре статистики?

Решение. Испытание – случайным образом выбирают 9 студентов. Событие A – среди 9 случайно отобранных студентов 5 занимаются научной работой на кафедре статистики.

Число всех возможных исходов испытания n конечно и равно числу всех возможных сочетаний из 15 чисел по 9. Поскольку не важен порядок отбора студентов, то $n = C_{15}^{9}$.

Число исходов, благоприятствующих событию A, можно посчитать так: выбор 5 студентов с кафедры статистики из 7 имеющихся можно осуществить C_7^5 способами и одновременно выбор еще 4 студентов с кафедры экономического анализа из 8 имеющихся можно осуществить C_8^4 способами. Согласно комбинаторному правилу умножения, число исходов благоприятствующих событию A равно $m = C_7^5 \cdot C_8^4$.

Поскольку выбор случаен, то все исходы равновозможны. Очевидно, что в данном случае применимо классическое определение вероятности:

$$P(A) = \frac{m}{n} = \frac{C_7^5 \cdot C_8^4}{C_{15}^9} = \frac{21 \cdot 70}{5005} \approx 0{,}29.$$

Геометрическое определение вероятности

Пусть проводится опыт с бесконечным числом равновозможных исходов. Например, материальная точка бросается на некоторую область D. Требуется определить вероятность того, что материальная точка окажется в меньшей области D_A, которая целиком находится в области D.

Исходами данного испытания являются координаты местоположения «брошенной» точки. Очевидно, что область D представляет собой непрерывное множество исходов рассматриваемого опыта. Заметим, что вероятность попадания точки в область D_A не зависит от формы и местоположения области D_A в области D, а только от размера области D_A. Таким образом, *вероятность события A* в данном испытании вычисляется по формуле:

$$P(A) = \frac{D_A}{D},$$

где D_A – мера той части области, которая благоприятствует событию A; D – мера всей области (длина, площадь, объем).

Пример 15. По району противника, который имеет форму эллипса с полуосями 7 км и 10 км, наносится удар ракетой. Определить вероятность того, что эпицентр взрыва окажется над танковым подразделением, площадь которого 4×3 км.

Решение. Испытание – наносится удар ракетой по району противника. Событие A – эпицентр взрыва оказался над танковым подразделением.

Вероятность события A в данном испытании вычисляется по формуле:

$$P(A)=\frac{D_A}{D},$$

где D_A = $S_{прямоугольника}$ – мера той части области (площадь танкового подразделения), которая благоприятствует событию A; D = $S_{эллипса}$ – мера всей области (площадь района противника) (рис. 1.6).

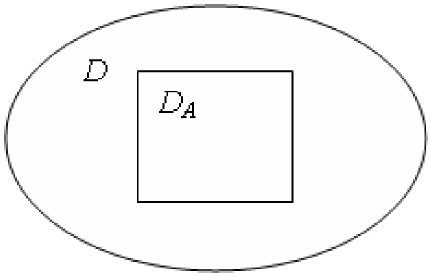


Рис. 1.6. Район противника

$$P(A)=\frac{S_{прямоугольника}}{S_{эллипса}}=\frac{3\cdot 4}{\pi\cdot 7\cdot 10}\approx 0{,}055.$$

Статистическое определение вероятности

Пусть испытание, с которым связано некоторое событие A, повторяется n раз, а само событие A осуществляется при этом k раз. Отношение $\frac{k}{n}$ принято называть *относительной частотой* события A.

Доказано, что при неограниченном увеличении числа испытаний ($n\to\infty$) относительная частота события A будет мало отличаться от вероятности появления этого события в отдельном испытании.

$$P(A)=\lim_{n\to\infty}\frac{k}{n}.$$

Введенная таким образом вероятность события называется *статистической*. Вероятность (в статистическом смысле) есть число, около которого колеблется наблюдаемая частота события. При достаточно большом n можно принять, что $P(A) \approx \frac{k}{n}$.

Пример 16. Стрелок сделал 40 выстрелов по мишени, при этом было зафиксировано 26 попаданий. Оценить вероятность попадания в мишень (событие A) этим стрелком.

Решение. Событие A – попадание в мишень. Общее число испытаний $n = 40$. Число испытаний $k = 26$, в которых наступило событие A. Отсюда следует:

$$P(A) = \frac{26}{40} = 0{,}65.$$

Рассмотренные выше *прямые* методы вычисления вероятностей не всегда эффективны. На практике часто применяют *косвенные* методы вычисления вероятностей, позволяющие находить вероятности одних событий по известным вероятностям других событий. В основе этих косвенных методов лежат основные теоремы и формулы теории вероятностей.

1.2. Основные теоремы и формулы теории вероятностей

Событие B называется **зависимым** от события A, если вероятность наступления события B зависит от того, произошло или нет событие A.

В теории вероятностей характеристикой связи событий A и B служит так называемая **условная вероятность** $P(B/A)$. Величина $P(B/A)$ может рассматриваться как вероятность наступления события B при условии, что событие A уже наступило.

Простейшими примерами связи событий A и B могут служить два крайних случая:

Наступление события A ведет к обязательному наступлению события B, т. е. $P(B/A) = 1$.

Пример 17. В сундуке лежат две монеты: золотая и серебряная. Наугад последовательно вынимают обе монеты. Какова вероятность, что вторая монета окажется серебряной, если первой вынули золотую монету?

Решение. Пусть событие *A* означает, что «вынули золотую монету», а событие *B* означает, что «вынули серебряную монету». По условию задачи заранее известно, что первой вынули золотую монету (событие *A* уже произошло) и в ящике осталась только серебряная монета. Поэтому $P(B/A)=\frac{m}{n}=\frac{1}{1}=1$, следовательно, событие *B* (вторая монета окажется серебряной) является достоверным, т. е. точно произойдет.

Наступление события *A* исключает возможность наступления события *B*, т. е. $P(B/A) = 0$.

Пример 18. В сундуке лежат две монеты: золотая и серебряная. Наугад последовательно вынимают обе монеты. Первой вынули золотую монету. Какова вероятность, что вторая монета окажется тоже золотой?

Решение. Пусть событие *A* означает, что «вынули золотую монету» и событие *B* означает, что «вынули серебряную монету». По условию задачи заранее известно, что первой вынули золотую монету (событие *A* уже произошло) и в сундуке осталась серебряная монета.

Поэтому:

$$P(B/A) = \frac{m}{n} = \frac{0}{1} = 0.$$

Следовательно, событие *B* (вторая монета окажется тоже золотой) является невозможным, т. е. точно не произойдет.

Теорема умножения вероятностей. Вероятность совместного наступления двух событий равна произведению вероятности одного из них на условную вероятность другого, вычисленную в предположении, что первое событие уже наступило:

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

Пример 19. Из 30 экзаменационных задач студент умеет решать только 25. На экзамене студент наугад последовательно выбирает две задачи. Какова вероятность того, что первую задачу он не решит, а со второй справится?

Решение. Пусть событие A означает, что «студент не решит первую задачу», т. е. ему попадется одна из пяти «плохих» задач. Тогда, исходя из классического определения вероятности, имеем:

$$P(A) = \frac{m}{n} = \frac{5}{30} = \frac{1}{6}.$$

Событие B означает, что «студент решит вторую задачу». Поскольку после наступления события A, одна из «плохих» задач уже извлечена, то остается всего 29 задач, из которых 25 студент умеет решать. Тогда, исходя из классического определения вероятности, имеем:

$$P(B/A) = \frac{m}{n} = \frac{25}{29}.$$

Искомую находим по теореме умножения вероятностей:

$$P(A \cdot B) = P(A) \cdot P(B/A) = \frac{1}{6} \cdot \frac{25}{29} \approx 0,14.$$

Теорема умножения вероятностей может быть обобщена на любое число множителей, например, для трех событий она имеет вид:

$$P(A \cdot B \cdot C) = P(A) \cdot P(B/A) \cdot P(C/AB).$$

Пример 20. Из шести имеющихся ключей к двери подходит только два. Случайным образом (без возвращения) выбирают ключ и пробуют открыть дверь. Ключи выбираются до тех пор, пока не появится нужный ключ. Какова вероятность, что к двери подойдет только четвертый ключ?

Решение. Пусть событие A означает, что «первый ключ не подходит». Так как всего ключей шесть, четыре из которых не подходят, то, исходя из классического определения вероятности, имеем:

$$P(A) = \frac{m}{n} = \frac{4}{6}.$$

Событие B означает, что «второй ключ не подходит». Так как после наступления события A останется всего пять ключей, из которых три не подходят, то, исходя из классического определения вероятности, имеем:

$$P(B/A) = \frac{m}{n} = \frac{3}{5}.$$

Событие C означает, что «третий ключ не подходит». Учитывая, что события A и B уже наступили и исходя из классического определения вероятности, имеем:

$$P(C/AB) = \frac{m}{n} = \frac{2}{4}.$$

Событие D означает, что «четвертый ключ подходит». Учитывая, что события A и B и C уже наступили и исходя из классического определения вероятности, имеем:

$$P(D/ABC) = \frac{m}{n} = \frac{2}{3}.$$

Событие «первый ключ не подходит, и второй ключ не подходит, и третий ключ не подходит, и четвертый ключ подходит» означает совместное наступление событий A, B, C, D, поэтому искомую вероятность находим по теореме умножения вероятностей:

$$P(A{\cdot}B{\cdot}C{\cdot}D)=P(A){\cdot}P(B/A){\cdot}P(C/AB){\cdot}P(D/ABC) = \frac{4}{6}\cdot\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{2}{3} \approx 0{,}13.$$

Событие B называется *независимым* от события A, если вероятность наступления события B не зависит от того, произошло или нет событие A.

Независимость событий означает, что наступление события A не изменяет вероятности появления события B, т. е. условная вероятность равна безусловной: $P(B/A) = P(B)$, и **теорема умножения вероятностей** имеет упрощенный вид:

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

Для независимых событий теорема умножения вероятностей может быть также обобщена на любое число множителей, например, для трех событий она имеет вид:

$$P(A{\cdot}B{\cdot}C) = P(A){\cdot}P(B){\cdot}P(C).$$

Пример 21. Абонент забыл последние три цифры телефонного номера, но помнит, что все они нечетные. Какова вероятность того, что абонент дозвонится с первого раза?

Решение. Пусть события A, B и C заключаются в том, что соответственно первая, вторая и третья из забытых цифр будут набраны верно.

Исходя из того, что события A, B, C независимы, а также учитывая, что из пяти нечетных цифр подходит в каждом случае только одна (классическое определение), вычислим искомую вероятность:

$$P(A \cdot B \cdot C) = P(A) \cdot P(B) \cdot P(C) = \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{1}{5} = 0{,}008.$$

Пример 22. Колесо фортуны разделено на 3 равных сектора, пронумерованных цифрами 1, 2, 3. Колесо крутят два раза. Событие A означает, что «в первый раз выпала цифра 3». Событие B означает, что «во второй раз выпала цифра не больше, чем в первый». Требуется найти $P(A)$, $P(B)$, $P(A \cdot B)$, $P(A/B)$ и определить, являются ли независимыми события A и B.

Решение. Для нахождения указанных вероятностей воспользуемся формулой классического определения вероятности. Опыт заключается в том, что «колесо крутят два раза» {все возможные исходы опыта 11, 12, 13, 21, 22, 23, 31, 32, 33}.

1. Событие A означает, что «в первый раз выпала цифра 3».
Исходы, благоприятные событию A: {31, 32, 33};

$$P(A) = \frac{m}{n} = \frac{3}{9} = \frac{1}{3}.$$

2. Событие B означает, что «во второй раз выпала цифра не больше, чем в первый».
Исходы, благоприятные событию B: {11, 21, 22, 31, 32, 33};

$$P(B) = \frac{m}{n} = \frac{6}{9} = \frac{2}{3}.$$

3. Событие $A \cdot B$ означает совместное наступление событий A и B. Другими словами, «в первый раз выпала цифра 3» и «во второй раз выпала цифра не больше чем в первый».
Исходы, благоприятные событию $A \cdot B$: {31, 32, 33};

$$P(A \cdot B) = \frac{m}{n} = \frac{3}{9} = \frac{1}{3}.$$

4. Событие A/B означает наступление события A при условии, что событие B уже наступило. В этом случае событие B становится достоверным. В рамках классического определения вероятности условную вероятность можно вычислять как отношение числа исходов, благоприятных появлению события A·B (совместному появлению событий A и B), к числу исходов, благоприятных появлению события B (отдельному появлению события B):

$$P(A/B) = \frac{P(A\cdot B)}{P(B)} = \frac{1/3}{2/3} = \frac{3}{6} = \frac{1}{2}.$$

5. События A и B не являются независимыми, так как $P(A\cdot B)=\frac{1}{3}$, $P(A)\cdot P(B)=\frac{1}{3}\cdot\frac{2}{3}=\frac{2}{9}$, следовательно $P(A\cdot B)\neq P(A)\cdot P(B)$.

Теорема сложения вероятностей несовместных событий. Вероятность суммы двух несовместных событий A и B равна сумме их вероятностей:

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

Пример 23. Из колоды в 36 карт наугад вынимают одну карту. Какова вероятность, что это будет семерка или король?

Решение. Пусть событие A означает, что «вынутая карта семерка», а событие B, что «вынутая карта король». Используя классическое определение, находим вероятности этих событий:

$P(A)=\frac{m}{n}=\frac{4}{36}=\frac{1}{9}$, так как всех карт 36, а семерок из них 4;

$P(B)=\frac{m}{n}=\frac{4}{36}=\frac{1}{9}$, так как всех карт 36, а королей из них 4.

Очевидно, что события A и B несовместны. Тогда, применяя теорему сложения вероятностей несовместных событий, получаем:

$$P(A+B)=P(A)+P(B) = \frac{1}{9}+\frac{1}{9} = \frac{2}{9} \approx 0{,}22.$$

Теорема сложения вероятностей несовместных событий может быть обобщена на любое число слагаемых:

$P(A_1 + A_2 + \ldots + A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$, *где* $A_1, A_2, \ldots, A_n$ – попарно несовместны.

Пример 24. На заочном отделении института обучаются 36 человек. Из них 8 человек на юридическом факультете, 12 человек на экономическом факультете, 10 человек на факультете подготовки психологов и 6 человек на лингвистическом факультете. Определить вероятность того, что наугад выбранный студент учится на юридическом, экономическом или лингвистическом факультете.

Решение. Пусть событие A означает, что «студент учится на юридическом факультете», событие B, что «студент учится на экономическом факультете», а событие C, что «студент учится на аграрном факультете». Используя классическое определение, находим вероятности этих событий:

$P(A)=\frac{m}{n}=\frac{8}{36}$, так как всех студентов 36, а юристов из них 8;

$P(B)=\frac{m}{n}=\frac{12}{36}$, так как всех студентов 36, а экономистов из них 12;

$P(C)=\frac{m}{n}=\frac{6}{36}$, так как всех студентов 36, а лингвистов из них 6.

