МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
Рубцовский индустриальный институт (филиал)
федерального государственного бюджетного образовательного
учреждения высшего образования
«Алтайский государственный технический университет им. И.И. Ползунова»
(РИИ АлтГТУ)

И.И. КУЛЕШОВА

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Методическое пособие
для студентов всех форм обучения направления «Экономика»

Рубцовск 2021

УДК 517.9

Кулешова, И.И. Теория вероятностей и математическая статистика: Методическое пособие для студентов всех форм обучения направления «Экономика» /И.И. Кулешова //Рубцовский индустриальный институт. – Рубцовск, 2021.-83 с. [ЭР].

Учебное пособие содержит теоретический материал по теории вероятностей и математической статистике с достаточным количеством примеров, помогающих самостоятельно изучить рассмотренные темы. Там, где это возможно, раскрывается экономический смысл математических понятий, приводятся простейшие приложения элементов теории вероятностей и математической статистики в экономике.

Рассмотрено и одобрено на заседании кафедры ПМ Рубцовского индустриального института
Протокол № 10 от 28.04.2021г.

СОДЕРЖАНИЕ

1. ПРЕДМЕТ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Наблюдаемые нами события (явления) можно подразделить на следующие три вида: достоверные, невозможные и случайные.

Достоверным называют событие, которое обязательно произойдет, если будет осуществлена определенная совокупность условий S. Например, если в сосуде содержится вода при нормальном атмосферном давлении и температуре 20°, то событие «вода в сосуде находится в жидком состоянии» есть достоверное. В этом примере заданные атмосферное давление и температура воды составляют совокупность условий S.

Невозможным называют событие, которое заведомо не произойдет, если будет осуществлена совокупность условий S. Например, событие «вода в сосуде находится в твердом состоянии» заведомо не произойдет, если будет осуществлена совокупность условий предыдущего примера.

Случайным называют событие, которое при осуществлении совокупности условий S может либо произойти, либо не произойти. Например, если брошена монета, то она может упасть так, что сверху будет либо герб, либо надпись. Поэтому событие «при бросании монеты выпал «герб» — случайное. Каждое случайное событие, в частности выпадение «герба», есть следствие действия очень многих случайных причин (в нашем примере: сила, с которой брошена монета, форма монеты и многие другие). Невозможно учесть влияние на результат всех этих причин, поскольку число их очень велико и законы их действия неизвестны. Поэтому теория вероятностей не ставит перед собой задачу предсказать, произойдет единичное событие или нет, — она просто не в силах это сделать.

1.1. Виды случайных событий

События называют *несовместными,* если появление одного из них исключает появление других событий в одном и том же испытании.

Пример 1.1. Из ящика, с деталями (стандартными и нестандартными) наудачу извлечена деталь. Появление стандартной детали исключает появление нестандартной детали. События «появилась стандартная деталь» и «появилась нестандартная деталь» — несовместные.

Несколько событий образуют *полную группу,* если в результате испытания появится хотя бы одно из них. Другими словами, появление хотя бы одного из событий полной группы есть достоверное событие. В частности, *если события, образующие полную группу, попарно несовместны, то в результате испытания появится одно и только одно из этих событий.* Этот частный случай представляет для нас наибольший интерес, поскольку используется далее.

Пример 1.2. Приобретены два билета денежно-вещевой лотереи. Обязательно произойдет одно и только одно из следующих событий: «выигрыш выпал на первый билет и не выпал на второй», «выигрыш не выпал на первый билет и выпал на второй», «выигрыш выпал на оба билета», «на оба билета выигрыш не выпал». Эти события образуют полную группу попарно несовместных событий.

События называют *равновозможными,* если есть основания считать, что ни одно из них не является более возможным, чем другое.

Пример 1.3. Появление «герба» и появление надписи при бросании монеты — равновозможные события. Действительно, предполагается, что монета изготовлена из однородного материала, имеет правильную цилиндрическую форму, и наличие чеканки не оказывает влияния на выпадение той или иной стороны монеты.

1.2. Классическое определение вероятности

Вероятность — одно из основных понятий теории вероятностей. Существует несколько определений этого понятия. Приведем определение, которое называют классическим. Далее укажем слабые стороны этого определения и приведем другие определения, позволяющие преодолеть недостатки классического определения.

Рассмотрим пример. Пусть в урне содержится 6 одинаковых, тщательно перемешанных шаров, причем 2 из них — красные, 3 — синие и 1 — белый. Очевидно, возможность вынуть наудачу из урны цветной (т. е. красный или синий) шар больше, чем возможность извлечь белый шар. Можно ли охарактеризовать эту возможность числом? Оказывается, можно. Это число и называют вероятностью события (появления цветного шара). Таким образом, вероятность есть число, характеризующее степень возможности появления события.

Поставим перед собой задачу дать количественную оценку возможности того, что взятый наудачу шар цветной. Появление цветного шара будем рассматривать в качестве события *A*. Каждый из возможных результатов испытания (испытание состоит в извлечении шара из урны) назовем *элементарным исходом (элементарным событием)*. Элементарные исходы обозначим через $\omega_1,\ \omega_2,\ \omega_3$ и т. д. В нашем примере возможны следующие 6 элементарных исходов: ω_1 — появился белый шар; $\omega_2,\ \omega_3$ — появился красный шар; $\omega_4,\ \omega_5,\ \omega_6$ — появился синий шар. Легко видеть, что эти исходы образуют полную группу попарно несовместных событий (обязательно появится только один шар) и они равновозможны (шар вынимают наудачу, шары одинаковы и тщательно перемешаны).

Те элементарные исходы, в которых интересующее нас событие наступает, назовем *благоприятствующими* этому событию. В нашем примере благоприятствуют событию *A* (появлению цветного шара) следующие 5 исходов: $\omega_2,\ \omega_3,\ \omega_4,\ \omega_5,\ \omega_6$.

Таким образом, событие *A* наблюдается, если в испытании наступает один, безразлично какой, из элементарных исходов, благоприятствующих *A;* в нашем примере *A* наблюдается, если наступит ω_2, или ω_3, или ω_4, или ω_5, или ω_6. В этом смысле событие *A* подразделяется на несколько элементарных событий

(ω_2, ω_3, ω_4, ω_5, ω_6); элементарное же событие не подразделяется на другие события. В этом состоит различие между событием A и элементарным событием (элементарным исходом).

Отношение числа благоприятствующих событию A элементарных исходов к их общему числу называют вероятностью события A и обозначают через $P(A)$. В рассматриваемом примере всего элементарных исходов 6; из них 5 благоприятствуют событию A. Следовательно, вероятность того, что взятый шар окажется цветным, равна $P(A) = 5/6$. Это число и дает ту количественную оценку степени возможности появления цветного шара, которую мы хотели найти. Дадим теперь определение вероятности.

Вероятностью события A называют отношение числа благоприятствующих этому событию исходов к общему числу всех равновозможных несовместных элементарных исходов, образующих полную группу. Итак, вероятность события A определяется формулой

$$P(A) = m/n,$$

где m — число элементарных исходов, благоприятствующих A;

n — число всех возможных элементарных исходов испытания.