Очевидно, что события A, B и C попарно несовместны. Тогда, применяя теорему сложения вероятностей несовместных событий, получаем:

$$P(A+B+C) = P(A) + P(B) + P(C) = \frac{8}{36}+\frac{12}{36}+\frac{6}{36}+\frac{26}{36}\approx 0{,}72.$$

В частности, если события $A_1, A_2, \ldots, A_n$ образуют *полную группу* попарно несовместных событий, то:

$$P(A_1, A_2, \ldots A_n) = P(A_1) + P(A_2) + \ldots + P(A_n) = 1.$$

Очевидно, что противоположные события A и $\overline{A}$ несовместны и тоже образуют полную группу, поэтому $P(A+\overline{A}) = P(A) + P(\overline{A}) = 1$, откуда получаем формулу для вычисления вероятности *противоположного* события:

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

Пример 25. Вероятность того, что ни один из трех экзаменов не будет сдан, равна 0,1. Найдите вероятность сдачи хотя бы одного из трех экзаменов.

Решение. Пусть событие A означает, что «ни один из трех экзаменов не будет сдан», тогда противоположное событие $\overline{A}$ означает, что «будет сдан хотя бы один из трех экзаменов». Используя формулу для вычисления вероятности противоположного события, находим:

$$P(\overline{A}) = 1 - P(A) = 1 - 0,1 = 0,9.$$

Теорема сложения вероятностей совместных событий. Вероятность суммы двух совместных событий A и B равна сумме вероятностей этих событий без вероятности их произведения:

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

Пример 26. Электрическая цепь состоит из двух последовательно соединенных элементов. Вероятность отказа первого элемента равна 0,1, второго – 0,4. Найти вероятность того, что за год откажет хотя бы один элемент.

Решение. Пусть событие A означает, что «откажет первый элемент», а событие B означает, что «откажет второй элемент». Событие «откажет хотя бы один элемент» означает, что или откажет первый элемент, или откажет второй элемент, или эти два события наступят одновременно (первый элемент и второй элемент откажут вместе). Следовательно, требуется определить сумму двух совместных событий A и B:

$$P(A + B) = P(A) + P(B) - P(A \cdot B) = 0,1 + 0,4 - 0,1 \cdot 0,4 = 0,46.$$

Теорема сложения вероятностей совместных событий может быть обобщена на любое число слагаемых, например для трех событий:

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

При увеличении числа слагаемых формула вероятности суммы совместных событий все более усложняется. В этом случае целесообразен переход к противоположному событию. В общем случае:

$$P(A_1 + A_2 + \ldots + A_n) = 1 - P(\overline{A}_2) \cdot P(\overline{A}_1) \cdot \ldots \cdot P(\overline{A}_n).$$

Пример 27. В банке установлены три системы защиты от ограбления. Вероятность, что при ограблении сработает первая, вторая и третья системы, равна 0,9, 0,8 и 0,7 соответственно.

Найти вероятность того, что при ограблении сработает хотя бы одна система защиты.

Решение. Пусть событие A означает, что «при ограблении сработает хотя бы одна система защиты».

Тогда противоположное событие $\overline{A}$ означает, что «ни одна система защиты не сработает», т. е. «не сработает первая система защиты, и не сработает вторая система защиты, и не сработает третья система защиты». Пусть события B, C, D означают, что сработает первая, вторая, третья системы защиты соответственно. Тогда противоположные события $\overline{B}$, $\overline{C}$, $\overline{D}$, означают, что не сработает первая, вторая, третья системы защиты соответственно. Вероятности событий B, C, D известны из условия задачи:

$$P(B) = 0{,}9,\ P(C) = 0{,}8,\ P(D) = 0{,}7.$$

Используя формулу для вычисления вероятности противоположного события, находим:

$$P(\overline{B}) = 1 - 0{,}9 = 0{,}1,\ P(\overline{C}) = 1 - 0{,}8 = 0{,}2,\ P(\overline{D}) = 1 - 0{,}7 = 0{,}3.$$

Поскольку события независимы, то $P(\overline{A}) = P(\overline{B})\cdot P(\overline{C})\cdot P(\overline{D})$.

Поэтому искомая вероятность (вероятность события A) будет равна:

$$P(A) = P(B + C + D) = 1 - P(\overline{A}) = 1 - P(\overline{B})\cdot P(\overline{C})\cdot P(\overline{D}) =$$
$$= 1 - 0{,}1\cdot 0{,}2\cdot 0{,}3 = 0{,}994.$$

При решении задачи теоремы сложения и умножения вероятностей часто применяются вместе.

Пример 28. Три студента сдают экзамен. Вероятность сдать экзамен на «отлично» для первого, второго и третьего студентов равна 0,7, 0,6 и 0,2 соответственно. Какова вероятность того, что экзамен будет сдан на «отлично»: 1) только одним студентом (событие B); 2) тремя студентами (событие C); 3) ни одним из студентов (событие D); 4) только двумя студентами (событие E); 5) только третьим студентом (событие F)?

Решение. Введем следующие обозначения:

A_1 – 1-й студент сдаст экзамен на отлично, $\overline{A}_1$ – 1-й студент не сдаст экзамен на отлично, $P(A_1) = 0{,}7$, $P(\overline{A}_1) = 1 - 0{,}7 = 0{,}3$;

A_2 – 2-й студент сдаст экзамен на отлично, $\overline{A}_2$ – 2-й студент не сдаст экзамен на отлично, $P(A_2) = 0{,}6$, $P(\overline{A}_1) = 1 – 0{,}6 = 0{,}4$;

A_3 – 3-й студент сдаст экзамен на отлично, $\overline{A}_3$ – 3-й студент не сдаст экзамен на отлично, $P(A_3) = 0{,}2$, $P(\overline{A}_3) = 1 – 0{,}2 = 0{,}8$.

События A_1, A_2, A_3 – совместны и независимы. Применяя теоремы сложения и умножения, находим вероятности событий B, C, D, E, F:

1) событие B – экзамен будет сдан на «отлично» только одним студентом (рис. 1.7).

$$P(B) = P(A_1)\cdot P(\overline{A}_2)\cdot P(\overline{A}_3)+P(\overline{A}_1)\cdot P(A_2)\cdot P(\overline{A}_3)+P(\overline{A}_1)\cdot P(\overline{A}_2)\cdot P(A_3);$$

$$P(B) = 0{,}7\cdot 0{,}4\cdot 0{,}8 + 0{,}3\cdot 0{,}6\cdot 0{,}8 + 0{,}3\cdot 0{,}4\cdot 0{,}2 = 0{,}392;$$

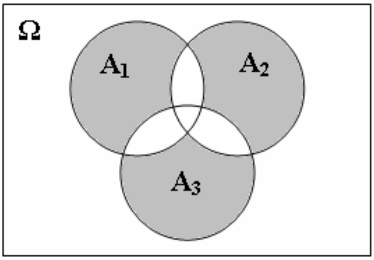


Рис. 1.7. Экзамен сдан на «отлично» только одним студентом

2) событие C – экзамен будет сдан на «отлично» тремя студентами (рис. 1.8).

$$P(C) = P(A_1)\cdot P(A_2)\cdot P(A_3) = 0{,}7\cdot 0{,}6\cdot 0{,}2 = 0{,}084;$$

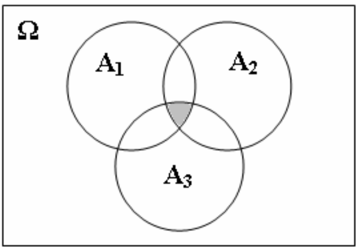


Рис. 1.8. Экзамен сдан на «отлично» тремя студентами

3) событие D – ни один из студентов не сдаст экзамен на «отлично» (рис. 1.9).

$$P(D) = P(\overline{A}_1)\cdot P(\overline{A}_2)\cdot P(\overline{A}_3) = 0{,}3\cdot 0{,}4\cdot 0{,}8 = 0{,}096.$$

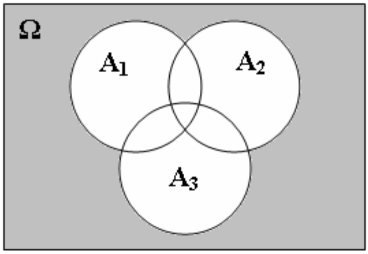


Рис. 1.9. Ни один из студентов не сдал экзамен на «отлично»

4) событие E – экзамен будет сдан на «отлично» только двумя студентами (рис. 1.10):

$$P(E) = P(A_1)\cdot P(A_2)\cdot P(\overline{A}_3) + P(\overline{A}_1)\cdot P(A_2)\cdot P(A_3) + P(A_1)\cdot P(\overline{A}_2)\cdot P(A_3).$$

$$P(E) = 0{,}7\cdot 0{,}6\cdot 0{,}8 + 0{,}3\cdot 0{,}6\cdot 0{,}2 + 0{,}7\cdot 0{,}4\cdot 0{,}2 = 0{,}428.$$

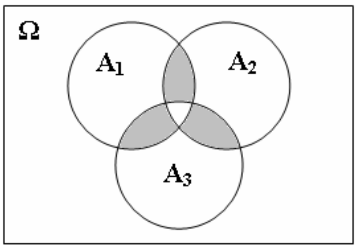


Рис. 1.10. Экзамен сдан на «отлично» только двумя студентами

5) событие F – экзамен будет сдан на «отлично» только третьим студентом (рис. 1.11):

$$P(E)=P(\overline{A}_1)\cdot P(\overline{A}_2)\cdot P(A_3) = 0{,}3\cdot 0{,}4\cdot 0{,}2 = 0{,}024.$$

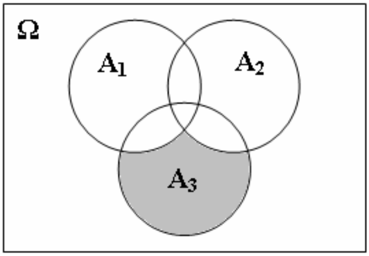


Рис. 1.11. Экзамен сдан на «отлично» только третьим студентом

Формула полной вероятности. Формула Байеса

Пусть некоторое событие A может произойти при наступлении одного и только одного из событий H_1, H_2, ..., H_n, называемых гипотезами и образующих полную группу несовместных событий.

Тогда вероятность события A определяется по **формуле полной вероятности**:

$$P(A) = P(H_1)\cdot P(A/H_1) + P(H_2)\cdot P(A/H_2) + \ldots + P(H_2)\cdot P(A/H_n)$$

или:

$$P(A)=\sum_{i=1}^{n} P(H_i)\cdot P(A/H_i),$$

где $P(H_i)$ – вероятности гипотез, $P(A/H_i)$ – условные вероятности наступления события A с каждой из гипотез, а $i = 1, 2 \ldots n$ – индекс, обозначающий номер соответствующей гипотезы[1].

[1] Количество слагаемых в формуле полной вероятности всегда равно количеству гипотез, выдвинутых по условию задачи.

Пример 29. Из пункта *O* ведут четыре дороги, проходящие через пункты H_1, H_2, H_3, H_4 соответственно (схема представлена на рис. 1.12). Какова вероятность, что путник, вышедший из пункта *O*, попадет в пункт *A*, если он пойдет по наугад выбранной дороге?

Решение. Событие *A* – путник доберется из пункта *O* в пункт *A* по наугад выбранной дороге.

Гипотезы H_1, H_2, H_3, H_4 – выбор путником одной из четырех дорог, проходящих через пункты H_1, H_2, H_3, H_4 соответственно. Очевидно, что данные гипотезы равновероятны (путник с равной вероятностью может выбрать одну из четырех дорог) и образуют полную группу несовместных событий:

$$P(H_1) = P(H_2) = P(H_3) = P(H_4) = \frac{1}{4} = 0{,}25.$$

Сделаем проверку (сумма вероятностей несовместных событий, образующих полную группу, есть событие достоверное):

$$P(H_1)+P(H_2)+P(H_3)+P(H_4) = 0{,}25+0{,}25+0{,}25+0{,}25 = 1.$$

Согласно схеме дорог находим условные вероятности попадания путника в пункт *A*:

$P(A/H_1) = 0$ – вероятность попасть в пункт *A*, если путник выбирает направление H_1;

$P(A/H_2)=\frac{1}{2} = 0{,}5$ – вероятность попасть в пункт *A*, если путник выбирает направление H_2;

$P(A/H_3) = 1$ – вероятность попасть в пункт *A*, если путник выбирает направление H_3;

$P(A/H_4) = 0{,}5$ – вероятность попасть в пункт *A*, если путник выбирает направление H_4.

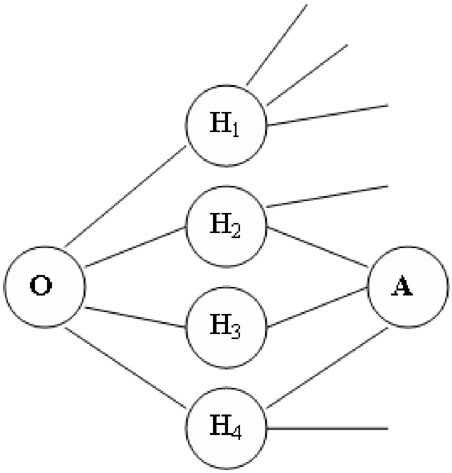


Рис. 1.12. Схема дорог

Найденные вероятности удобно записать в таблицу:

i	1	2	3	4
$P(H_i)$	0,25	0,25	0,25	0,25
$P(A/H_i)$	0	0,5	1	0,5

Применяя формулу полной вероятности, получим:

$$P(A) = P(H_1)\cdot P(A/H_1) + P(H_2)\cdot P(A/H_2) + P(H_3)\cdot P(A/H_3) + $$
$$+ P(H_4)\cdot P(A/H_4)$$

$$P(A) = 0{,}25\cdot 0 + 0{,}25\cdot 0{,}5 + 0{,}25\cdot 1 + 0{,}25\cdot 0{,}5 = 0{,}5.$$

Рассмотрим ситуацию, противоположную предыдущей (в некотором смысле). Пусть событие A уже произошло. Определим вероятность того, что событие A произошло именно с гипотезой H_i. Условная вероятность гипотезы H_i при условии того, что событие A уже произошло, определяется по *формуле Байеса*:

$$P(H_i/A) = \frac{P(H_i)\cdot P(A/H_i)}{\sum_{i=1}^{n} P(H_i)\cdot P(A/H_i)}.$$

Следует заметить, что знаменателем формулы Байеса является формула полной вероятности:

$$\sum_{i=1}^{n} P(H_i)\cdot P(A/H_i).$$

Формула Байеса часто интерпретируется как формула, позволяющая по *априорным* (известным до проведения опыта) вероятностям гипотез $P(H_i)$ и условным вероятностям $P(A/H_i)$ вычислить *апостериорные* (найденные после опыта, т. е. при условии, что событие A уже произошло) вероятности $P(H_i/A)$.