Здесь предполагается, что элементарные исходы несовместны, равновозможны и образуют полную группу. Из определения вероятности вытекают следующие ее свойства:

Свойство 1. *Вероятность достоверного события равна единице.*

Действительно, если событие достоверно, то каждый элементарный исход испытания благоприятствует событию. В этом случае $m = n$, следовательно,

$$P(A) = m/n = n/n = 1.$$

Свойство 2. *Вероятность невозможного события равно нулю.*

Действительно, если событие невозможно, то ни один из элементарных исходов испытания не благоприятствует событию. В этом случае $m = 0$, следовательно,

$$P(A) = m/n = 0/n = 0.$$

Свойство 3. *Вероятность случайного события есть положительное число, заключенное между нулем и единицей.*

Действительно, случайному событию благоприятствует лишь часть из общего числа элементарных исходов испытания. В этом случае $0<m<n$, значит, $0<m/n<1$, следовательно,

$$0 < P(A) < 1.$$

Итак, вероятность любого события удовлетворяет двойному неравенству

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

Далее приведены теоремы, которые позволяют по известным вероятностям одних событий находить вероятности других событий.

1.3. Ограниченность классического определения вероятности. Статистическая вероятность

Классическое определение вероятности предполагает, что число элементарных исходов испытания конечно. На практике же весьма часто встречаются испытания, число возможных исходов которых бесконечно. В таких случаях классическое определение неприменимо. Уже это обстоятельство указывает на ограниченность классического определения. Отмеченный недостаток может быть преодолен, в частности, введением геометрических вероятностей и, конечно, использованием аксиоматической вероятности.

Наиболее слабая сторона классического определения состоит в том, что очень часто невозможно представить результат испытания в виде совокупности элементарных событий. Еще труднее указать основания, позволяющие считать элементарные события равновозможными. Обычно о равновозможности элементарных исходов испытания говорят из соображений симметрии. Так, например, предполагают, что игральная кость имеет форму правильного многогранника (куба) и изготовлена из однородного материала. Однако задачи, в которых можно исходить из соображений симметрии, на практике встречаются весьма редко. По этой причине наряду с классическим определением вероятности используют и другие определения, в частности статистическое определение: *в качестве статистической вероятности события принимают относительную частоту или число, близкое к ней.* Например, если в результате достаточно большого числа испытаний оказалось, что относительная частота весьма близка к числу 0,4, то это число можно принять за статистическую вероятность события.

Легко проверить, что свойства вероятности, вытекающие из классического определения, сохраняются и при статистическом определении вероятности. Действительно, если событие достоверно, то $m = n$ и относительная частота

$$m/n = n/n = 1,$$

т. е. статистическая вероятность достоверного события (так же как и в случае классического определения) равна единице.

Если событие невозможно, то $m=0$ и, следовательно, относительная частота

$$0/n===0,$$

т. е. статистическая вероятность невозможного события равна нулю.

Для любого события $0 \le m \le n$, и, следовательно, относительная частота

$$0 \le m/n \le 1,$$

т. е. статистическая вероятность любого события заключена между нулем и единицей.

Для существования статистической вероятности события A требуется:

а) возможность, хотя бы принципиально, производить неограниченное число испытаний, в каждом из которых событие A наступает или не наступает;

б) устойчивость относительных частот появления A в различных сериях достаточно большого числа испытаний.

1.4. Комбинаторика, ее основные понятия и правила

1.4.1. Понятие комбинаторики

Комбинаторика — это раздел математики, изучающий вопросы о том, сколько комбинаций (соединений) определенного вида можно составить из данных элементов.

Комбинаторика широко применяется в математической логике, теории чисел, теории массового обслуживания, теории управляющих систем и вычислительной технике, а также во многих других разделах науки и техники. Особое место она занимает в теории вероятностей.

При решении ряда вероятностных задач успешно используются два основных правила комбинаторики (правило сложения и правило умножения), а также применяются ее основные понятия — размещения, перестановки и сочетания — и их число.

Установим прежде всего главные правила, на которых базируются многие утверждения комбинаторики.

Основными правилами комбинаторики являются правило сложения (суммы) и правило умножения (произведения). Рассмотрим вначале правило умножения на примере следующей задачи.

1.4.2. Основные правила комбинаторики

Правило умножения. Предположим, что требуется выполнить одно за другим k действий. Если первое действие можно выполнить n_1 способами, второе — n_2 способами, третье — n_3 способами и так до *k-го* действия, которое может быть выполнено n_k способами, то все k действий вместе могут быть выполнены числом способов, равным $n_1 \cdot n_2 \cdot n_3 \cdot ... \cdot n_k$.

Пример 1.4. В первой группе учится 20 человек, во второй - 25 человек. Для участия в конференции необходимо выбрать по одному человеку из каждой группы. Сколькими способами это можно сделать?

Решение. Из 20 учащихся первой группы может быть выбран любой учащийся, т.е. первое действие может быть совершено 20 способами. Аналогично из 25 учащихся второй группы также может быть выбран любой человек, т.е. второе действие может быть выполнено 25 способами. Следовательно, по правилу умножения оба действия — выбор из первой и второй групп по одному человеку — могут быть осуществлены $20 \cdot 25 = 500$ способами.

Изменим условие примера 1 и посмотрим, каким окажется его решение.

Пример 1.5. В первой группе учится 20 человек, во второй — 25 человек. Произвольно выбирают одного человека из какой-то группы. Сколькими разными способами это можно сделать?

Решение. Из первой группы человека можно выбрать 20 способами, из второй группы — 25 способами. Всего способов 20 + 25 = 45.

При решении примера 2 существенным оказывается то, что оба действия (выбор человека из первой и выбор человека из второй группы) не могут быть выполнены одновременно, поскольку они взаимно исключают друг друга. В данном случае должно быть выполнено либо первое действие, либо второе, а не первое, а затем второе. Рассуждения при решении примера 2 лежат в основе правила сложения.

Правило сложения. Если два действия взаимно исключают друг друга, причем одно из них может быть выполнено m_1 способами, а другое — m_2 способами, то выполнить одно любое из этих действий можно $m_1 + m_2$ способами.

Это правило распространяется на любое конечное число действий.

1.4.3. Основные элементы комбинаторики

Комбинации элементов могут быть составлены различными способами: в виде размещений, перестановок и сочетаний. Эти три вида соединений (комбинаций) являются основными понятиями комбинаторики. Прежде чем перейти к рассмотрению этих понятий, познакомимся с понятием факториала натурального числа.

Факториалом натурального числа *n* называется произведение всех натуральных чисел до *n* включительно, т.е.

$$n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n.$$

Например,

$$2! = 1 \cdot 2 = 2$$

$$3! = 1 \cdot 2 \cdot 3 = 6$$

$$4! = 1 \cdot 2 \cdot 3 \cdot 4 = 24 \text{ и т.д.}$$

По определению принято, что $0! = 1$. Пусть имеется три различных элемента *a, b, c.* Будем составлять из этих элементов различные комбинации.

Выберем из трех элементов *a, b, c* два элемента с учетом их расположения. Получим

$$a\,b,\ b\,a,\ b\,c,\ c\,b,\ a\,c,\ c\,a.$$

Составленные комбинации называются размещениями из трех элементов по два.

Размещениями из *n* элементов по *m* в каждом называются такие комбинации из *m* элементов, взятых из *n* элементов, которые отличаются друг от друга либо самими элементами, либо порядком их расположения. Например, размещения *a b* и *b a* отличаются порядком расположения элементов, а размещения *a b* и *c b* — элементами *a* и *c.*

Очевидно, что $m < n$.

На практике чаще представляют интерес не сами размещения, а их число. Для того чтобы понять формулу для числа размещений, рассмотрим следующий пример.

Пример 1.6. Руководство некоторого банка должно выбрать трех человек из 12 на три различные должности. Все 12 кандидатов имеют равные шансы на занятие той или иной должности. Сколько существует всевозможных вариантов для выбора?

Решение. Используем правило умножения. Первого человека нужно выбрать из 12 кандидатов, следовательно, существует 12 способов выбора, второго человека выбирают из оставшихся 11, поэтому для осуществления второго действия имеется 11 способов. И, наконец, третьего человека выбираем из оставшихся десяти. Очевидно, что число способов осуществления третьего действия равно 10. По правилу умножения можно составить $12 \cdot 11 \cdot 10 = 1320$ различных вариантов по три человека в каждом.

С другой стороны, число всевозможных вариантов выбора равно числу размещений из 12 человек по три. Очевидно, что комбинации, состоящие из трех человек, будут размещениями, так как они будут отличаться либо самими людьми, либо расположением этих людей соответственно их должностям (по условию примера — должности разные). Таким образом, число размещений из 12 человек по 3 равно $12 \cdot 11 \cdot 10 = 1320$.

Пример 1.7. Сколько можно записать трехзначных чисел, используя без повторения все девять (кроме нуля) цифр?

Решение. Необходимо составить различные размещения из девяти цифр по три и найти их число. Итак,

$$A_9^3 = \frac{9!}{(9-3)!} = \frac{9!}{6!} = 7 \cdot 8 \cdot 9 = 504.$$

Таким образом, можно составить 504 различных числа.

Изменим условие примера 4 и предположим, что при составлении чисел цифры могут повторяться, например, будем считать, что возможны числа 121, 144, 555 и т.д. Понятно, что в этом случае количество различных чисел будет значительно больше. Новые полученные комбинации чисел будут называться размещениями с повторениями.

Размещениями с повторениями из n элементов по m называются размещения из n элементов по m, которые могут содержать любой из n элементов и любое (не большее m) число раз.

Таким образом, каждое размещение с повторением может состоять не только из различных m элементов, но и из m каких угодно и сколь угодно повторяющихся элементов.

Число размещений с повторениями из n элементов по m вычисляется по следующей формуле:

$$A^m_{n(\text{повт}))} = n^m. \tag{1.1}$$

Пример 1.8. Сколько можно записать трехзначных чисел, используя все девять (кроме нуля) цифр? Цифры могут повторяться.

Решение. Число размещений с повторениями из девяти цифр по три равно

$$A^3_{9(\text{повт})} = 9^3 = 729.$$

Таким образом, можно составить 729 различных чисел.

Рассмотрим пример 4 с дополнительным условием.

Пример 1.9. Руководство некоторого банка должно выбрать трех человек из 12 на три различные должности. Все 12 кандидатов имеют равные шансы на занятие той или иной должности, при этом допускается занятие одним человеком не только одной, но и двух и даже трех должностей. Сколько существует всевозможных вариантов для выбора?

Решение. Как и в примере 1.4, речь идет о размещениях. Однако в этом случае допускается занятие одним человеком нескольких должностей, поэтому здесь используются размещения с повторениями. Таким образом,

$$A^3_{12(\text{повт})} = 12^3 = 1728,$$

т.е. существует 1728 вариантов выбора.

В условии всех предыдущих примеров предполагалось, что $m < n$. Рассмотрим отдельно случай, когда $m = n$. Размещения, соответствующие этому случаю, называются перестановками.

Выберем из трех элементов *a, b, c* три элемента с учетом их расположения. Получим

$$a\ b\ c,\ b\ a\ c,\ b\ c\ a,\ c\ b\ a,\ a\ c\ b,\ c\ a\ b.$$

Составленные комбинации называются перестановками из трех элементов.

Перестановками из *n* элементов называются комбинации**,** которые состоят из всех n элементов и отличаются лишь порядком расположения этих элементов. Иначе, перестановки — это размещения из n элементов по n. Нетрудно получить формулу для числа перестановок:

$$P_n = n! \qquad (1.2)$$

Пример 1.10. Представитель торговой фирмы ежедневно просматривает 5 изданий, в которых исследуются спрос и предложение на определенные товары. Сколько существует способов просмотра, если выбор изданий случаен?

Решение. Каждый день все 5 изданий должны быть просмотрены, может меняться лишь порядок просмотра, поэтому число способов просмотра равно $P_5 = 5! = 120$.

Пример 1.11. На полке стоят 7 учебников, 3 из которых по эконометрике. Сколько существует способов расстановки семи различных книг на полке так, чтобы 3 учебника по эконометрике стояли рядом?

Решение. Зафиксируем расположение трех учебников по эконометрике. Тогда остальные четыре учебника плюс связка из 3 учебников можно расставить на полке $P_5 = 5! = 120$ различными способами. В свою очередь учебники по эконометрике также можно переставить между собой $P_3 = 3! = 6$ способами. По правилу умножения число способов расстановки для семи книг равно $P_5 \cdot P_3 = 120 \cdot 6 = 720$.

В рассмотренных ранее примерах предполагалось, что все элементы в одной перестановке различны. Если допустить наличие одинаковых элементов в перестановке, то получим перестановки с повторениями. Как известно, перестановки — это частный случай размещений, в связи с чем, учитывая формулу (1.1), можно получить следующую формулу для числа перестановок с повторениями

$$P_{n(\text{повт}))} = A^n_{n(\text{повт}))} = n^n \qquad (1.3)$$

Пример 1.12. Сколько различных чисел можно составить из трех цифр 3, 4, 5?

Решение. Применяя формулу (1.3), получим $P_{3(\text{повт})} = 3^3 = 27$. Решим эту же задачу другим способом, без применения формулы (1.3).

Числа, состоящие из трех предложенных цифр, можно разделить на три группы. В первую группу входят числа, в которых все цифры разные, например, 345, 534 и т.д. Очевидно, что их число равно 3!=6. Вторая группа состоит из чисел, в которых имеется две одинаковых цифры, например, с двумя цифрами 3 существует 6 различных чисел: 334, 343, 433, 335, 353, 533. Аналогично существует по 6 различных чисел с двумя цифрами 4 и 5. Следовательно, вторая группа состоит из 18 чисел. Наконец, третья группа составлена из чисел, в которых все три цифры одинаковы, т.е. 333, 444, 555. Таким образом, общее количество чисел равно 6 + 18 + 3 = 27.

Для решения ряда задач бывает полезна формула, позволяющая найти число различных комбинаций из n элементов, среди которых есть конкретное число одинаковых элементов.

Пусть среди n элементов есть n_1 элементов одного вида, n_2 элементов второго вида и т.д. и n_k элементов k-го вида, тогда число перестановок из этих чисел равно:

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

где $n_1 + n_2 + \ldots + n_k = n$.

Пример 1.13. Сколько существует различных шестизначных чисел, в которых цифра 1 повторяется два раза, цифра 3 — один раз, цифра 5 — три раза?

Решение. Из условия задачи ясно, что n = 6, n_1 = 2, n_2 = 1, n_3 = 3ю По формуле (1.3) получим $P_6(2,\ 1,\ 3) = \frac{6!}{2!\cdot 1!\cdot 3!} = 60$.

Рассмотрим два различных размещения, которые состоят из одних и тех же элементов. Тогда они обязательно должны отличаться порядком расположения этих элементов. Однако часто бывает, что нет необходимости учитывать этот порядок, т.е. размещения, которые отличаются лишь расположением элементов, считать равными. В этом случае полученные комбинации будут называться *сочетаниями.*

Выберем из трех элементов *a, b, c* два элемента без учета их расположения. Получим

a b, b c, a c.

Составленные комбинации называются сочетаниями из трех элементов по два.

Сочетаниями из *n* элементов по *m* в каждом называются такие комбинации *m* элементов, взятых из *n* элементов, которые отличаются друг от друга по крайней мере одним элементом.

Очевидно, что $m \leq n$.

Число сочетаний из *n* элементов по *m* определяется по следующей формуле:

$$C_n^m = \frac{n!}{m!(n-m)!}, \tag{1.5}$$

где $0 \leq m \leq n$.