Пример 30. Пусть сохраняются условия предыдущей задачи (пример 29), но дополнительно известно, что событие A уже произошло (путник попал из пункта O в пункт A). Какова вероятность, что он выбрал дорогу, проходящую через пункт H_4 (была реализована гипотеза H_4)?

Используя формулу Байеса, находим:

$$P(H_4/A) = \frac{P(H_4)\cdot P(A/H_4)}{\sum_{i=1}^{4} P(H_i)\cdot P(A/H_i)} = \frac{0,25\cdot 0,5}{0,5} = 0,25.$$

Пример 31. Стрелковый клуб получил 12 винтовок, из которых 8 пристрелянных, а 4 нет. Вероятность попадания в цель из пристрелянной винтовки равна 0,8, а из не пристрелянной 0,5. Стрелок делает один выстрел из винтовки. Найти вероятность того, что: 1) стрелок попал в цель, выбрав винтовку наугад; 2) если стрелок попал в цель, выбрав винтовку наугад, то он стрелял из пристрелянной винтовки.

Решение. Событие A – стрелок поразил цель, выбрав винтовку наугад. Гипотеза H_2 – стрелок возьмет пристрелянную винтовку. Гипотеза H_2– стрелок возьмет не пристрелянную винтовку.

Очевидно, что данные гипотезы образуют полную группу несовместных событий. Применяя классическое определение вероятности, находим:

$P(H_1) = \frac{m}{n} = \frac{8}{12} = 0,67$ – вероятность того, что винтовка пристрелянная;

$P(H_2) = \frac{m}{n} = \frac{4}{12} = 0{,}33$ – вероятность того, что винтовка не пристрелянная.

Сделаем проверку (сумма вероятностей событий, образующих полную группу, есть событие достоверное):

$$P(H_1) + P(H_2) = \frac{8}{12} + \frac{4}{12} = 1.$$

Согласно условию задачи запишем условные вероятности попадания стрелком в цель:

$P(A/H_1) = 0{,}8$ – вероятность попасть в цель, если стрелок выбирает пристрелянную винтовку;

$P(A/H_2) = 0{,}5$ – вероятность попасть в цель, если стрелок выбирает не пристрелянную винтовку.

Применяя формулу полной вероятности, находим вероятность того, что стрелок попал в цель, выбрав винтовку наугад:

$$P(A) = P(H_1)\cdot P(A/H_1) + P(H_2)\cdot P(A/H_2) = 0{,}67\cdot 0{,}8 + 0{,}33\cdot 0{,}5 = 0{,}701.$$

Применяя формулу Байеса, находим вероятность того, что если стрелок попал в цель, выбрав винтовку наугад, то он стрелял из пристрелянной винтовки:

$$P(H_1/A) = \frac{P(H_1)\cdot P(A/H_1)}{\sum_{i=1}^{2} P(H_i)\cdot P(A/H_i)} = \frac{0{,}67\cdot 0{,}8}{0{,}701} = 0{,}76.$$

Формула Бернулли. Формула Пуассона

Рассмотрим одну из главных схем теории вероятностей, имеющую прикладное значение не только в разнообразных областях знания, но и в самой теории вероятностей как математической науке. Эта схема состоит в том, что рассматривается *последовательность взаимно независимых испытаний*[1]. Вероятность того или иного результата в каждом из этих испытаний не зависит от того, какие результаты наступили или наступят в остальных.

---

[1] В слово «испытание» вкладывается весьма разнообразный смысл. Так, если производят стрельбу по некоторой цели, то под испытанием подразумевают каждый отдельный выстрел. Если изучают результаты вакцинации от гриппа в детском учреждении, то под испытанием понимают обследование отдельного ребенка и т. д.

На практике часто приходится рассматривать последовательности испытаний с двумя исходами: стрелок поразит цель или промахнется; изделие окажется годным или дефектным; лотерейный билет будет выигрышным или не выигрышным; вакцина будет полезной или бесполезной и т. д.

Частный случай схемы повторных независимых испытаний, в котором каждое испытание может закончиться только одним из двух исходов, называют *схемой Бернулли*. Ее изучение ведет свое начало от известного швейцарского ученого Якова Бернулли, жившего в конце XVII в.

Пусть производится n одинаковых и независимых испытаний, причем в каждом испытании некоторое событие A может наступить («успех») с вероятностью $P(A) = p$ или не наступить («неудача») с вероятностью $P(\overline{A}) = 1 - p$. Тогда вероятность того, что в n испытаниях событие A появится k раз вычисляется по *формуле Бернулли*:

$$P_n(k) = C_n^k \cdot p^k \cdot (1-p)^{n-k}.$$

Пример 32. Прививка от гриппа дает положительный результат в 70 % случаев. Найти вероятность того, что в группе из десяти человек для двух она будет бесполезной.

Решение. По условию задачи событие A (прививка бесполезна) может наступить с вероятностью:

$$P(A) = p = 1 - \frac{70\ \%}{100\ \%} = 0{,}3.$$

Очевидно, имеет место схема Бернулли:

p	$1-p$	n	k
0,3	0,7	10	2

Поэтому получаем:

$$P_{10(2)} = C_{10}^2 \cdot (0{,}3)^2 \cdot (0{,}7)^{10-2} = 0{,}233474 \approx 0{,}23,$$

где $P_{10(2)}$ – вероятность того, что для двух человек из десяти прививка от гриппа окажется бесполезной.

Пример 33. Студент может правильно решить задачу по математике с вероятностью 0,6. Найти вероятность того, что из 10 предложенных задач он решит правильно:1) ровно три; 2) хотя бы одну; 3) не более двух.

Решение. Из условия задачи очевидно, что имеет место схема Бернулли, где:

1) $p = 0{,}6,\ 1–p = 0{,}4, n = 10,\ k = 3$, поэтому:

$$P_{10(3)} = C_{10}^{3} \cdot 0{,}6^{3} \cdot 0{,}4^{10-3} \approx 0{,}042.$$

2) $p = 0{,}6,\ 1 - p = 0{,}4,\ n = 10,\ k \geq 1.$ Событие «студент решит хотя бы одну задачу» означает, что студент решит одну задачу или более (2 или 3 или 4 или 5 или 6 или 7 или 8 или 9 или 10). Чтобы избежать громоздких вычислений, находим вероятность противоположного события («студент не решит ни одной задачи») и вычитаем ее из единицы:

$$P_{10(k\geq 1)} = 1 - P_{10(0)} = 1 - C_{10}^{0} \cdot (0{,}6)^{0} \cdot (0{,}4)^{10-0} = 1 - (0{,}4)^{10}.$$

3) $p = 0{,}6,\ 1 - p = 0{,}4,\ n = 10,\ k \leq 2.$ Событие «студент решит не более двух задач» означает, что студент решит две задачи, или решит одну задачу, или не решит ни одной задачи. Поэтому вероятность этого события (согласно теореме сложения вероятностей для попарно несовместных событий) равна:

$$P_{10(k\leq 2)} = P_{10(2)} + P_{10(1)} + P_{10(0)}.$$

Применение формулы Бернулли при большом числе испытаний n и малом значении вероятности p приводит к трудоемким вычислениям, поэтому в таком случае удобнее пользоваться приближенной *формулой Пуассона*[1]:

$$P_{n(k)} \approx \frac{\lambda^{k}}{k!} \cdot e^{-\lambda},\ где\ \lambda = n \cdot p,\ где\ e \approx 2{,}72.$$

Эта формула дает хорошее приближение при $p \leq 0{,}1$ и $\lambda = n \cdot p \leq 10$. События, для которых применима формула Пуассона, называют *редкими,* так как вероятность их наступления

[1] С. Пуассон (1781–1840) – французский математик.

достаточно мала. Число $\lambda = n{\cdot}p$ можно интерпретировать как среднее число появления рассматриваемого события в n испытаниях.

Пример 34. МВД России отправило 200 сотрудников в служебную командировку на Северный Кавказ. Вероятность получить ранение во время боевых действий равна 0,025. Найти вероятность того, что будет ранено ровно четыре человека.

Решение. Поскольку вероятность ранения p = 0,025 мала, количество отправленных сотрудников n = 200 велико, а $\lambda = n{\cdot}p = 5 \leq 10,$ можно воспользоваться формулой Пуассона:

$$P_{200(4)} \approx \frac{5^4}{4!} \cdot 2{,}72^{-5} \approx 0{,}18.$$

1.3. Дискретные случайные величины

Наряду со случайным событием одним из основных понятий теории вероятностей является понятие **случайной величины** – величины, численное значение которой может меняться в зависимости от результата испытания[1].

Случайной называют величину, которая в результате испытания примет одно и только одно из возможных значений, наперед неизвестное и зависящее от ряда причин, которые нельзя учесть заранее.

Примерами случайных величин могут служить: оценка на экзамене – целое положительное число (от 2 до 5); количество баллов за тест Айзенка – целое положительное число (не более 160)[2]; продолжительность безотказной работы компьютера – любое неотрицательное число и т. д. Обозначают случайные величины обычно строчными буквами греческого алфавита – ξ, η, ζ или про-

---

[1] Случайное событие – качественная характеристика испытания, случайная величина – количественная характеристика испытания. Например, сбой в работе полиграфа – это случайное событие, а время его бесперебойной работы – случайная величина.

[2] Ганс Айзенк – выдающийся немецко-британский психолог.

писными буквами латинского алфавита – X, Y, Z, …, а их возможные значения – x, y, z, …, которые снабжают (при необходимости) индексами.

В зависимости от возможных значений все случайные величины можно разделить на два класса: **дискретные** и **непрерывные**.

Дискретной называют случайную величину, которая принимает отдельные, изолированные друг от друга значения с определенными вероятностями. Число возможных значений дискретной случайной величины может быть конечным или бесконечным, но счетным.

Под **законом распределения вероятностей случайной величины** понимается всякое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями.

Знание закона распределения позволяет заранее установить, какое значение случайной величины будет появляться чаще, а какое реже и насколько. Закон распределения случайной величины может быть представлен рядом распределения (таблично):

x_i	x_1	x_2	…	x_n
p_i	p_1	p_2	…	p_n

Здесь x_i – возможные значения случайной величины X, а p_i – соответствующие этим значениям вероятности. Очевидно, что справедливо следующее равенство (т. е. то, что случайная величина X в результате испытания примет одно из значений x_1, x_2, ..., x_n , есть событие достоверное):

$$\sum_{i=1}^{n} p_i = p_1 + p_2 + \ldots + p_n = 1.$$

Закон распределения случайной величины может быть представлен графически – многоугольником (полигоном) распределения вероятностей[1]:

[1] Многоугольник распределения лежит не ниже оси OX, так как все $p_i \geq 0$.

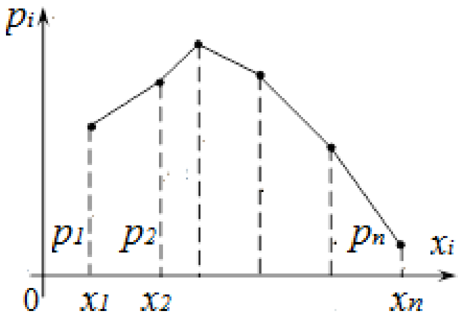


Рис. 1.13. Полигон

Закон распределения случайной величины может быть представлен аналитически. Если вероятности событий, в результате которых случайная величина примет одно из своих возможных значений, вычисляются по определенным формулам – формуле Бернулли, формуле Пуассона и др., то законы распределений имеют определенное название: Биномиальный, Пуассоновский и т. д. На практике бывает достаточно знание лишь основных суммарных характеристик случайной величины: математического ожидания, дисперсии, среднего квадратического отклонения.

Математическим ожиданием дискретной случайной величины называют сумму произведений всех ее возможных значений на соответствующие им вероятности[1]. Обычно обозначают: $M(X)$, μ, a.

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

Геометрический смысл: $M(X)$ является абсциссой центра тяжести (ЦТ) площади, ограниченной сверху полигоном распределения, а снизу осью OX (рис. 1.14).

[1] Математическое ожидание приближенно равно среднему арифметическому возможных значений случайной величины, причем тем точнее, чем больше число испытаний.

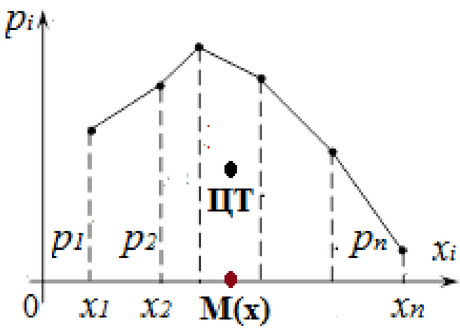


Рис. 1.14. Геометрический смысл M(X)

$M(X)$ обладает следующими свойствами:

$M(C) = C$; $M(CX) = C \cdot M(X)$, где C – *const*;

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

Характеристиками рассеивания возможных значений случайной величины относительно ее математического ожидания служат дисперсия и среднее квадратическое отклонение.

Дисперсией дискретной случайной величины называют математическое ожидание квадрата отклонения случайной величины от ее математического ожидания. Обозначают $D(X)$.

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

$D(X)$ обладает следующими свойствами:

$D(C) = 0$; $D(CX) = C^2 \cdot D(X)$, где C – *const*;

$D(X + Y + Z) = D(X) + D(Y) + D(Z)$,

где X; Y; Z – независимые случайные величины.

Поскольку размерность дисперсии – квадрат размерности случайной величины, то для оценки меры рассеивания случайной величины (самой величины отклонения) используют среднее квадратическое отклонение.

Среднее квадратическое отклонение имеет размерность случайной величины и равна квадратному корню из дисперсии. Обозначают σ(X).

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

Пример 35. Известны законы распределения числа выбитых очков двумя стрелками.

Надо определить, кто из них стреляет лучше[1].

x_i	8	9	10
p_i	0,3	0,5	0,2

y_i	8	9	10
p_i	0,5	0,1	0,4

Решение. Пусть случайная величина X – число очков, выбитых первым стрелком, а Y – число очков, выбитых вторым стрелком.

$$M(X) = 8\cdot 0{,}3 + 9\cdot 0{,}5 + 10\cdot 0{,}2 = 8{,}9;$$

$$M(Y) = 8\cdot 0{,}5 + 9\cdot 0{,}1 + 10\cdot 0{,}4 = 8{,}9.$$

Математические ожидания равны, поэтому для оценки стрельбы определим кучность стрельбы у каждого стрелка.

$$D(X) = 64\cdot 0{,}3 + 81\cdot 0{,}5 + 100\cdot 0{,}2 – 8{,}9^2= 0{,}49;$$

$$D(Y)= 64\cdot 0{,}5 + 81\cdot 0{,}1 + 100\cdot 0{,}4 – 8{,}9^2= 0{,}89.$$

$$\sigma(X) =\sqrt{0{,}49} = 0{,}7;\ \sigma(Y) =\sqrt{0{,}89} = 0{,}94.$$

Очевидно, что первый стрелок стреляет лучше, так как мера рассеивания его стрельбы меньше.