Числа C_n называются **биномиальными коэффициентами,** так как они входят в разложение формулы бинома Ньютона:

$$(a+b)^n = \sum_{m=0}^{n} C_n^m a^m b^{n-m} = b^n + C_n^1 a^1 b^{n-1} +$$
$$+ C_n^2 a^2 b^{n-2} + \ldots + C_n^m a^m b^{n-m} + \ldots + a^n.$$

Числа C_n^m обладают следующими свойствами:

1) $C_n^0 = C_n^n = 1,\ C_n^1 = n$;

2) $\sum_{m=0}^{n} C_n^m = 2^n$;

3) $C_n^m = C_n^{n-m}$, где $m = 0,\ 1,\ \ldots,\ n$, удобно применять ри $m > \frac{n}{2}$;

4) $C_n^m = C_{n-1}^{m-1} + C_{n-1}^m$, где $m = 0,\ 1,\ \ldots,\ n-1$.

Пример 1.14. Решить пример 3, предполагая, что кандидаты выбираются на одинаковые должности.

Решение. В виду того, что должности одинаковые, порядок выбора кандидатов значения не имеет, поэтому при решении применяем формулу (1.4) для числа сочетаний и получаем

$$C_{12}^3 = \frac{12!}{3!(12-3)!} = \frac{12!}{3!9!} = \frac{10 \cdot 11 \cdot 12}{1 \cdot 2 \cdot 3} = 220.$$

Изменим условие примера 11, предполагая, что один человек может занимать не только одну, но и две или даже три должности, однако должности остаются одинаковыми. Тогда появляются комбинации, называемые сочетаниями с повторениями.

Сочетания с повторениями из *n* элементов по *m* могут состоять не только из *m* различных элементов, но и из *m* каких угодно и сколько угодно раз повторяющихся элементов, взятых из данного множества *n* элементов.

Число сочетаний с повторениями из *n* элементов по *m* может быть вычислено по формуле

$$C_{n(\text{повт})}^m = C_{n+m-1}^m = \frac{(n+m-1)!}{m!(n-1)!}. \tag{1.6}$$

Необходимо обратить внимание на то, что в формуле (1.6) число m может быть и больше n. Например, число способов выбора 6 авторучек, если в продаже имеется 4 вида авторучек, может быть определено по формуле

$$C_{4(\text{повт})}^{6} = C_{4+6-1}^{6} = C_{9}^{6} = \frac{9!}{6!\cdot 3!} = 84.$$

Очевидно, что в этом примере $m = 6$, а $n = 4$.

2. ТЕОРЕМЫ СЛОЖЕНИЯ И УМНОЖЕНИЯ ВЕРОЯТНОСТЕЙ

2.1. Теорема сложения вероятностей несовместных событий

Суммой A+B двух событий A и B называют событие, состоящее в появлении события *A,* или события *B,* или обоих этих событий. Например, если из орудия произведены два выстрела и *A* — попадание при первом выстреле, *B* — попадание при втором выстреле, то *A+B* — попадание при первом выстреле, или при втором, или в обоих выстрелах.

В частности, если два события *A* и *B* — несовместные, то *A+B* — событие, состоящее в появлении одного из этих событий, безразлично какого.

Суммой нескольких событий называют событие, которое состоит в появлении хотя бы одного из этих событий. Например, событие *A+B+C* состоит в появлении одного из следующих событий: *A, B, C, A* и *B, A* и *C, B* и *C, A* и *B* и *C*.

Пусть события *A* и *B* — несовместные, причем вероятности этих событий известны. Как найти вероятность того, что наступит либо событие *A,* либо событие *B?* Ответ на этот вопрос дает теорема сложения.

Теорема. *Вероятность появления одного из двух несовместных событий, безразлично какого, равна сумме вероятностей этих событий:*

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

Доказательство. Введем обозначения: n — общее число возможных элементарных исходов испытания; m_1 — число исходов, благоприятствующих событию *A;* m_2 — число исходов, благоприятствующих событию *B.*

Число элементарных исходов, благоприятствующих наступлению либо события *A*, либо события *B,* равно $m_1 + m_2$. Следовательно,

$$P(A+B) = (m_1 + m_2)/n = m_1/n + m_2/n.$$

Приняв во внимание, что $m_1/n = P(A)$ и $m_2/n = P(B)$ — окончательно получим

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

Следствие. *Вероятность появления одного из нескольких попарно несовместных событий, безразлично какого, равна сумме вероятностей этих событий:*

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

Доказательство. Рассмотрим три события: *A, B* и С. Так как рассматриваемые события попарно несовместны, то появление одного из трех событий, *A, B* и С, равносильно наступлению одного из двух событий, *A*+*B* и *C,* поэтому в силу указанной теоремы

$$P(A+B+C)=P\left[(A+B)+C\right]=P(A+B)+P(C)=$$
$$=P(A)+P(B)+P(C).$$

Для произвольного числа попарно несовместных событий доказательство проводится методом математической индукции.

Пример 2.1. В урне 30 шаров: 10 красных, 5 синих и 15 белых. Найти вероятность появления цветного шара.

Решение. Появление цветного шара означает появление либо красного, либо синего шара.

Вероятность появления красного шара (событие *A*).

$$P(A) = 10/30 = 1/3.$$

Вероятность появления синего шара (событие *B*).

$$P(B) = 5/30 = 1/6.$$

События *A* и *B* несовместны (появление шара одного цвета исключает появление шара другого цвета), поэтому теорема сложения применима.

Искомая вероятность

$$P(A+B)=P(A)+P(B)=1/3+1/6=1/2.$$

Пример 2.2. Стрелок стреляет по мишени, разделенной на 3 области. Вероятность попадания в первую область равна 0,45, во вторую — 0,35. Найти вероятность того, что стрелок при одном выстреле попадет либо в первую, либо во вторую область.

Решение. События *A* — «стрелок попал в первую область» и *B* — «стрелок попал во вторую область» — несовместны (попадание в одну область исключает попадание в другую), поэтому теорема сложения применима.

Искомая вероятность

$$P(A+B)=P(A)+P(B)=0,45+0,35=0,80.$$

Полная группа событий

Теорема. *Сумма вероятностей событий* A_1 A_2, ..., *A*,,, *образующих полную группу, равна единице:*

$$P(A_1)+P(A_2)+...+P(A_n)=1.$$

Доказательство. Так как появление одного из событий полной группы достоверно, а вероятность достоверного события равна единице, то

$$P(A_1+A_2+...+A_n)=1. \qquad (2.1)$$

Любые два события полной группы несовместны, поэтому можно применить теорему сложения:

$$P(A_1+A_2+...+A_n)=P(A_1)+P(A_2)+...+P(A_n). \qquad (2.2)$$

Сравнивая (2.1) и (2.2), получим

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

Пример 2.3. Консультационный пункт института получает пакеты с контрольными работами из городов *A, B* и *C.* Вероятность получения пакета из города *A* равна 0,7, из города *B* — 0,2. Найти вероятность того, что очередной пакет будет получен из города *C.*

Решение. События «пакет получен из города *A*», «пакет получен из города *B*», «пакет получен из города *C*» образуют полную группу, поэтому сумма вероятностей этих событий равна единице:

$$0{,}7 + 0{,}2 + p=1.$$

Отсюда искомая вероятность

$$p = 1-0{,}9=0{,}1.$$

Противоположные события

Противоположными называют два единственно возможных события, Образующих полную группу. Если одно из двух противоположных событий обозначено через *A*, то другое принято обозначать $\overline{A}$.

Пример 2.4. Попадание и промах при выстреле по цели — противоположные события. Если *A* — попадание, то $\overline{A}$ — промах.

Пример 2.5. Из ящика наудачу взята деталь. События «появилась стандартная деталь» и «появилась нестандартная деталь» — противоположные.

Теорема. *Сумма вероятностей противоположных событий равна единице:*

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

Доказательство. Противоположные события образуют полную группу, а сумма вероятностей событий, образующих полную группу, равна единице.

Замечание 1. Если вероятность одного из двух противоположных событий обозначена через *p,* то вероятность другого события обозначают через q. Таким образом, в силу предыдущей теоремы

$$p+q=1.$$

Пример 2.6. Вероятность того, что день будет дождливым, *p*=0,7. Найти вероятность того, что день будет ясным.

Решение. События «день дождливый» и «день ясный» — противоположные, поэтому искомая вероятность

$$q=1-p=1-0{,}7=0{,}3.$$

Замечание 2. При решении задач на отыскание вероятности события *A* часто выгодно сначала вероятность события *A,* а затем найти искомую вероятность по формуле

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

2.2. Теорема умножения вероятностей

Произведение событий

Произведением двух событий A и B называют событие AB, состоящее в совместном появлении (совмещении) этих событий. Например, если A — деталь годная, B — деталь окрашенная, то AB — деталь годна и окрашена. *Произведением нескольких событий* называют событие, состоящее в совместном появлении всех этих событий. Например, если A, B, C — появление «герба» соответственно в первом, втором и третьем бросаниях монеты, то ABC — выпадение «герба» во всех трех испытаниях.

Условная вероятность

Случайное событие определено как событие, которое при осуществлении совокупности условий S может произойти или не произойти. Если при вычислении вероятности события никаких других ограничений, кроме условий S, не налагается, то такую вероятность называют *безусловной;* если же налагаются и другие дополнительные условия, то вероятность события называют *условной.* Например, часто вычисляют вероятность события B при дополнительном условии, что произошло событие A. Заметим, что и безусловная вероятность, строго говоря, является условной, поскольку предполагается осуществление условий S.

Условной вероятностью $P_A(B)$ называют вероятность события B, вычисленную в предположении, что событие A уже наступило.

Пример 2.7. В урне 3 белых и 3 черных шара. Из урны дважды вынимают по одному шару, не возвращая их обратно. Найти вероятность появления белого шара при втором испытании (событие B), если при первом испытании был извлечен черный шар (событие A).

Решение. После первого испытайия в урне осталось 5 шаров, из них 3 белых. Искомая условная вероятность

$$P_A(B) = 3/5.$$

Этот же результат можно получить по формуле

$$P_A(B)=P(AB)/P(A) \qquad (P(A)>0). \tag{2.3}$$

Действительно, вероятность появления белого шара при первом испытании

$$P(A) = 3/6 = 1/2.$$

Найдем вероятность $P(AB)$ того, что в первом испытании появится черный шар, а во втором — белый. Общее число исходов — совместного появления двух шаров, безразлично какого цвета, равно числу размещений $A_6^2 = 6 \cdot 5 = 30$. Из этого числа исходов событию AB благоприятствуют $3 \cdot 3 = 9$ исходов. Следовательно,

$$P(AB) = 9/30 = 3/10.$$

Искомая условная вероятность

$$P_A(B) = P(AB)/P(A) = (3/10)/(1/2) = 3/5.$$

Как видим, получен прежний результат.

Исходя из классического определения вероятности, формулу (2.3) можно доказать. Это обстоятельство и служит основанием для следующего общего (применимого не только для классической вероятности) определения.

Условная вероятность события B при условии, что событие A уже наступило, по определению, равна

$$P_A(B) = P(AB)/P(A) \qquad (P(A) > 0).$$

Рассмотрим два события: A и B; пусть вероятности $P(A)$ и $P_A(B)$ известны. Как найти вероятность совмещения этих событий, т. е. вероятность того, что появится и событие A и событие В? Ответ на этот вопрос дает теорема умножения.

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

$$P(AB) = P(A)P_A(B).$$

Доказательство. По определению условной вероятности, по формуле (2.3)

$$P_A(B) = P(AB)/P(A).$$

Отсюда

$$P(AB) = P(A)P_A(B).$$

Замечание. Применив данную формулу к событию BA, получим

$$P(BA)=P(B)P_B(A),$$

или, поскольку событие BA не отличается от события AB,

$$P(AB)=P(B)P_B(A).$$

(**)

Следовательно,

$$P(A)P_A(B)=P(B)P_B(A). \qquad (2.5)$$

Следствие. *Вероятность совместного появления нескольких событий равна произведению вероятности одного из них на условные вероятности всех остальных, причем, вероятность каждого последующего события вычисляется в предположении, что все предыдущие события уже появились:*

$$P(A_1A_2A_3...A_n) = P(A_1)P_{A_1}(A_2)P_{A_1A_2}(A_3)...$$
$$...P_{A_1A_2...A_{n-1}}(A_n),$$

где $P_{A_1A_2...A_{n-1}}(A_n)$ — вероятность события A_n, вычисленная в предположении, что события $A_1, A_2, ..., A_{n-1}$ наступили. В частности, для трех событий

$$P(ABC) = P(A)P_A(B)P_{AB}(C).$$

Заметим, что порядок, в котором расположены события, может быть выбран любым, т. е. безразлично какое событие считать первым, вторым и т.д.

Пример 2.8. У сборщика имеется 3 конусных и 7 эллиптических валиков. Сборщик взял один валик, а затем второй. Найти вероятность того, что первый из взятых валиков — конусный, а второй — эллиптический.

Решение. Вероятность того, что первый валик окажется конусным (событие A),

$$P(A) = 3/10.$$

Вероятность того, что второй валик окажется эллиптическим (событие B), вычисленная в предположении, что первый валик — конусный, т. е. условная вероятность

$$P_A(B) = 7/9.$$

По теореме умножения, искомая вероятность

$$P(AB) = P(A)\, P_A(B) = (3/10) \cdot (7/9) = 7/30.$$

Заметим, что, сохранив обозначения, легко найдем: $P(B) = 7/10$, $P_в(A)=3/9$, $P(B)P_B(A)=7/30$, что наглядно иллюстрирует справедливость равенства (2.5).

Пример 2.9. В урне 5 белых, 4 черных и 3 синих шара. Каждое испытание состоит в том, что наудачу извлекают один шар, не возвращая его обратно. Найти вероятность того, что при первом испытании появится белый шар (событие A), при втором — черный (событие B) и при третьем — синий (событие С).

Решение. Вероятность появления белого шара в первом испытании

$$P(A) = 5/12.$$

Вероятность появления черного шара во втором испытании, вычисленная в предположении, что в первом испытании появился белый шар, т. е. условная вероятность

$$P_A(B) = 4/11.$$

Вероятность появления синего шара в третьем испытании, вычисленная в предположении, что в первом испытании появился белый шар, а во втором — черный, т. е. условная вероятность

$$P_{AB}(C) = 3/10.$$

Искомая вероятность

$$P(ABC)=P(A)P_A(B)P_{AB}(C)=(5/12)\cdot(4/11)\cdot(3/10)=1/22.$$

Независимые события. Теорема умножения для независимых событий

Пусть вероятность события B не зависит от появления события A.

Событие В называют независимым от события А, если появление события A не изменяет вероятности события B, т. е. если условная вероятность события B равна его безусловной вероятности:

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

Отсюда

$$P_B(A)=P(A),$$

т. е. условная вероятность события A в предположении, что наступило событие B, равна его безусловной вероятности. Другими словами, событие A не зависит от события B.

Итак, если событие B не зависит от события А, то и событие A не зависит от события B; это означает, что свойство независимости событий взаимно.