Пример 36. Абитуриент сдает два вступительных экзамена: по истории и математике. Составить закон распределения случай-

---

[1] Математика : учебное пособие. Ч. 2 / сост. А. В. Мамасуев, Н. М. Дубинина, С. А. Скворцова. М. : Академия экономической безопасности МВД России, 2009. С. 71.

ной величины *X*, числа полученных пятерок, если вероятность получения пятерки по истории – 0,8, а по математике – 0,6. Найти $M(X)$, $D(X)$, $\sigma(X)$. Построить многоугольник распределения вероятностей.

Решение. Пусть событие *A* означает, что «история сдана на 5», а событие *B* – «математика сдана на 5». Тогда $P(A) = 0{,}8$, а $P(B) = 0{,}6$.

Очевидно, что возможными значениями *X* (числа полученных пятерок) являются 0, 1, 2.

$$P(X = 0) = P(\overline{A})\cdot P(\overline{B}) = (1 - 0{,}8)\cdot(1 - 0{,}6) = 0{,}2\cdot 0{,}4 = 0{,}08;$$

$$P(X=1) = P(\overline{A})\cdot P(B) + P(A)\cdot P(\overline{B}) = 0{,}2\cdot 0{,}6 + 0{,}8\cdot 0{,}4 = 0{,}44;$$

$$P(X = 2) = P(A)\cdot P(B) = 0{,}8\cdot 0{,}6 = 0{,}48.$$

Контроль:

$$\sum_{i=1}^{3} P_i = 0{,}08 + 0{,}44 + 0{,}48 = 1.$$

Представим закон распределения *X*:

x_i	0	1	2
p_i	0,08	0,44	0,48

Найдем числовые характеристики данной случайной величины:

$$M(X) = 0\cdot 0{,}08 + 1\cdot 0{,}44 + 2\cdot 0{,}48 = 1{,}4;$$

$$D(X) = 0^2\cdot 0{,}08 + 1^2\cdot 0{,}44 + 2^2\cdot 0{,}48 - 1{,}4^2 = 2{,}36 - 1{,}96 = 0{,}4;$$

$$\sigma(X) = \sqrt{D(X)} = \sqrt{0{,}4} = 0{,}6.$$

Закон распределения *X* в графическом виде (рис. 1.15):

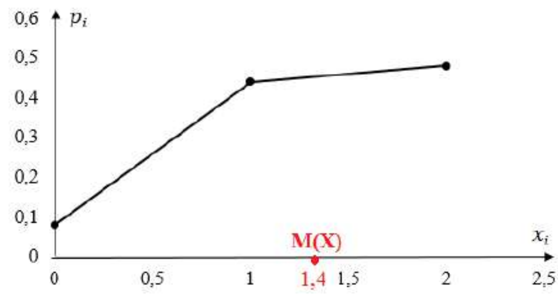


Рис. 1.15. Многоугольник распределения вероятностей

1.4. Непрерывные случайные величины

Непрерывной называют случайную величину, возможные значения которой целиком заполняют некоторый интервал.

При описании непрерывной случайной величины принципиально невозможно выписать все ее значения, принадлежащие даже самому узкому интервалу[1]. Поэтому, в отличие от дискретной случайной величины, задание значений непрерывной случайной величины и соответствующих им вероятностей недопустимо. В этой связи «универсальной» формой представления закона распределения является *функция распределения* $F(x)$, которая полностью характеризует как дискретную, так и непрерывную случайную величину.

Функцией распределения $F(x)$ называется вероятность того, что случайная величина X примет значение меньше наперед заданного действительного числа x:

$$F(x) = P(X < x).$$

Геометрически функцию распределения можно интерпретировать как вероятность того, что случайная величина X примет значение, которое изображается на числовой оси точкой, лежащей левее точки x.

$F(x)$ обладает следующими свойствами:

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

2) $F(x)$ – неубывающая функция, так как $F(x_2) \ge F(x_1)$, если $x_2 \ge x_1$;

3) если возможные значения X принадлежат интервалу $(a; b)$, то:

$$F(x) = 0 \text{ при } x \le a;$$

$$F(x) = 1 \text{ при } x \ge b;$$

---

[1] Эти значения образуют несчетное множество – «континуум» (от лат. continuum – «непрерывное»). Например, можно говорить, что снаряд пролетит 4523,1208 м только с вероятностью 0, так как практически невозможно установить, произошло это событие или нет. В связи с тем, что измерение производится с ограниченной точностью, в качестве измеренной величины можно лишь указать границы некоторого интервала, содержащего искомое значение.

4) вероятность того, что случайная величина X примет значение, заключенное в интервале $(a;\ b)$, равна приращению функции распределения на этом интервале.

$$P(a \le X \le b) = F(b) - F(a).$$

Так как вероятность того, что непрерывная случайная величина X примет одно определенное значение, равна нулю, то справедливо утверждение[1]:

$$P(a \le X \le b) = P(a < X < b) = P(a < X \le b) = P(a \le X < b).$$

Функцию $F(x)$ называют **интегральной** функцией распределения или интегральным законом распределения.

Пример 37. Случайная величина X задана интегральной функцией распределения:

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

Вычислить $P(0{,}5 < X < 0{,}6)$.

Решение.

$$P(0{,}5<X<0{,}6) = F(0{,}6) - F(0{,}5) = \frac{0{,}6}{2} - \frac{0{,}5}{2} = 0{,}3 - 0{,}25 = 0{,}05.$$

Можно представить графиком (рис. 1.16):

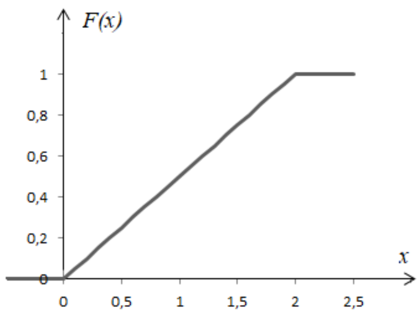


Рис. 1.16. Интегральная функция распределения

[1] Гмурман В. Е. Теория вероятностей и математическая статистика : учебное пособие для вузов. изд. 6-е. М. : Высшая школа, 1998. С. 113.

Однако по $F(x)$ трудно судить о характере распределения непрерывной случайной величины в окрестности той или иной точки числовой оси. Поэтому более наглядное представление дает первая производная от интегральной функции распределения: $f(x) = F'(x)$.

Функцию $f(x)$ называют **плотностью распределения вероятностей**. Плотность распределения вероятностей (далее – плотность распределения) $f(x)$ показывает, как часто появляется случайная величина X в окрестности точки x (некоторого действительного числа) при повторении опытов.

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

$f(x)$ обладает следующими свойствами:

1) $f(x) \geq 0$, т. е. функция плотности распределения неотрицательна;

2) кривая распределения всегда расположена не ниже оси OX;

3) вероятность того, что случайная величина примет значение, принадлежащее интервалу $(-\infty; +\infty)$, равна единице[1]:

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

Геометрический смысл: вся площадь, ограниченная сверху кривой распределения, а снизу осью OX, равна 1.

Зная $f(x)$, можно вычислить вероятность попадания случайной величины на некоторый интервал $(a; b)$[2]:

$$\mathrm{P}(a < X < b) = \int_{a}^{b} f(x)dx = F(b) - F(a).$$

Геометрический смысл: вероятность того, что непрерывная случайная величина примет значение, принадлежащее интервалу

---

[1] Очевидно, что такое событие достоверно.

[2] Интервал $(a; b)$ составляет некоторую долю от общей вероятности $P(-\infty;+\infty) = 1$.

$(a; b)$, равна площади криволинейной трапеции, ограниченной сверху кривой распределения (графиком $f(x)$), снизу осью OX, слева прямой $x = a$, справа прямой $x = b$ (рис. 1.17).

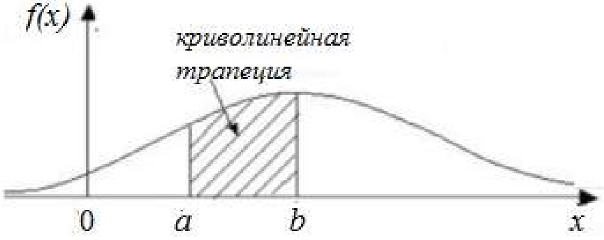


Рис. 1.17. Геометрический смысл $\int_a^b f(x)dx$

Функцию $f(x)$ называют **дифференциальной** функцией распределения. Плотность распределения $f(x)$ является формой задания закона распределения непрерывной случайной величины[1].

Математическое ожидание непрерывной случайной величины X, возможные значения которой принадлежат интервалу $(-\infty; +\infty)$, т. е. всей оси OX, определяется равенством:

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

В частности, если все возможные значения X принадлежат интервалу $(a; b)$:

$$\mathrm{M(X)} = \int_{a}^{b} x \cdot f(x)dx.$$

Дисперсия непрерывной случайной величины X, возможные значения которой принадлежат интервалу $(-\infty; +\infty)$, т. е. всей оси OX, определяется равенством:

[1] Очевидно, что $f(x)$ неприемлема для задания дискретной случайной величины.

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

В частности, если все возможные значения X принадлежат интервалу $(a;\ b)$:

$$D(X) = \int_{a}^{b} (x - M(X))^2 \cdot f(x)dx = \int_{a}^{b} x^2 \cdot f(x)dx - M^2(X).$$

Следует заметить, что свойства $M(X)$ и $D(X)$ дискретных случайных величин сохраняются и для непрерывных величин.

Среднее квадратическое отклонение непрерывной случайной величины определяется так же, как и для дискретной величины[1]: $\sigma(X) = \sqrt{D(X)}$.

Пример 38. Непрерывная случайная величина X задана интегральной функцией распределения:

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

Найти плотность распределения $f(x)$, вычислить вероятность того, что в результате испытания X примет значение, принадлежащее интервалу (0,5; 1), определить $M(X)$, $D(X)$, $\sigma(X)$.

Решение. Учитывая, что $f(x)$ является первой производной от интегральной функции распределения, находим плотность распределения:

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

Очевидно, что вероятность того, что X примет значение, принадлежащее интервалу (0,5; 1), можно вычислить двумя способами.

I способ:

$$P(0{,}5 < X < 1) = F(1) - F(0{,}5) = 1^2 - 0{,}5^2 = 1 - 0{,}25 = 0{,}75.$$

---

[1] Гмурман В. Е. Теория вероятностей и математическая статистика. С. 126.

II способ:

$$P(0,5 < X < 1) = \int_{0,5}^{1} 2xdx = \frac{2\cdot 1^2}{2} - \frac{2\cdot 0,5^2}{2} = 1 - 0,25 =$$

$$= 0,75 \cdot M(X) = \int_{0}^{1} x\cdot 2xdx = \frac{2\cdot 1^3}{3} - \frac{2\cdot 0^3}{3} = \frac{2}{3}.$$

$$D(x) = \int_{0}^{1} x^2\cdot 2xdx - M^2(X) = \left(\frac{2\cdot 1^4}{4} - \frac{2\cdot 0^4}{4}\right) - \frac{2^2}{3^2} = \frac{1}{18} = 0,06.$$

$$\sigma(X) = \sqrt{0,06} = 0,24.$$

Можно представить графиком (рис. 1.18):

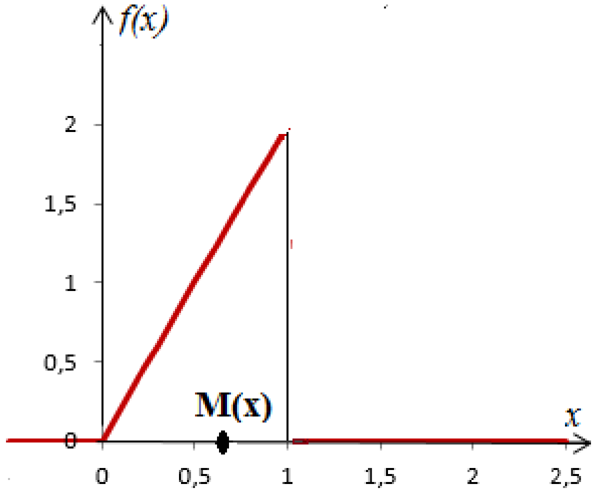


Рис. 1.18. Дифференциальная функция распределения

При решении задач, которые выдвигает практика, наиболее часто приходится сталкиваться с равномерным, показательным и нормальным распределениями случайной величины. Считают, что случайная величина подчиняется *нормальному* закону распределения, если представляется возможным рассматривать эту величину как сумму достаточно большого числа случайных величин[1].

Суммируемые величины должны быть независимы (или слабо зависимы) и каждая из этих величин должна играть в общей сумме примерно одинаковую (ничтожно малую роль). Например, если ведется стрельба из ПМ, то величина отклонения точки попадания от точки прицеливания – есть случайная величина, подчиняющаяся нормальному закону. Очевидно, что на данную случайную величину влияет целый ряд случайных факторов: зрение стрелка, атмосферное давление, пристрелянность оружия и др.

Непрерывная случайная величина распределена нормально, если ее плотность распределения имеет вид:

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

Нормальное распределение определяется двумя параметрами: $M(X)$ и σ.

Кривая нормального распределения характеризуется рядом особенностей: 1) кривая имеет колоколообразную форму; 2) если через наивысшую точку кривой провести прямую, параллельную оси OY, то кривая будет симметрична относительно этой прямой; 3) отклонения случайной величины от математического ожидания, имеющие одинаковые абсолютные значения, но разные знаки, встречаются одинаково часто; 4) чем больше значение отклонения, тем реже оно встречается.

Для любых комбинаций $M(X)$ и σ можно построить свой график нормального распределения. Однако существует «эталонное» нормальное распределение (универсальное, не зависящее от масштаба и единиц измерения), которое называется стандартным.

---

[1] При этом суммируемые случайные величины могут подчиняться каким угодно законам распределения.

Стандартное нормальное распределение имеет следующие параметры: $M(X) = 0$, $\sigma = 1$. В этом случае площадь под кривой распределения может быть интерпретирована как вероятность (рис. 1.19).

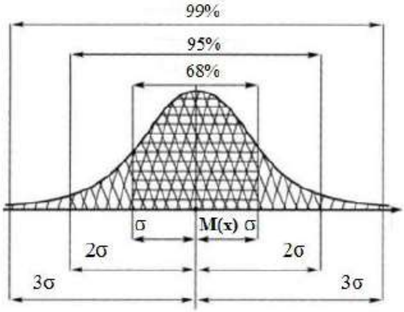


Рис. 1.19. Стандартное нормальное распределение

Для того чтобы воспользоваться стандартным нормальным распределением, масштаб реальных данных «подгоняют» под эталон с помощью процедуры «нормализация», которая имеет вид:

$$Z = \frac{x - M(X)}{\sigma} .$$

Данная процедура позволяет внести коррективы в интерпретацию случайной величины. Теперь это не просто наблюдаемая величина, а отклонение от $M(X)$, выраженное в σ.

Правило трех сигм. Если случайная величина распределена по нормальному закону, то ее отклонение от математического ожидания практически (99 %) не превышает $\pm 3\sigma$.

Пример 39. Записать закон распределения нормально распределенной случайной величины, если $M(X) = 10$, $\sigma = 2$. Вычислить $P(12<X<14)$.

Решение. Подставляем $M(X) = 10$, $\sigma = 2$ в закон нормального распределения, имеем:

$$\mathrm{f}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x - M(X))^2}{2\sigma^2}} = \frac{1}{2\sqrt{2\pi}} e^{-\frac{(x-10)^2}{8}} .$$

Зная $f(x)$, можно вычислить вероятность попадания случайной величины на некоторый интервал $(a;b)$:

$$\mathrm{P}(a < X < b) \ = \int\limits_{a}^{b} f(x)dx = \int\limits_{12}^{14} \frac{1}{2\sqrt{2\pi}} e^{-\frac{(x-10)^2}{8}} dx.$$

Интеграл от плотности нормального распределения не берется в элементарных функциях, поэтому его выражают через *Ф(x)* – функцию Лапласа:

$$P(a < X < b) = \Phi(Z_2) - \Phi(Z_1).$$

С помощью процедуры «нормализация» находим:

$$Z_1 = \frac{a - M(X)}{\sigma} = \frac{12-10}{2} = 1, \quad Z_2 = \frac{b - M(X)}{\sigma} = \frac{14-10}{2} = 2.$$

Значения функции Лапласа, приведенные в таблице (*Приложение 1*), представляют собой вероятность попадания случайной величины в интервал от 0 до Z. Учитывая, что $\Phi(-x) = -\Phi(x)$, легко вычислить вероятность попадания случайной величины в любой интервал[1] (рис. 1.20):

$$P(12 < X < 14) = \Phi(2) - \Phi(1) = 0{,}47725 - 0{,}34134 = 0{,}1359.$$

Z	Ф(z)	Z	Ф(z)	Z	Ф(z)	Z	Ф(z)	Z	Ф(z)	Z	Ф(z)
0,00	0,00000	0,50	0,19146	1,00	0,34134	1,50	0,43319	2,00	0,47725	3,00	0,49865
0,01	0,00399	0,51	0,19497	1,01	0,34375	1,51	0,43448	2,02	0,47831	3,05	0,49886
0,02	0,00798	0,52	0,19847	1,02	0,34614	1,52	0,43574	2,04	0,47932	3,10	0,49903

Рис. 1.20. Нахождение значения функции Лапласа

Нормальный закон играет особую роль и при решении прикладных задач из сферы психологии.

Следует заметить, что уже со второй половины XIX в. измерительные методы в психологии разрабатывались на основе следующего принципа: если индивидуальная изменчивость некоторого свойства есть результат воздействия множества случайных причин, то закономерность встречаемости значений этого свойства соответствует кривой нормального распределения.

[1] Это значение равно величине площади стандартной нормальной кривой от 0 до соответствующего значения Z.

Контрольные задания к главе 1

1. Из букв слова «математика» наугад выбирают одну букву. Какова вероятность, что это окажется: а) буква «м»; б) буква «а»; в) гласная буква; г) согласная буква?

2. Бросают два игральных кубика. Какова вероятность, что сумма выпавших очков окажется равной: а) 2; б) 3; в) 4; г) 7; д) 12?

3. Абонент забыл две цифры нужного ему номера и набирает их наудачу. Найти вероятность того, что номер будет набран правильно, если абонент помнит, что эти цифры различные.

4. Среди 20 студентов учебной группы, в которой 10 девушек, разыгрывается 5 билетов в театр. Определить вероятность того, что среди обладателей билетов окажутся три девушки.

5. В прямоугольном броневом щите размерами 2 на 1 м имеется невидимая для противника амбразура размерами 10 на 10 см. Определить вероятность того, что пуля, попавшая в щит, попадет в амбразуру, если попадание в любую точку щита равновозможно.

6. Из колоды в 36 карт случайным образом последовательно выбирают три карты. Какова вероятность, что этими картами окажутся шестерка, червовая десятка и король?

7. Студент разыскивает формулу в трех справочниках. Вероятность того, что формула окажется в первом, втором, третьем справочниках, соответственно равна 0,6; 0,7; 0,8. Найти вероятность того, что формула окажется: а) только в одном справочнике; б) хотя бы в одном справочнике; в) во всех трех справочниках; г) ни в одном справочнике; д) только во втором справочнике.

8. Вероятность того, что клиент банка не погасит кредит в период экономического роста, равна 0,04, а период экономического кризиса – 0,13. Предположим, что период экономического роста может наступить с вероятностью 0,65. Найти вероятность того, что случайно выбранный клиент банка не погасит кредит.

9. Вероятность попадания в цель при выстреле из первого, второго и третьего орудий соответственно равна 0,4; 0,6; 0,5. Для разрушения моста достаточно одного попадания в цель из любого

орудия. Какова вероятность, что мост будет разрушен? Какова вероятность, что мост разрушили выстрелом из третьего орудия?

10. Вероятность прибытия англоязычной группы туристическое агентство оценивает величиной 0,6. Прибывает шесть групп. Найти вероятность того, что: а) из прибывших групп ровно четыре англоязычные; б) среди прибывших групп будет не более двух англоязычных.

11. Вероятность госпитализации пациента при эпидемии гриппа равна 0,002. Найти вероятность того, что из 2 000 заболевших поликлиника направит на госпитализацию не более 5 пациентов.

12. Непрерывная случайная величина X задана интегральной функцией распределения:

$$\mathrm{F}(x) = \begin{cases} 0, \text{ где } x \le 0 \\ x^2/4, \text{ где } 0 < x \le 2 \\ 1, \text{ где } x > 2. \end{cases}$$

Найти плотность распределения $f(x)$, вычислить вероятность того, что в результате испытания X примет значение, принадлежащее интервалу (1; 2), определить $M(X)$, $D(X)$.

ГЛАВА 2. Математическая статистика

2.1. Основные понятия и определения

Математическая статистика позволяет обосновать ответ на вопрос: случайно или закономерно изучаемое явление? Установление статистических закономерностей, присущих массовым случайным явлениям, основано на изучении статистических данных[1].

Статистические данные – данные о том, какие значения принял в результате эксперимента (опыта, испытания) исследуемый признак (случайная величина *X*)[2].

Различные значения признака называют **вариантами**.

Различают **качественные** (спортивная квалификация, стратегия поведения в конфликте и др.) и **количественные** (уровень интеллекта, успеваемость и др.) признаки. В отличие от количественных, качественные признаки не выражаются числовыми значениями.

Генеральная совокупность – множество, включающее все однородные объекты, которым присущи (или не присущи) определенные количественные и (или) качественные признаки.

Полное обследование генеральной совокупности часто бывает неэкономично, а иногда невозможно. В таком случае из генеральной совокупности делают выборку, т. е. исследуют только некоторые объекты.

Выборочная совокупность (выборка) – совокупность случайно отобранных объектов из генеральной совокупности.

Для того чтобы по объектам выборки можно было достаточно уверенно судить об объектах генеральной совокупности, необходимо, чтобы выборка была **репрезентативной**.

---

[1] Кремер Н. Ш. Теория вероятностей и математическая статистика : учебник для вузов. М. : Юнити-Дана, 2003. С. 264.

[2] Признак – это свойство, характерная черта или иная особенность объекта.

Выборка будет репрезентативной, если каждый объект отобран случайно из генеральной совокупности или все объекты имеют одинаковую вероятность попасть в выборку.

Статистическое наблюдение (эксперимент) – выбор объекта из генеральной совокупности и измерение значения признака.

Обычно данные, полученные во время статистического наблюдения, весьма беспорядочны. Их трудно анализировать. Поэтому, как правило, первым шагом любого анализа являются систематизация наблюдаемых данных и их наглядное представление.

Пример 1. В результате тестирования по математике группа из 28 человек набрала следующие баллы: 3, 3, 3, 2, 2, 5, 2, 3, 2, 2, 3, 4, 3, 2, 2, 4, 3, 2, 2, 3, 4, 2, 3, 4, 3, 3, 4, 3. Построить статистическое распределение выборки и полигон.

Решение. Очевидно, что в нашем случае исследуемый признак (случайная величина X) – количество баллов за тест, отражающее математические способности. Различные значения признака (варианты): 2, 3, 4, 5. Объем выборки $n = 28$.

Проранжируем[1] и сгруппируем исходные данные так, чтобы каждая группа соответствовала определенному варианту (x_i). Число элементов в каждой группе называют **частотой** варианта (n_i), а отношение частоты варианта к объему выборки – **относительной частотой** варианта[2].

Статистическое распределение выборки – перечень вариантов и соответствующих им частот или относительных частот[3]. Ранжированный перечень вариантов и соответствующих им

---

[1] Ранжирование – расположение значений признака (вариант) по возрастанию.

[2] Сумма всех частот равна объему выборки. Сумма всех относительных частот равна 1.

[3] В теории вероятностей распределение – это соответствие между возможными значениями случайной величины и их вероятностями, а в статистике – соответствие между наблюдаемыми значениями случайной величины и их частотами (относительными частотами).

частот или относительных частот называют также **дискретным вариационным рядом**.

В итоге получаем статистическое распределение выборки (табл. 2.1).

Таблица 2.1

Статистическое распределение выборки

x_i	2	3	4	5	Контроль
n_i	10	12	5	1	$\sum_{i=1}^{4} n_i=28$
$\frac{n_i}{n}$	$\frac{10}{28}$	$\frac{12}{28}$	$\frac{5}{28}$	$\frac{1}{28}$	$\sum_{i=1}^{4} \frac{n_i}{n}=1$

Дискретный вариационный ряд изображают графически с помощью полигона (рис. 2.1) – ломаной, отрезки которой соединяют точки: $(x_1; n_1)$, $(x_2; n_2)$,..., $(x_к; n_к)$, $(x_1; \frac{n_1}{n})$, $(x_2; \frac{n_2}{n})$,..., $(x_к; \frac{n_к}{n})$.

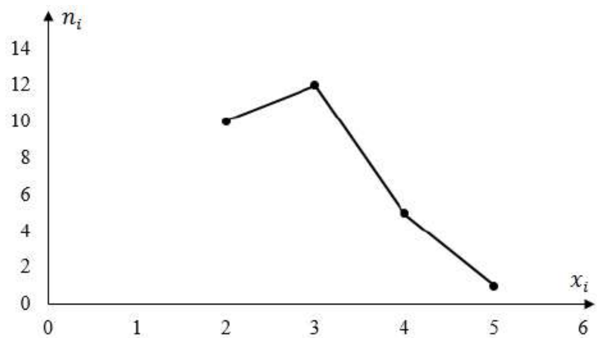


Рис. 2.1. Полигон частот

Если число значений признака велико или случайная величина является непрерывной и не удается выделить характерные черты варьирования ее значений[1], то строят **интервальный вариационный ряд**:

1. Промежуток варьирования наблюдаемых значений признака разбивают на ряд отдельных интервалов и подсчитывают количество значений признака, попадающего в каждый интервал.
2. Количество интервалов может быть определено по формуле Стерджесса: $k = 1+3{,}322lgn$, где n – объем выборки.
3. Считают, что отдельные интервалы имеют одинаковую длину h:

$$h = \frac{x_{max} - x_{min}}{k}.$$

где x_{max} – наибольшее значение признака, а x_{min} – наименьшее значение признака.

Пример 2. Средний балл успеваемости по математике в группе из 28 человек имеет следующие значения: 2,58; 2,65; 2,67; 2,72; 2,73; 2,75; 2,84; 2,87; 3,00; 3,06; 3,06; 3,06; 3,06; 3,07; 3,13; 3,13; 3,17; 3,17; 3,36; 3,38; 3,50; 3,60; 3,79; 3,79; 3,92; 4,08; 4,30; 4,77. Построить интервальный вариационный ряд и гистограмму.

Решение. Определяем количество интервалов по формуле Стерджесса: $k = 1 + 3{,}322lgn = 5{,}807 \approx 6$. Определяем длину интервала:

$$h = \frac{x_{max} - x_{min}}{k} = \frac{4{,}77 - 2{,}58}{6} = 0{,}365.$$

В качестве начального значения интервального вариационного ряда используем $x_{min} = 2{,}58$. Разобьем промежуток изменения признака на 6 интервалов с $h = 0{,}365$. Например, верхняя граница первого интервала находится так: $x_{min} + h = 2{,}58 + 0{,}365 = 2{,}945$.

Подсчитаем, сколько значений признака (вариант) попадает в каждый интервал:

---

[1] Отдельные значения случайной величины могут мало отличаться друг от друга, а одинаковые значения встречаются редко.

Интервал	n_i
2,580–2,945	8
2,945–3,310	10
3,310–3,675	4
3,675–4,040	3
4,040–4,405	2
4,405–4,770	1

Строим **гистограмму частот** – фигуру, состоящую из прямоугольников с основанием h и высотами n_i (рис. 2.2) или $\frac{n_i}{n}$.

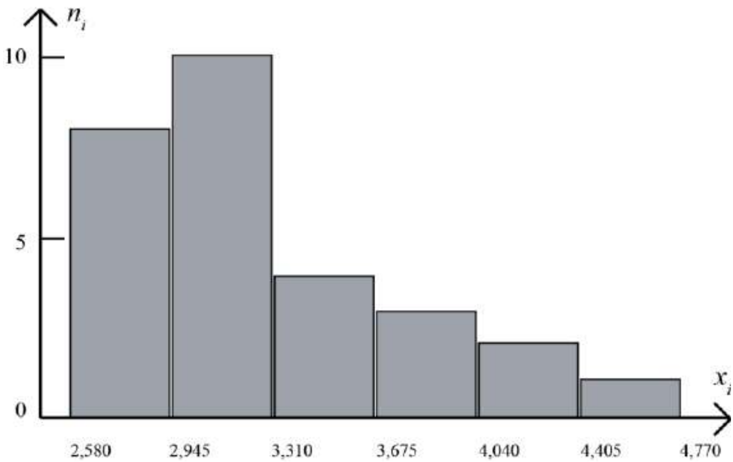


Рис. 2.2. Гистограмма частот

2.2. Основные статистические показатели

Напомним, что под **распределением признака** понимают закономерность встречаемости разных его значений.

Параметры распределения – это его числовые характеристики, показывающие, где в «среднем» располагаются значения признака и насколько они изменчивы[1]. В реальных психологических исследованиях оперируют не параметрами, а их оценками, которые называются **выборочными параметрами** или **статистиками**[2].

Оценки могут быть **точечными** и **интервальными**.

Точечной оценкой называется число, которое используется в качестве оценки неизвестного параметра генеральной совокупности[3]. Все точечные оценки можно разделить на три группы:

1. **Средние статистики** (меры центральной тенденции) – это статистические показатели, которые дают усредненную характеристику совокупности объектов по определенному признаку. Основные средние статистики: *среднее значение* (выборочное среднее)[4], *медиана*, *мода*.