Для независимых событий теорема умножения $P(AB) = P(A)\ P_A(B)$ имеет вид

$$P(AB)=P(A)P(B) \tag{2.6}$$

т. е. вероятность совместного появления двух независимых событий равна произведению вероятностей этих событий.

Равенство (2.4) принимают в качестве определения независимых событий.

Два события называют независимыми, если вероятность их совмещения равна произведению вероятностей этих событий; в противном случае события называют *зависимыми*.

На практике о независимости событий заключают по смыслу задачи. Например, вероятности поражения цели каждым из двух орудий не зависят от того, поразило ли цель другое орудие, поэтому события «первое орудие поразило цель» и «второе орудие поразило цель» независимы.

Пример 2.10. Найти вероятность совместного поражения цели двумя орудиями, если вероятность поражения цели первый орудием (событие A) равна 0,8, а вторым (событие B) — 0,7.

Решение. События A и В независимые, поэтому, по теореме умножения, искомая вероятность

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

Замечание 1. Если события A и B независимы, то независимы также события A и $\bar{B}$, $\bar{A}$ и B, $\bar{A}$ и $\bar{B}$. Действительно,

$$A = A\bar{B} + AB.$$

Следовательно,

$$P(A) = P\left(A\bar{B}\right) + P(AB), \text{ или } P(A) = P\left(A\bar{B}\right) + P(A)P(B).$$

Отсюда

$$P\left(A\bar{B}\right) = P(A)\left[1 - P(B)\right], \text{ или } P\left(A\bar{B}\right) = P(A)P\left(\bar{B}\right).$$

т. е. события A и B независимы.

Независимость событий $\bar{A}$ и B, $\bar{A}$ и $\bar{B}$ — следствие доказанного утверждения.

Несколько событий называют попарно независимыми, если каждые два из них независимы. Например, события A, B, C попарно независимы, если независимы события A и B, A и C, B и C.

Для того чтобы обобщить теорему умножения на несколько событий, введем, понятие независимости событий в совокупности.

Несколько событий называют независимыми в совокупности (или просто *независимыми),* если независимы каждые два из них и независимы каждое событие и все возможные произведения остальных. Например, если события A_1, A_2, A_3 независимы в совокупности, то независимы события A_1 и A_2, A_1 и A_3, A_2 и A_3; A_1 и A_2A_3, A_2 и A_1A_3, A_3 и A_1A_2. Из сказанного следует, что если события независимы в совокупности, то условная вероятность появления любого события из них, вычисленная в предположении, что наступили какие-либо другие события из числа остальных, равна его безусловной вероятности.

Подчеркнем, что если несколько событий независимы попарно, то отсюда еще не следует их независимость в совокупности. В этом смысле требование независимости событий в совокупности сильнее требования их попарной независимости.

Поясним сказанное на примере. Пусть в урне имеется 4 шара, окрашенные: один — в красный цвет (A), один — в синий цвет (B), один — в черный цвет (C) и один — во все эти три цвета (ABC). Чему равна вероятность того, что извлеченный из урны шар имеет красный, цвет?

Так как из четырех шаров два имеют красный цвет, то $P(A)=2/4=1/2$. Рассуждая аналогично, найдем $P(B) = 1/2$, $P(C)=1/2$. Допустим теперь, что взятый шар имеет синий цвет, т. е. событие B уже произошло. Изменится ли вероятность того, что извлеченный шар имеет красный цвет, т. е. изменится ли вероятность события A? Из двух шаров, имеющих синий цвет, один шар имеет и красный цвет, поэтому вероятность события A по-прежнему равна 1/2. Другими словами, условная вероятность события A, вычисленная в предположении, что наступило событие B, равна его безусловной вероятности. Следовательно, события A и B независимы. Аналогично придем к выводу, что события A и С, B и C независимы. Итак, события A, B и C попарно независимы.

Независимы ли эти события в совокупности? Оказывается, нет. Действительно, пусть извлеченный шар имеет два цвета, например синий и черный. Чему равна вероятность того, что этот шар имеет и красный цвет? Лишь один шар окрашен во все три цвета, поэтому взятый шар имеет и красный цвет. Таким образом, допустив, что события B и C произошли, приходим к выводу, что событие A обязательно наступит. Следовательно, это событие достоверное и вероятность его равна единице. Другими словами, условная вероятность $P_{BC}(A)=1$ события A не равна его безусловной вероятности $P(A) = 1/2$. Итак, попарно независимые события A, B, C не являются независимыми в совокупности.

Приведем теперь следствие из теоремы умножения.

Следствие. *Вероятность совместного появления нескольких событий, независимых в совокупности, равна произведению вероятностей этих событий.*

Доказательство. Рассмотрим три события: A, B к C. Совмещение событий A, B и C равносильно совмещению событий AB и C, поэтому

$$P\left(ABC\right)=P\left(AB\cdot C\right).$$

Так как события A, B и C независимы в совокупности, то независимы, в частности, события AB и C, а также A и B. По теореме умножения для двух независимых событий имеем:

$$P(A\,B \cdot \mathrm{C}) = P(AB)\,P(C) \text{ и } P(AB) = P(A)\,P(B).$$

Итак, окончательно получим

$$P\,(ABC) = P\,(A)\,P\,(B)\,P\,(\mathrm{C}).$$

Для произвольного n доказательство проводится методом математической индукции.

Замечание. Если события $A_1, A_2, ..., A_{n-1}$ независимы в совокупности, то и противоположные им события $\overline{A}_1, \overline{A}_2, ..., \overline{A}_n$ также независимы в совокупности.

Пример 2.11. Найти вероятность совместного появления герба при одном бросании двух-монет.

Решение. Вероятность появления герба первой монеты (событие A)

$$P(A) = 1/2.$$

Вероятность появления герба второй монеты (событие B)

$$P(B) = 1/2.$$

События A и B независимые, поэтому искомая вероятность по теореме умножения равна

$$P(AB) = P(A)P(B) = 1/2 \cdot 1/2 = 1/4.$$

Пример 2.12. Имеется 3 ящика, содержащих по 10 деталей. В первом ящике 8, во втором 7 и в третьем 9 стандартных деталей. Из каждого ящика наудачу вынимают по одной детали. Найти вероятность того, что все три вынутые детали окажутся стандартными.

Решение. Вероятность того, что из первого ящика вынута стандартная деталь (событие A).

$$P(A) = 8/10 = 0{,}8.$$

Вероятность того, что из второго ящика вынута стандартная деталь (событие B),

$$P(B) = 7/10 = 0{,}7.$$

Вероятность того, что из третьего ящика вынута стандартная деталь (событие C),

$$P(C) = 9/10 = 0{,}9.$$

Так как события A, B и C независимые в совокупности, то искомая вероятность (по теореме умножения) равна

$$P(ABC) = P(A)\, P(B)\, P(C) = 0{,}8 \cdot 0{,}7 \cdot 0{,}9 = 0{,}504.$$

Приведем пример совместного применения теорем сложения и умножения.

Пример 2.13. Вероятности появления каждого из трех независимых событий A_1, A_2, A_3 соответственно равны p_1, p_2, p_3. Найти вероятность появления только одного из этих событий.

Решение. Заметим, что, например, появление только первого события A_1 равносильно появлению события $A_1, \overline{A}_2, \overline{A}_3$ (появилось первое и не появились второе и третье события). Введем обозначения:

B_1—появилось только событие A_1 т. е. $B_1 = A_1\overline{A}_2\overline{A}_3$;

B_2 — появилось только событие A_2, т. е $B_2 = A_2\overline{A}_1\overline{A}_3$;

B_3—появилось только событие A_3, т. е $B_3 = A_3\overline{A}_1\overline{A}_2$.