2. **Статистики рассеяния** (меры рассеяния) – это статистические показатели, характеризующие различия между отдельными значениями выборки. Они позволяют судить о степени однородности полученного множества. Наиболее используемые в психологических исследованиях показатели: *выборочная дисперсия*[5], *стандартное отклонение* (выборочное среднее квадратическое отклонение)[6], *размах* (диапазон разброса данных).

3. **Статистики отклонения формы распределения** – это статистические показатели, которые позволяют количественно

---

[1] Параметры распределения – параметры генеральной совокупности: математическое ожидание, дисперсия, среднее квадратическое отклонение. Параметр – фиксированное, но неизвестное число, вычисленное для всей совокупности.

[2] Параметр, вычисленный по данным выборки.

[3] Интервальной оценкой называется вычисленный на основе выборки интервал значений случайной величины, который с известной вероятностью содержит истинное значение оцениваемого параметра генеральной совокупности.

[4] Выборочное среднее является лучшей (эффективной, состоятельной, несмещенной) оценкой математического ожидания генеральной совокупности.

[5] Выборочная дисперсия является лучшей оценкой дисперсии генеральной совокупности.

[6] Стандартное отклонение является лучшей оценкой среднего квадратического отклонения генеральной совокупности.

оценить меру отклонения изучаемого распределения от нормального распределения: *асимметрия*, *эксцесс*.

Цель нахождения статистических показателей – сделать выводы и принять стратегические (для дальнейшего анализа) решения, основанные на имеющихся данных.

Пример 3. Число опечаток, сделанных студентами на заключительном экзамене по машинописи, составило: 3, 11, 5, 12, 9, 8, 16, 10, 9. Рассчитайте основные статистические показатели.

Решение. Для автоматизации процесса нахождения основных статистических показателей используем табличный процессор Excel:

1. В ячейки С2:С10 вводим результаты девяти измерений и вычисляем *выборочное среднее* $\overline{X}$: в ячейку С11 вводим =СРЗНАЧ(С2:С10) (рис. 2.3). В результате округления получаем $\overline{X} = 9,2$. Формула для ручного подсчета:

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

	A	B	C	D
1		№	X	
2		1	3	
3		2	11	
4		3	5	
5		4	12	
6		5	9	
7		6	8	
8		7	16	
9		8	10	
10		9	9	
11		Среднее	=СРЗНАЧ(C2:C10	9R x 1C

Рис. 2.3. Среднее значение

2. Вычисляем **моду Мо**[1]: в ячейку C12 вводим =МОДА(C2:C10). В результате получаем: $Mo = 9$.

3. Вычисляем **медиану Ме**[2]: в ячейку C13 вводим =МЕДИАНА(C2:C10). В результате получаем: $Me = 9$.

4. Вычисляем **выборочную дисперсию**[3] S^2: в ячейку C14 вводим =ДИСП.В(C2:C10). В результате получаем: $S^2 = 14,4$.

5. Вычисляем **стандартное отклонение**[4] S: в ячейку C15 вводим =СТАНДОТКЛОН.В(C2:C10). В результате получаем $S = 3,8$. Формула для ручного подсчета:

$$S = \sqrt{\frac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \ldots + (x_n - \overline{x})^2}{n - 1}}\,.$$

6. Вычисляем размах выборки[5] W: в ячейку C16 вводим =МАКС(C2:C10)-МИН(C2:C10). В результате получаем $W = 13$.

7. Вычисляем **асимметрию** распределения[6] A: в ячейку C17 вводим =СКОС(C2:C10). В результате получаем $A = 0,06$.

8. Вычисляем **эксцесс** распределения[7] E: в ячейку C17 вводим =ЭКСЦЕСС(C2:C10). В результате получаем $E = 0,56$. Положительное значение эксцесса свидетельствует о более «острой» вершине по сравнению с вершиной нормального распределения[8].

---

[1] Мода – значение выборки, наблюдаемое наибольшее число раз.

[2] Медиана – значение выборки, которое делит вариационный ряд на две равные части.

[3] Выборочная дисперсия – это среднее всех квадратов отклонений результатов наблюдений от их среднего значения.

[4] Стандартное отклонение описывает типичное расстояние от среднего значения для отдельных результатов измерений.

[5] Размахом выборки называется разность между наибольшим значением и наименьшим значением этой выборки.

[6] Числовое отображение степени отклонения графика кривой данного распределения от симметричного нормального распределения.

[7] Числовое отображение остроты пика (вершины) кривой данного распределения относительно пика кривой нормального распределения.

[8] Для нормального распределения $A = 0$ и $E = 0$.

9. Итак, получаем основные статистические показатели (рис. 2.4).

11	Среднее	9,2
12	Мода	9
13	Медиана	9
14	Дисперсия	14,4
15	Станд.отклон.	3,8
16	Размах	13
17	Асимметрия	0,06
18	Эксцесс	0,56

Рис. 2.4. Статистические показатели

2.3. Статистическая проверка статистических гипотез

Метод проверки статистических гипотез позволяет на основе имеющейся информации сделать выбор между двумя альтернативными предположениями, называемыми гипотезами. Проверку статистических гипотез можно рассматривать как один из компонентов принятия решений.

Статистическая гипотеза – всякое утверждение о генеральной совокупности, проверяемое по выборке.

Нулевая гипотеза – утверждение об отсутствии статистически значимых различий или связи.

Альтернативная гипотеза – утверждение о наличии статистически значимых различий или связи.

Уровень значимости α – вероятность ошибки, связанной с обобщением результата измерения, полученного по выборке, на всю генеральную совокупность. Уровень значимости α – достаточно малое число, которое может принимать значения из промежутка (0; 1).

Например, $\alpha = 0,05$; $\alpha = 0,01$ и т. п.

Статистический критерий – строгое математическое правило, согласно которому отвергается или принимается та или иная статистическая гипотеза с известным уровнем значимости[1].

Различают параметрические и непараметрические критерии. Параметрические критерии предназначены для проверки гипотез о параметрах распределений, вид которых известен[2]. В частности, многие параметрические критерии опираются на предположение, что распределение, которому принадлежат исследуемые выборки, соответствует нормальному. Расчет статистики параметрического критерия основывается на параметрах распределения (среднее, дисперсия, стандартное отклонение и др.).

Статистика непараметрического критерия основана на оперировании рангами или частотами. Непараметрические критерии могут применяться для распределений, вид которых неизвестен.

Ошибка первого рода, возникающая при проверке статистических гипотез, заключается в отклонении верной гипотезы: нулевую гипотезу отвергают, когда она верна.

Ошибка второго рода, возникающая при проверке статистических гипотез, заключается в принятии ложной гипотезы: нулевую гипотезу не отвергают, когда она ложна.

Этапы проверки гипотезы:

1) сформулировать нулевую (основную) и альтернативную (конкурирующую) гипотезы;

---

[1] Задохина Н. В. Статистические методы в деятельности эксперта-криминалиста. С. 196.

[2] Вид распределения исследуемой величины считается известным, если может быть определено, какова частота тех или иных значений этой величины в общей совокупности. Различают теоретическое (предельное или эталонное) и эмпирическое (опытное) распределения. Теоретическое распределение – распределение результатов измерений исследуемой величины, которое получилось бы, если число этих измерений было бы бесконечно велико. Эмпирическое распределение (статистическое распределение выборки) – соответствие между измеренными значениями исследуемой величины, записанными в порядке возрастания, и их частотами.

2) задать уровень значимости;

3) выбрать статистический критерий;

4) найти критическое значение критерия и построить критическую область;

5) найти эмпирическое значение критерия (статистику);

6) сделать вывод. Если эмпирическое значение критерия попало в критическую область, то нулевую гипотезу следует отклонить и принять альтернативную[1].

Указанная последовательность шагов применима ко всем критериям.

Пример 4. Университет ввел программу психологической подготовки студентов к итоговой аттестации (с целью корректировки психоэмоционального напряжения). Из предыдущих исследований известно, что у обучающихся в среднем частота сердечных сокращений перед экзаменом составляет 100 уд/мин.

Повлияла ли программа психологической подготовки на психоэмоциональное напряжение обучающихся, если у 9 случайно отобранных студентов, сдававших итоговую аттестацию, сердце билось(уд/мин): 96, 93, 99, 97, 100, 94, 98, 95, 96.

Решение. Задача сводится к проверке равенства выборочного среднего заданному значению, в качестве которого выступает гипотетическое генеральное среднее (математическое ожидание) распределения:

1. Формулируем нулевую и альтернативную гипотезы:

H_0: Программа психологической подготовки не повлияла на психоэмоциональное напряжение обучающихся;

H_1: Программа психологической подготовки повлияла на психоэмоциональное напряжение обучающихся.

2. Выбираем уровень значимости: $\alpha = 0{,}05$.

[1] Задохина Н. В. Статистические методы в деятельности эксперта-криминалиста. С. 195.

3. Выбираем статистический критерий: t-критерий[1].

4. Находим критическое значение критерия и строим критическую область**:** в нашем случае находим двустороннее значение критерия и строим двустороннюю критическую область.

Для нахождения двустороннего критического значения критерия (рис. 2.5) используем таблицу критических точек распределения Стьюдента (*Приложение 2*).

В нашем случае: $\alpha = 0{,}05$, $k = n - 1 = 9 - 1 = 8$. Получаем: $t_{кр}=2{,}31$.

Число степеней свободы *к*	**Уровень значимости α (двусторонняя критическая область)**					
	0,10	**0,05**	**0,02**	**0,01**	**0,002**	**0,001**
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	4,04

Рис. 2.5. Нахождение критического значения критерия

[1] t-критерий – это общее название для методов статистической проверки гипотез (статистических критериев), которые основаны на распределении Стьюдента. При малом количестве измерений распределение может более или менее отклоняться от нормального. Эта дополнительная ненадежность устраняется с помощью симметричного t-распределения. Распределение Стьюдента (t-распределение) зависит только от объема выборки (степеней свободы) и не зависит от параметров генеральной совокупности. При увеличении количества измерений вид t-распределения приближается к нормальному распределению. t-критерий может применяться для исследования малых выборок ($n<30$). Для применения t-критерия необходимо, чтобы выборки извлекались из совокупности, имеющей нормальное распределение.

Строим двустороннюю критическую область (рис. 2.6).

Рис. 2.6. Критическая область

5. Находим эмпирическое значение критерия (статистику):

$$t_{эмп} = \frac{\overline{X} - M(X)}{\left(\frac{S}{\sqrt{n}}\right)},$$

где $M(X)=100$ *уд./мин* – гипотетическое генеральное среднее.

$$\overline{X} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{96 + 93 + 99 + \ldots + 96}{9} = 96{,}44 \text{ уд./мин.}$$

$$S = \sqrt{\frac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \ldots + (x_n - \overline{x})^2}{n-1}} =$$

$$= \sqrt{\frac{(96-96{,}44)^2 + (93-96{,}44)^2 + \ldots + (96-96{,}44)^2}{9-1}} = 2{,}3 \text{ уд./мин.}$$

Учитывая, что $n = 9$, находим эмпирическое значение критерия (рис. 2.7):

$$t_{эмп} = \frac{\overline{X} - M(X)}{\left(\frac{S}{\sqrt{n}}\right)} = \frac{96{,}44-100}{\left(\frac{2{,}3}{3}\right)} = -4{,}64.$$

6. Формулируем вывод: так как $t_{эмп}$ попало в критическую область, H_0 отклоняется.

Вывод: программа психологической подготовки повлияла на психоэмоциональное напряжение обучающихся.

Рис. 2.7. Эмпирическое значение критерия

Пример 5. В двух студенческих группах проводилось измерение уровня интеллекта по тесту Векслера (табл. 2.2). Есть ли различия в степени однородности показателей интеллекта между группами?

Таблица 2.2

Результаты теста Векслера

№	*I*	*II*
1	110	98
2	100	116
3	114	128
4	115	120
5	106	100
6	120	134
7	124	112
8	116	100
9	118	128
10	122	110

Решение. Задача сводится к сравнению дисперсий двух независимых выборок:

1. Формулируем нулевую и альтернативную гипотезы:

H_0: нет различий в степени однородности показателей интеллекта между группами (дисперсии распределений, из которых извлечены выборки статистически значимо не отличаются);

H_1: есть различия в степени однородности показателей интеллекта между группами (дисперсии распределений, из которых извлечены выборки статистически значимо отличаются).

2. Выбираем уровень значимости: $\alpha = 0,05$.
3. Выбираем статистический критерий: F-критерий[1].
4. Находим критическое значение критерия:

Для нахождения критического значения F-критерия (рис. 2.8) используем функцию Excel – FРАСПОБР (альфа; степени свободы_1; степени свободы_2), которая имеет три аргумента: уровень значимости α; степени свободы первой выборки[2]; степени свободы второй выборки.

В нашем случае: $\alpha = 0,05$, $n - 1 = 10 - 1 = 9$; $n - 1 = 10 - 1 = 9$.

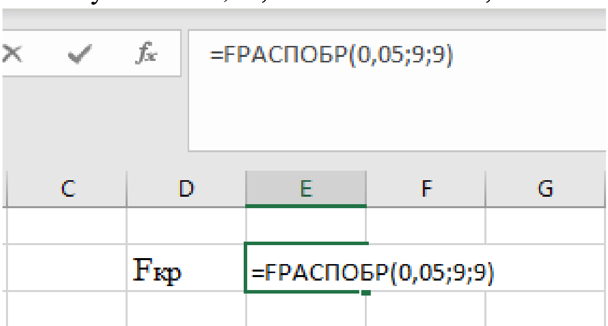


Рис. 2.8. Критическое значение критерия

[1] Для применения F-критерия необходимо, чтобы выборки извлекались из совокупности, имеющей нормальное распределение.

[2] Первой считается та выборка, которая имеет большую дисперсию S^2.

Получаем: $F_{кр} = 3,18.$

5. Находим эмпирическое значение критерия (статистику):

$$F_{эмп} = \frac{S_б^2}{S_м^2} \quad ,$$

где: $S_б^2$ – большая выборочная дисперсия; $S_м^2$ – меньшая выборочная дисперсия.

Для нахождения выборочной дисперсии (рис. 2.9) используем функцию Excel – ДИСП.В (число1;…), которая в нашем случае имеет десять аргументов – результаты измерения, входящие в соответствующую выборку.

Вводим в ячейку B12 формулу =ДИСП.В(B2:B11). Копируем эту формулу в ячейку C12. Получаем: $S_м^2$= 54,94 и $S_б^2$ = 166,27.

B12 | =ДИСП.В(B2:B11)

	A	B	C
1	**№**	**I**	**II**
2	1	110	98
3	2	100	116
4	3	114	128
5	4	115	120
6	5	106	100
7	6	120	134
8	7	124	112
9	8	116	100
10	9	118	128
11	10	122	110
12	**Дисперсия**	**54,94**	**166,27**

Рис. 2.9. Нахождение выборочной дисперсии

Далее находим эмпирическое значение критерия:

$$F_{эмп} = \frac{S_б^2}{S_м^2} = \frac{54,94}{166,27} = 3,0.$$

6. Формулируем вывод: $F_{эмп} < F_{кр}$ ($F_{эмп}$ не попало в критическую область), поэтому H_0 не отклоняется.