Таким образом, чтобы найти вероятность появления только одного из событий A_1, A_2, A_3, будем искать вероятность $P(B_1 + B_2 + B_3)$ появления одного, безразлично какого из событий $B_1 + B_2 + B_3$.

Так как события B_1, B_2, B_3 несовместны, то применима теорема сложения

$$P(B_1+B_2+B_3)=P(B_1)+P(B_2)+P(B_3). \qquad (2.7)$$

Остается найти вероятности каждого из событий B_1, B_2, B_3.

События A_1, A_2, A_3 независимы, следовательно, независимы события $A_1, \bar{A}_2, \bar{A}_3$, поэтому к ним применима теорема умножения

$$P(B_1)=P(A_1\bar{A}_2\bar{A}_3)=P(A_1)P(\bar{A}_2)P(\bar{A}_3)=p_1q_2q_3.$$

Аналогично,

$$P(B_2)=P(A_2\bar{A}_1\bar{A}_3)=P(A_2)P(\bar{A}_1)P(\bar{A}_3)=p_2q_1q_3;$$

$$P(B_3)=P(A_3\bar{A}_1\bar{A}_2)=P(A_3)P(\bar{A}_1)P(\bar{A}_2)=p_3q_1q_2.$$

Подставив эти вероятности в (2.7), найдем искомую вероятность появления только одного из событии A_1, A_2, A_3:

$$P(B_1+B_2+B_3)=p_1q_2q_3+p_2q_1q_3+p_3q_1q_2.$$

2.3. Вероятность появления хотя бы одного события

Пусть в результате испытания могут появиться *n* событий, независимых в совокупности, либо некоторые из них (в частности, только одно или ни одного), причем вероятности появления каждого из событий известны. Как найти вероятность того, что наступит хотя бы одно из этих событий? Например, если в результате испытания могут появиться три события, то появление хотя бы одного из этих событий означает наступление либо одного, либо двух, либо трех событий. Ответ на поставленный вопрос дает следующая теорема.

Теорема. *Вероятность появления хотя бы одного из событий $A_1, A_2, ..., A_n$, независимых в совокупности, равна разности между единицей и произведением вероятностей противоположных событий $\bar{A}_1, \bar{A}_2, ..., \bar{A}_n$:*

$$P(A)=1 — q_1q_2, \ \dots \ q_n. \qquad (2.8)$$

Доказательство. Обозначим через *А* событие, состоящее в появлении хотя бы одного из событий $A_1, A_2, ..., A_n$. События *А* и $\bar{A}_1\bar{A}_2 \dots \bar{A}_n$ (ни одно из событий не наступило) противоположны, следовательно, сумма их вероятностей равна единице:

$$P(A)+P(\bar{A}_1\bar{A}_2 \dots \bar{A}_n)=1$$

Отсюда, пользуясь теоремой умножения, получим

$$P(A)=1-P(\bar{A}_1\bar{A}_2 \dots \bar{A}_n)=1-P(\bar{A}_1)P(\bar{A}_2)\dots P(\bar{A}_n),$$

или

$$P(A)=1 — q_1q_2, \ \dots \ q_n.$$

Частный случай. *Если события $A_1, A_2, ..., A_n$ имеют одинаковую вероятность, равную р, то вероятность появления хотя бы одного из этих событий*

$$P(A)=1-q^n. \qquad (2.9)$$

Пример 2.14. Вероятности попадания в цель при стрельбе из трех орудий таковы: $p_1 = 0,8$; $p_2 = 0,7$; p_3==0,9. Найти вероятность хотя бы одного попадания (событие *А)* при одном залпе из всех орудий.

Решение. Вероятность попадания в цель каждым из орудий не зависит от результатов стрельбы из других орудий, поэтому рассматриваемые события A_1 (попадание первого орудия), A_2 (попадание второго орудия) и A_3 (попадание третьего орудия) независимы в совокупности.

Вероятности событий, противоположных событиям A_1, A_2 и A_3 (т.е. вероятности промахов), соответственно равны:

$$q_1 = 1 - p_1 = 1 - 0,8 = 0,2;\; q_2 = 1 - p_2 = 1 - 0,7 = 0,3;$$

$$q_3 = 1 - p_3 = 1 - 0,9 = 0,1.$$

Искомая вероятность

$$P(A) = 1 - q_1 q_2 q_3 = 1 - 0,2 \cdot 0,3 \cdot 0,1 = 0,994.$$

Пример 2.15. В типографии имеется 4 плоскопечатные машины. Для каждой машины вероятность того, что она работает в данный момент, равна 0,9. Найти вероятность того, что в данный момент работает хотя бы одна машина (событие *А*).

Решение. События «машина работает» и «машина не работает» (в данный момент) — противоположные, поэтому сумма их вероятностей равна единице:

$$p + q = 1.$$

Отсюда вероятность того, что машкиа в данный момент не работает, равна

$$q = 1 - p = 1 - 0,9 = 0,1.$$

Искомая вероятность

$$P(A) = 1 - q^4 = 1 - 0,1^4 = 0,9999.$$

Так как полученная вероятность весьма близка к единице, то на основании следствия из принципа практической невозможности мало вероятных событий мы вправе заключить, что в данный момент работает хотя бы одна из машин.

Пример 2.16. Вероятность того, что при одном выстреле стрелок попадает в цель, равна 0,4. Сколько выстрелов должен произвести стрелок, чтобы с вероятностью не менее 0,9 он попал в цель хотя бы один раз?

Решение. Обозначим через *А* событие «при *п* выстрелах стрелок попадает в цель хотя бы один раз». События, состоящие в попадании в цель при первом, втором выстрелах и т.д., независимы в совокупности, поэтому применима формула (2.9)

$$P(A) = 1 - q^n.$$

Приняв во внимание, что, по условию, $P(A) \ge 0,9$, $p = 0,4$ (следовательно, q=1—0,4 = 0,6), получим

$$1 - 0,6^n \ge 0,9;\text{ отсюда } 0,6^n \le 0,1.$$

Прологарифмируем это неравенство по основанию 10:

$$n\lg 0,6 \le \lg 0,1.$$

Отсюда, учитывая, что lg0,6 < 0, имеем

$$n \ge \lg 0,1/\lg 0,6 = -1/1,7782 = -1/(-0,2218) = 4,5.$$

Итак, $n \ge 5$, т.е. стрелок должен произвести не менее 5 выстрелов.

Пример 2.17. Вероятность того, что событие появится хотя бы один раз в трех независимых в совокупности испытаниях, равна 0,936. Найти вероятность появления события в одном испытании (предполагается, что во всех испытаниях вероятность появления события одна и та же).

Решение. Так как рассматриваемые события независимы в совокупности, то применима формула (2.9)

$$P(A) = 1 - q^n.$$

По условию, $P(A) = 0,936$; $n = 3$. Следовательно,

$$0,936 = 1 - q^3, \text{ или } q^3 = 1 - 0,936 = 0,064.$$

Отсюда $q = \sqrt[3]{0,064} = 0,4$.

Искомая вероятность

$$p = 1 - q = 1 - 0,4 = 0,6.$$

2.4. Следствия теорем сложения и умножения

Теорема сложения вероятностей совместных событий

Была рассмотрена теорема сложения для несовместных событий. Здесь будет изложена теорема сложения для совместных событий.

Два события называют *совместными,* если появление одного из них не исключает появления другого в одном и том же испытании.

Пример 2.18. *А* — появление четырех очков при бросании игральной кости; В — появление четного числа очков. События *A* и *B*— совместные.

Пусть события *A* и *B* совместны, причем даны вероятности этих событий и вероятность их совместного появления. Как найти вероятность события $A + B$, состоящего в том, что появится хотя бы одно из событий *A* и *B*? Ответ на этот вопрос дает теорема сложения вероятностей совместных событий.