Вывод: различий в степени однородности показателей интеллекта между группами нет.

Пример 6. При исследовании скорости реакции двух групп пациентов на звуковой сигнал в различных условиях (условия *A* и условия *B*) были получены следующие данные (в сотых долях секунды).

Условия А	11,5	12,5	13,5	12,5	14,0	11,0
Условия В	12,5	14,0	13,5	14,0	16,2	15,0

Проверить, оказала ли разница условий влияние на скорость реакции пациентов.

Решение. Задача сводится к сравнению средних значений двух независимых выборок:

1. Формулируем нулевую и альтернативную гипотезы:

H_0: разница условий не оказала влияние на скорость реакции пациентов (средние распределений[1], из которых извлечены выборки, равны);

H_1: разница условий оказала влияние на скорость реакции пациентов (среднее распределение, из которого извлечена выборка, соответствующая условию *B*, больше среднего распределения, из которого извлечена выборка, соответствующая условию *A*).

2. Выбираем уровень значимости: $\alpha = 0{,}05$ (по условию).

3. Выбираем статистический критерий: t-критерий[2] для разности двух средних независимых выборок (двухвыборочный t-тест).

---

[1] Среднее распределения – генеральное среднее или математическое ожидание.

[2] Для применения критерия необходимо, чтобы выборки извлекались из совокупностей, имеющих нормальное распределение.

Проверим выборки с помощью F-критерия (рис. 2.10) и убедимся, что дисперсии не имеют статистически значимых различий (см. Пример 5).

$F_{эмп} < F_{кр}$ ($F_{эмп}$ не попало в критическую область), поэтому можем применить двухвыборочный t-тест с одинаковыми дисперсиями[1].

11 =C9/B9

A	B	C
№ измерения	**А**	**В**
1	11,5	12,5
2	12,5	14,0
3	13,5	13,5
4	12,5	14,0
5	14,0	16,2
6	11,0	15,0
Среднее	**12,5**	**14,2**
Дисперсия	**1,3**	**1,6**
Fкр	**5,05**	
Fэмп	**1,25**	

Рис. 2.10. Нахождение эмпирического значения F-критерия

4. Находим критическое значение критерия: в нашем случае находим правостороннее значение критерия. Для нахождения правостороннего критического значения критерия используем таблицу *Приложения 2* или выражение Excel =

[1] Если $F_{эмп} > F_{кр}$, т. е. дисперсии значимо различаются, то применяют двухвыборочный t-тест с различными дисперсиями.

ABS(СТЬЮДЕНТ.ОБР (альфа; степени свободы)). В нашем случае: $\alpha = 0,05$, $k = n_1 + n_2 - 2 = 12 - 2 = 10$, где n_1, n_2 – объемы выборок. Получаем: $t_{кр.пр}=1,81$.

5. Находим эмпирическое значение критерия (статистику):

1) выбираем пакет «Анализ данных[1]», который содержит группа инструментов «Анализ» вкладки «Данные» табличного процессора Excel (рис. 2.11);

2)

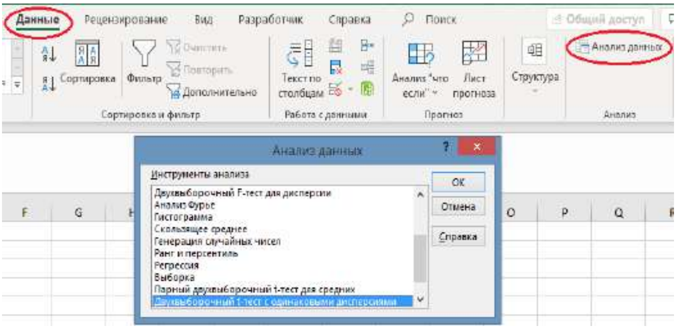


Рис. 2.11. Пакет инструментов Excel «Анализ данных»

3) в диалоговом окне пакета инструментов «Анализ данных» выбираем «Двухвыборочный t-тест с одинаковыми дисперсиями» и вводим необходимые данные (рис. 2.12);

[1] Если пакет «Анализ данных» не активен, то следует выполнить последовательность шагов: Файл-Параметры-Надстройки-Пакет анализа-Перейти…– Пакет анализа-Ок.

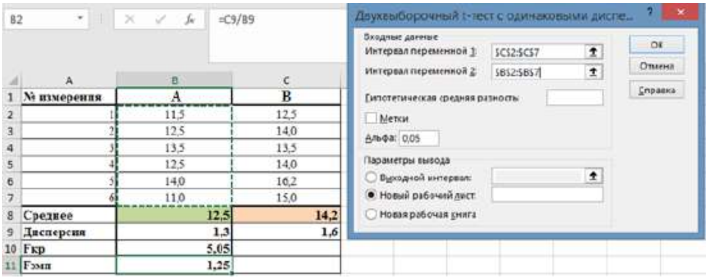


Рис. 2.12. Двухвыборочный t-тест с одинаковыми дисперсиями

4) среди полученных результатов, которые отображаются на новом листе Excel, определяем эмпирическое значение t-критерия (t-статистику): $t_{эмп}=2,4368... \approx 2,44$, $t_{кр} = 1,8124 \approx 1,81$ (рис. 2.13).

	A	B	C	D
1	Двухвыборочный t-тест с одинаковыми дисперсиями			
2				
3		*Переменная 1*	*Переменная 2*	
4	Среднее	14,2	12,5	
5	Дисперсия	1,62	1,3	
6	Наблюдения	6	6	
7	Объединенная дисперсия	1,46		
8	Гипотетическая разность средних	0		
9	df	10		
10	t-статистика	2,436874261		
11	P(T<=t) одностороннее	0,017517907		
12	t критическое одностороннее	1,812461123		
13	P(T<=t) двухстороннее	0,035035814		
14	t критическое двухстороннее	2,228138852		

Рис. 2.13. Нахождение эмпирического значения двухвыборочного t-критерия

6. Формулируем вывод: $t_{эмп} > t_{кр}$ ($t_{эмп}$ попало в критическую область), поэтому H_0 отклоняется.

Вывод: разница условий оказала влияние на скорость реакции пациентов (скорость реакции пациентов в условиях *B* медленнее).

2.4. Корреляционно-регрессионный анализ

Статистической называют зависимость, при которой изменение одной из случайных величин влечет изменение распределения другой. Статистическую зависимость называют **корреляционной**, если при изменении одной из случайных величин изменяется среднее значение другой.

С корреляционным анализом тесно связан **регрессионный** анализ. В корреляционном анализе оценивается сила связи между переменными, в регрессионном – форма. Регрессионный анализ позволяет аналитически представить связь между переменными. Эта аналитическая связь называется **уравнением регрессии**. Регрессия – «движение» назад[1].

Пусть проводится некоторый эксперимент, целью которого является исследование зависимости величины y от другой величины x: $y = \varphi(x)$. В результате эксперимента получены несколько пар значений x и y. Таким образом, на основании экспериментальных данных требуется установить вид зависимости.

Простое проведение через все экспериментальные точки некоторой кривой, являющейся графиком этой зависимости, лишено смысла, так как вид этой зависимости будет меняться от одной серии измерений к другой из-за наличия случайных погрешностей.

Очевидно, что необходимо подобрать функцию, график которой проходил бы как можно ближе к экспериментальным точкам.

---

[1] Термин «регрессия» появился в 1885 г. в работе английского антрополога Френсиса Гальтона.

Такую функцию называют **аппроксимирующей**[1]. Аппроксимирующая функция должна быть достаточно простой и адекватно отражать зависимость между переменными. Один из методов нахождения таких функций – метод наименьших квадратов (МНК).

Пусть вид аппроксимирующей функции $\varphi(x) = \varphi(x, a, b, c, \ldots)$ известен. Параметры $a, b, c, \ldots$ аппроксимирующей функции подбираются так, чтобы:

1) функция располагалась наиболее близко к экспериментальным точкам;

2) сумма квадратов отклонений ординат экспериментальных точек от соответствующих ординат графика аппроксимирующей функции была минимальной:

$$S(a, b, c, \ldots) = \sum_{i=1}^{n} (y_i - \varphi(x_i, a, b, c, \ldots))^2 \to min.$$

Пример 7. Построить уравнение линейной регрессии, определяющее взаимосвязь между ростом и стопой человека, если в результате наблюдения составлена таблица зависимости роста и длины стопы (табл. 2.3). Сделать вывод о характере корреляционной связи между переменными x и y.

Таблица 2.3

Зависимость роста и длины стопы

Х	**У**
22,5	164
22,8	163
23,0	165
23,0	163

[1] В качестве аппроксимирующей функции не рекомендуется использовать многочлен высокой степени, так как график такой функции «петляет», плохо отражая главную тенденцию.

Окончание табл. 2.3

23,0	165
23,5	168
24,0	175
24,0	174
24,0	165
24,2	175
24,5	178
24,5	169
24,8	179

Решение. Используем табличный процессор Excel для автоматизации процесса нахождения параметров уравнения линейной регрессии (a, b) методом наименьших квадратов.

1. В ячейки B2:O3 вводим табличные данные (рис. 2.14):

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1																
2		Х	22,5	22,8	23,0	23,0	23,0	23,5	24,0	24,0	24,0	24,2	24,5	24,5	24,8	
3		У	164	163	165	163	165	168	175	174	165	175	178	169	179	
4																

Рис. 2.14. Исходные данные

2. По данным таблицы строим точечную диаграмму (рис. 2.15):

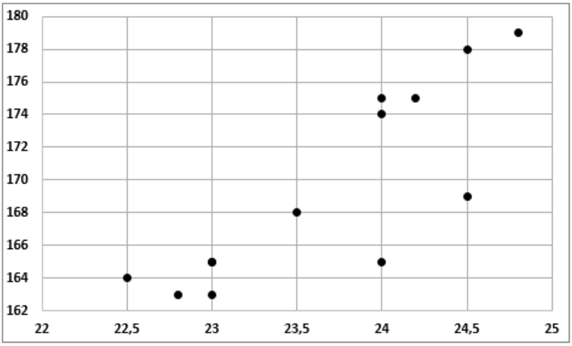


Рис. 2.15. Экспериментальные точки

3. Подбираем аппроксимирующую функцию МНК: Выделяем правой клавишей мыши экспериментальные точки и вызываем контекстное меню. Выбираем команду «Добавить линию тренда». В диалоговом окне «Формат линии тренда» выбираем линейную зависимость. В нижней части диалогового окна «Формат линии» тренда ставим два флажка: показывать уравнение на диаграмме; поместить на диаграмму величину достоверности аппроксимации R^2.

Уравнение регрессии имеет вид (рис. 2.16):

$$y = 6{,}6299x+12{,}486;\ R^2 = 0{,}7057.$$

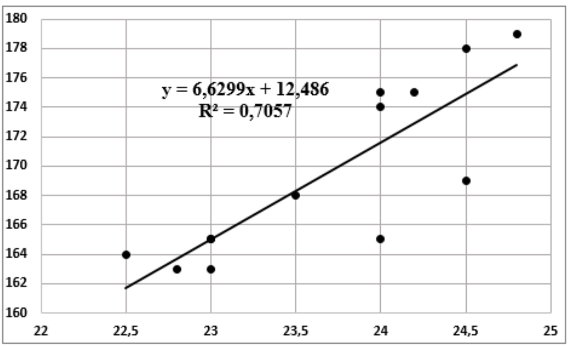
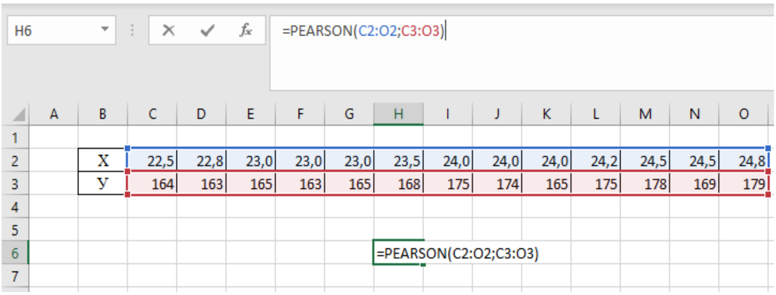


Рис. 2.16. Построение уравнения линейной регрессии

4. Делаем вывод о характере корреляционной связи между переменными x и y. С помощью табличного процессора Excel вычисляем коэффициент корреляции Пирсона r [1] (рис. 2.17): в ячейку H4 вводим =PEARSON(C2:O2;C3:O3). Аргументами функции PEARSON(массив 1; массив 2) служат элементы первой и второй выборки соответственно.

В результате получаем $r = 0{,}84$, что свидетельствует о положительной (достаточно сильной) корреляции между переменными x и y.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1															
2		X	22,5	22,8	23,0	23,0	23,0	23,5	24,0	24,0	24,0	24,2	24,5	24,5	24,8
3		У	164	163	165	163	165	168	175	174	165	175	178	169	179
4															
5															
6								=PEARSON(C2:O2;C3:O3)							
7															

Рис. 2.17. Нахождение коэффициента корреляции Пирсона

Чем ближе значение коэффициента корреляции r к единице (по модулю), тем теснее корреляционная связь.

Если $r = 1$, то наблюдается полная положительная корреляция.

Если $r = -1$, то наблюдается полная отрицательная корреляция.

Так как исходные данные являются выборочными, то необходимо оценить значимость коэффициента корреляции.

1. Формулируем нулевую и альтернативную гипотезы:

H_0: корреляционная связь статистически не значима[2];

H_1: корреляционная связь статистически значима[3];

---

[1] Предварительно необходимо проверить, что выборки извлечены из совокупностей, имеющих нормальное распределение.

[2] $r_г = 0$ – коэффициент корреляции в генеральной совокупности равен нулю.

[3] $r_г \neq 0$ – коэффициент корреляции в генеральной совокупности не равен нулю.

2. Выбираем уровень значимости: $\alpha = 0,05$.
3. Выбираем статистический критерий: t-критерий.
4. Находим критическое значение критерия:

В нашем случае: $\alpha = 0,05$, $n - 2 = 13 - 2 = 11$. Получаем: $t_{кр} = 2,2$ (формула=СТЬЮДЕНТ.ОБР.2Х (0,05;11).

5. Находим эмпирическое значение критерия (статистику):

$$t_{эмп} = \frac{|r|}{\sqrt{\frac{1 - r^2}{n - 2}}} = \frac{0,84}{\sqrt{\frac{1 - 0,84^2}{13 - 2}}} = 5,13,$$

где: $n = 13$ – число элементов выборки; r – коэффициент корреляции.

Строим двустороннюю критическую область и отмечаем $t_{эмп}$ (рис. 2.18).

Рис. 2.18. Двусторонняя критическая область

6. Формулируем вывод: так как $t_{эмп}$ попало в критическую область, H_0 отклоняется и принимается H_1.

Статистическая надежность уравнения регрессии проверяется с помощью критерия Фишера (F-критерия). Если оценивается уравнение линейной регрессии, то эмпирическое значение F-критерия рассчитывается по формуле[1]:

---

[1] Горелова Г. В. Теория вероятностей и математическая статистика в примерах и задачах с использованием Excel : учебное пособие для вузов. Ростов н/Д : Феникс, 2005. С. 281.

$$F_{эмп} = \frac{r^2}{1 - r^2} \cdot (n - 2) = \frac{0{,}84^2}{1 - 0{,}84^2} \cdot (13 - 2) = 26{,}36.$$

Критическое значение критерия $F_{кр}$ находим с помощью табличного процессора MS Excel: $F_{кр} = 4{,}84$. Для вычисления используем формулу: =FРАСПОБР(α; k_1; k_2). где, $\alpha = 0{,}05$; $k_1 = k$[1] $= 1$; $k_2 = n - k - 1 = 13 - 1 - 1 = 11$.

$F_{эмп} > F_{кр}$, следовательно, уравнение регрессии статистически надежно.

[1] k – число параметров при переменной x (количество факторов).

Контрольные задания к главе 2

1. Данные о весе (фунтах) выборки из 11 человек составили: 180, 156, 162, 181, 170, 177, 205, 206, 180, 183, 196. Найти среднее выборочное, дисперсию, стандартное отклонение, моду, медиану.

2. Выборка из 25 оценок при проверке способностей к изучению языков имеет: $\overline{X} = 75$, $s = 7$. Используя двусторонний t-критерий при уровне значимости 0,01, проверьте гипотезу, что данная выборка взята из распределения с $M(X) = 78$.

3. При обследовании двух групп студентов с помощью теста по выявлению способностей к научной деятельности получены следующие результаты: $n_1 = 20$; $S_1^2 = 36$ и $n_2 = 15$; $S_2^2 = 64$. Применив F-критерий при уровне значимости 0,1, проверьте нулевую гипотезу о том, что дисперсии распределений, из которых извлечены выборки, равны (выборки извлекались из генеральных совокупностей, имеющих нормальное распределение).

4. В результате эксперимента получено семь значений функции ***y*** при соответствующих значениях аргумента *x* (табл. 1).

Таблица 1

Результаты

x	92	94	96	98	100	102	104
y	69	34	49	64	54	44	59

Сделайте вывод о характере корреляционной связи между переменными *x* и *y*.

5. В результате эксперимента получено семь значений функции ***y*** при соответствующих значениях аргумента *x* (табл. 2).

Таблица 2

Результаты

x	7	23	35	42	46	48	49
y	12	30	26	29	24	29	36

Сделайте вывод о характере корреляционной связи между переменными *x* и *y*.

6. В результате эксперимента получено семь значений функции y при соответствующих значениях аргумента x (табл. 3).

Таблица 3

Результаты

x	92	94	96	98	100	102	104
y	34	39	44	49	54	59	64

С помощью МНК подберите функциональную зависимость вида $y = ax + b$; сделайте вывод о характере корреляционной связи между переменными x и y.

7. В результате эксперимента получено семь значений функции y при соответствующих значениях аргумента x (табл. 4).

Таблица 4

Результаты

x	36	47	49	52	54	68	76
y	118	108	95	82	77	52	47

С помощью МНК подберите функциональную зависимость вида $y = a \cdot ln(x) + b$; сделайте вывод о характере корреляционной связи между переменными x и y.

ОТВЕТЫ

К главе 1

1. а) 0,2; б) 0,3; в) 0,5; г) 0,5.
2. а) $\frac{1}{36}$; б) $\frac{1}{18}$; в) $\frac{1}{12}$; г) $\frac{1}{6}$; д) $\frac{1}{36}$.
3. $\frac{1}{90}$.
4. 0,348.
5. 0,005.
6. 0,0004.
7. а) 0,188; б) 0,976; в) 0,336; г) 0,024; д) 0,056.
8. 0,0715.
9. 0,5; 1/3.
10. а) 0,311; б) 0,201.
11. 0,815.
12. $P(1 < X < 2) = 0,75$; $M(X) = 4/3$; $D(X) = 2/9$.

К главе 2

1. $\overline{X} = 181$, $S^2 = 255$, $S = 16$.
2. $t_{кр} = \pm 2,79$, $t_{эмп} = -2,14$, H_0 не отклоняется.
3. $F_{кр} = 1,88$, $F_{эмп} = 1,78$, H_0 не отклоняется.
4. $r = 0,03$ корреляционная связь практически отсутствует.
5. $r = 0,74$ положительная (достаточно сильная) корреляция.
6. $y = 2,5x - 196$; r = 1 полная положительная корреляция.
7. $y = -105,5\ln(x) + 501,83$; $r = -0,97$ отрицательная сильная корреляция.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Гмурман, В. Е. Теория вероятностей и математическая статистика : учебное пособие для вузов / В. Е. Гмурман. – 6-е изд. – М. : Высшая школа, 1998. – 479 с.

2. Горелова, Г. В. Теория вероятностей и математическая статистика в примерах и задачах с использованием Excel : учебное пособие для вузов / Г. В. Горелова, И. А. Кацко. – 3-е изд., перераб. и доп. – Ростов н/Д : Феникс, 2005. – 480 с.

3. Задохина, Н. В. Математика и информатика. Решение логико-познавательных задач : учебное пособие / Н. В. Задохина. – М. : Юнити-Дана, 2015. – 127 с.

4. Задохина, Н. В. Статистические методы в деятельности эксперта-криминалиста / Н. В. Задохина // Образование. Наука. Научные кадры. – 2020. – № 3. – С. 195–197.

5. Кремер, Н. Ш. Теория вероятностей и математическая статистика : учебник для вузов / Н. Ш. Кремер. – М. : Юнити-Дана, 2003. – 543 с.

6. Левин, С. Ф. Философские проблемы и статистические методы фундаментальной метрологии / С. Ф. Левин // Метафизика. – 2012. – № 3 (5). – С. 93–94.

7. Математика : учебное пособие. Ч. 2 / сост. А. В. Мамасуев, Н. М. Дубинина, С. А. Скворцова. – 2-е изд., испр. и доп. – М. : Академия экономической безопасности МВД России, 2009. – 122 с.

8. Терентьева, А. Н. Автоматизация сбора и обработки данных психологических исследований / А. Н. Терентьева, Н. В. Задохина // Совершенствование профессиональной подготовки психологов для подразделений органов внутренних дел : Межведомственная научно-практическая конференция : сборник материалов. – М., 2020. – С. 419–422.

ПРИЛОЖЕНИЕ 1

Таблица значений интегральной функции Лапласа

$$\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_0^z e^{\frac{-t^2}{2}}dt$$

Z	Ф(z)	Z	Ф(z)	Z	Ф(z)
0,00	0,00000	1,00	0,34134	2,00	0,47725
0,01	0,00399	1,01	0,34375	2,02	0,47831
0,02	0,00798	1,02	0,34614	2,04	0,47932
0,03	0,01197	1,03	0,34849	2,06	0,48030
0,04	0,01595	1,04	0,35083	2,08	0,48124
0,05	0,01994	1,05	0,35314	2,10	0,48214
0,06	0,02392	1,06	0,35543	2,12	0,48300
0,07	0,02790	1,07	0,35769	2,14	0,48382
0,08	0,03188	1,08	0,35993	2,16	0,48461
0,09	0,03586	1,09	0,36214	2,18	0,48537
0,10	0,03983	1,10	0,36433	2,20	0,48610
0,11	0,04380	1,11	0,36650	2,22	0,48679
0,12	0,04776	1,12	0,36864	2,24	0,48745
0,13	0,05172	1,13	0,37076	2,26	0,48809
0,14	0,05567	1,14	0,37286	2,28	0,48870
0,15	0,05962	1,15	0,37493	2,30	0,48928
0,16	0,06356	1,16	0,37698	2,32	0,48983
0,17	0,06749	1,17	0,37900	2,34	0,49036
0,18	0,07142	1,18	0,38100	2,36	0,49086
0,19	0,07535	1,19	0,38298	2,38	0,49134
0,20	0,07926	1,20	0,38493	2,40	0,49180
0,21	0,08317	1,21	0,38686	2,42	0,49224
0,22	0,08706	1,22	0,38877	2,44	0,49266
0,23	0,09095	1,23	0,39065	2,46	0,49305
0,24	0,09483	1,24	0,39251	2,48	0,49343
0,25	0,09871	1,25	0,39435	2,50	0,49379
0,26	0,10257	1,26	0,39617	2,52	0,49413
0,27	0,10642	1,27	0,39796	2,54	0,49446
0,28	0,11026	1,28	0,39973	2,56	0,49477

Таблица значений интегральной функции Лапласа (продолжение)

$$\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_0^z e^{\frac{-t^2}{2}}dt$$

Z	Ф(z)	Z	Ф(z)	Z	Ф(z)
0,29	0,11409	1,29	0,40147	2,58	0,49506
0,30	0,11791	1,30	0,40320	2,60	0,49534
0,31	0,12172	1,31	0,40490	2,62	0,49560
0,32	0,12552	1,32	0,40658	2,64	0,49585
0,33	0,12930	1,33	0,40824	2,66	0,49609
0,34	0,13307	1,34	0,40988	2,68	0,49632
0,35	0,13683	1,35	0,41149	2,70	0,49653
0,36	0,14058	1,36	0,41309	2,72	0,49674
0,37	0,14431	1,37	0,41466	2,74	0,49693
0,38	0,14803	1,38	0,41621	2,76	0,49711
0,39	0,15173	1,39	0,41774	2,78	0,49728
0,40	0,15542	1,40	0,41924	2,80	0,49744
0,41	0,15910	1,41	0,42073	2,82	0,49760
0,42	0,16276	1,42	0,42220	2,84	0,49774
0,43	0,16640	1,43	0,42364	2,86	0,49788
0,44	0,17003	1,44	0,42507	2,88	0,49801
0,45	0,17364	1,45	0,42647	2,90	0,49813
0,46	0,17724	1,46	0,42785	2,92	0,49825
0,47	0,18082	1,47	0,42922	2,94	0,49836
0,48	0,18439	1,48	0,43056	2,96	0,49846
0,49	0,18793	1,49	0,43189	2,98	0,49856
0,50	0,19146	1,50	0,43319	3,00	0,49865
0,51	0,19497	1,51	0,43448	3,05	0,49886
0,52	0,19847	1,52	0,43574	3,10	0,49903
0,53	0,20194	1,53	0,43699	3,15	0,49918
0,54	0,20540	1,54	0,43822	3,20	0,49931
0,55	0,20884	1,55	0,43943	3,25	0,49942
0,56	0,21226	1,56	0,44062	3,30	0,49952
0,57	0,21566	1,57	0,44179	3,35	0,49960
0,58	0,21904	1,58	0,44295	3,40	0,49960
0,59	0,22240	1,59	0,44408	3,45	0,49972

Таблица значений интегральной функции Лапласа (продолжение)

$$\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_0^z e^{\frac{-t^2}{2}}dt$$

0,60	0,22575	1,60	0,44520	3,50	0,49977
0,61	0,22907	1,61	0,44630	3,53	0,49981
0,62	0,23237	1,62	0,44738	3,60	0,49984
0,63	0,23565	1,63	0,44845	3,65	0,49987
0,64	0,23891	1,64	0,44950	3,70	0,49989
0,65	0,24215	1,65	0,45053	3,75	0,49991
0,66	0,24537	1,66	0,45154	3,80	0,49993
0,67	0,24857	1,67	0,45254	3,85	0,49994
0,68	0,25175	1,68	0,45352	3,90	0,49995
0,69	0,25490	1,69	0,45449	3,95	0,49996
0,70	0,25804	1,70	0,45543	4,00	0,49997
0,71	0,26115	1,71	0,45637	4,05	0,49997
0,72	0,26424	1,72	0,45728	4,10	0,49998
0,73	0,26730	1,73	0,45818	4,15	0,49998
0,74	0,27035	1,74	0,45907	4,20	0,49999
0,75	0,27337	1,75	0,45994	4,25	0,49999
0,76	0,27637	1,76	0,46080	4,30	0,49999
0,77	0,27935	1,77	0,46164	4,35	0,49999
0,78	0,28230	1,78	0,46246	4,40	0,49999
0,79	0,28524	1,79	0,46327	4,45	0,50000
0,80	0,28814	1,80	0,46407	4,50	0,50000
0,81	0,29103	1,81	0,46485	4,55	0,50000
0,82	0,29389	1,82	0,46562	4,60	0,50000
0,83	0,29673	1,83	0,46638	4,65	0,50000
0,84	0,29955	1,84	0,46712	4,70	0,50000
0,85	0,30234	1,85	0,46784	4,75	0,50000
0,86	0,30511	1,86	0,46856	4,80	0,50000
0,87	0,30875	1,87	0,46926	4,85	0,50000
0,88	0,31057	1,88	0,46995	4,90	0,50000
0,89	0,31327	1,89	0,47062	4,95	0,50000
0,90	0,31594	1,90	0,47128	5,00	0,50000

Таблица значений интегральной функции Лапласа (окончание)

$$\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_0^z e^{\frac{-t^2}{2}}dt$$

0,91	0,31859	1,91	0,47193		
0,92	0,32121	1,92	0,47257		
0,93	0,32381	1,93	0,47320		
0,94	0,32639	1,94	0,47381		
0,95	0,32894	1,95	0,47441		
0,96	0,33147	1,96	0,47500		
0,97	0,33398	1,97	0,47558		
0,98	0,33646	1,98	0,47615		
0,99	0,33891	1,99	0,47670		

ПРИЛОЖЕНИЕ 2

Критические точки распределения Стьюдента

Число степеней свободы *к*	**Уровень значимости α (двусторонняя критическая область)**					
	0,10	**0,05**	**0,02**	**0,01**	**0,002**	**0,001**
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	4,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,96
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,74
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,46	2,76	3,40	3,66
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,37
∞	1,64	1,96	2,33	2,58	3,09	3,29
	0,05	**0,025**	**0,01**	**0,005**	**0,001**	**0,0005**
	Уровень значимости α (односторонняя критическая область)					

Учебное издание

Задохина Нина Владимировна,
кандидат педагогических наук

Дубинина Наталья Михайловна,
кандидат юридических наук, доцент

Страхов Андрей Александрович

Цыганкова Яна Валерьевна,
кандидат юридических наук

Математика для психологов: элементы теории вероятностей и математической статистики

Редактор *Мирзоева Л. С.*
Корректор *Лосева О. С., Титова В. П.*
Компьютерная верстка *Мирзоева Л. С.*

Московский университет МВД России имени В.Я. Кикотя
117997, г. Москва, ул. Академика Волгина, д. 12

Подписано в печать 28.12.2021	Формат 60×84 1/16	Тираж	245 экз.
Заказ № 93	Цена договорная	Объем	2,34 уч.-изд. л.
			5,40 усл. печ. л.