Теорема. *Вероятность появления хотя бы одного из двух совместных событий равна сумме вероятностей этих событий без вероятности их совместного появления:*

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

Доказательство. Поскольку события *A* и *B*, по условию, совместны, то событие $A + B$ наступит, если наступит одно из следующих трех несовместных событий: $A\bar{B}$, $\bar{A}B$ или AB. По теореме сложения вероятностей несовместных событий,

$$P(A + B) = P(A\bar{B}) + P(\bar{A}B) + P(AB). \qquad (2.10)$$

Событие *A* произойдет, если наступит одно из двух несовместных событий: $A\bar{B}$ или AB. По теореме сложения вероятностей несовместных событий имеем

$$P(A) = P(A\bar{B}) + P(AB).$$

Отсюда

$$P\left(A\bar{B}\right)=P\left(A\right)-P\left(AB\right). \tag{2.11}$$

Аналогично имеем

$$P\left(B\right)=P\left(A\bar{B}\right)+P\left(AB\right).$$

Отсюда

$$P\left(A\bar{B}\right)=P\left(B\right)-P\left(AB\right). \tag{2.12}$$

Подставив (2.11) и (2.12) в (2.10), окончательно получим

$$P\left(A+B\right)=P\left(A\right)+P\left(B\right)-P\left(AB\right). \tag{2.13}$$

Замечание 1. При использовании полученной формулы следует иметь в виду, что события A и B могут быть как независимыми, так и зависимыми.

Для независимых событий

$P\ (A + B) = P\ (A) + P\ (B)\text{-}P\ (A)\ P\ (B);$

для зависимых событий

$P\ (A + B) = P\ (A) + P\ (B)\text{-}P\ (A)\ P_A\ (B).$

Замечание 2. Если события A и B несовместны, то их совмещение есть невозможное событие и, следовательно, $P(AB)=0,$ Формула (2.13) для несовместных событий принимает вид

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

Мы вновь получили теорему сложения для несовместных собыий. Таким образом, формула (2.13) справедлива как для совместных, так и для несовместных событий.

Пример 2.19. Вероятности попадания в цель при стрельбе первого и второго орудий соответственно равны: p_1 = 0,7; p_2 = 0,8. Найти вероятность попадания при одном залпе (из обоих орудий) хотя бы одним из орудий.

Решение. Вероятность попадания в цель каждым из орудий не зависит от результата стрельбы из другого орудия, поэтому события A (попадание первого орудия) и B (попадание второго орудия) независимы.

Вероятность события AB. (оба орудия дали попадание)

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

Искомая вероятность

$P\ (A + B)=P\ (A)+P\ (B)—P\ (AB)$ = 0,7 + 0,8-0,56 = 0,94.

Замечание 3. Так как в настоящем примере события A и B независимые, то можно было воспользоваться формулой $P=1-q_1q_2$. В самом деле, вероятности событий, противоположных событиям A и $B,$ т. е. вероятности промахов, таковы:

$q_1=1-p_1=1-0,7=0,3;\qquad q_2=1-p_2=1-0,8=0,2.$

Искомая вероятность того, что при одном залпе хотя бы одно орудие даст попадание, равна

$$P=1-q_1q_2=1-0,3\cdot 0,2=0,94.$$

Как и следовало ожидать, получен тот же результат.

2.5. Формула полной вероятности

Пусть событие A может наступить при условии появления одного из несовместных событий $B_1, B_2, ..., B_n$, которые образуют полную группу. Пусть известны вероятности этих событий и условные вероятности $P_{B_1}(A), P_{B_2}(A), ..., P_{B_n}(A)$ события A. Как найти вероятность события A? Ответ на этот вопрос дает следующая теорема.

Теорема. *Вероятность события A, которое может наступить лишь при условии появления одного из несовместных событий* $B_1, B_2, ..., B_n$, *образующих полную группу, равна сумме произведений вероятностей каждого из этих событий на соответствующую условную вероятность события A:*

$$P(A) = P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A) + ... + P(B_n)P_{B_n}(A).$$

Эту формулу называют «формулой полной вероятности».

Доказательство. По условию, событие A может наступить, если наступит одно из несовместных событий $B_1, B_2, ..., B_n$. Другими словами, появление события A означает осуществление одного, безразлично какого, из несовместных событий $B_1A, B_2A, ..., B_nA$. Пользуясь для вычисления вероятности события A теоремой сложения, получим

$$P(A) = P(B_1A) + P(B_2A) + ... + P(B_nA) \quad . \tag{2.14}$$

Остается вычислить каждое из слагаемых. По теореме умножения вероятностей зависимых событий имеем

$$P(B_1A) = P(B_1)P_{B_1}(A); \quad P(B_2A) = P(B_2)P_{B_2}(A); \ldots;$$
$$P(B_nA) = P(B_n)P_{B_n}(A).$$

Подставив правые части этих равенств в соотношение (2.14), получим формулу полной вероятности

$$P(A) = P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A) + ...$$
$$... + P(B_n)P_{B_n}(A).$$

Пример 2.20. Имеется два набора деталей. Вероятность того, что деталь первого набора стандартна, равна 0,8, а второго — 0,9. Найти вероятность того, что взятая наудачу деталь (из наудачу взятого набора) — стандартная.

Решение. Обозначим через A событие «извлеченная деталь стандартна».

Деталь может быть извлечена либо из первого набора (событие B_1), либо из второго (событие B_2).

Вероятность того, что деталь вынута из первого набора $P(B_1) = 1/2$.

Вероятность того, что деталь вынута из второго набора, $P(B_2) = 1/2$.

Условная вероятность того, что из первого набора будет извлечена стандартная деталь, $P_{B_1}(A) = 0{,}8$.

Условная вероятность того, что из второго набора будет извлечена стандартная деталь, $P_{B_2}(A) = 0{,}9$.

Искомая вероятность того, что извлеченная наудачу деталь - стандартная, по формуле полной вероятности равна

$$P(A)=P(B_1)P_{B_1}(A)+P(B_2)P_{B_2}(A)=$$
$$=0,5\cdot 0,8+0,5\cdot 0,9=0,85.$$

Пример 2.21. В первой коробке содержится 20 радиоламп, из них 18 стандартных; во второй коробке — 10 ламп, из них 9 стандартных. Из второй коробки наудачу взята лампа и переложена в первую. Найти вероятность того, что лампа, наудачу извлеченная из первой коробки, будет стандартной.

Решение. Обозначим через *A* событие «из первой коробки извлечена стандартная лампа».

Из второй коробки могла быть извлечена либо стандартная лампа (событие B_1), либо нестандартная (событие B_2).

Вероятность того, что из второй коробки извлечена стандартная лампа, $P(B_1)=9/10$.

Вероятность того, что из второй коробки извлечена нестандартная лампа, $P(B_2)=1/10$.

Условная вероятность того, что из первой коробки извлечена стандартная лампа, при условии, что из второй коробки в первую была переложена стандартная лампа, равна $P_{B_1}(A)=19/21$.

Условная вероятность того, что из первой коробки извлечена стандартная лампа, при условии, что из второй коробки в первую была переложена нестандартная лампа, равна $P_{B_2}(A)=18/21$.

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

$$P(A)=P(B_1)P_{B_1}(A)+P(B_2)P_{B_2}(A)=(9/10)(19/21)+$$
$$+(1/10)\cdot(18/21)=0,9$$

2.6. Вероятность гипотез. Формулы Бейеса

Пусть событие *A* может наступить при условии появленля одного из несовместных событий $B_1,\ B_2,\ ...,\ B_n$, образующих полную группу. Поскольку заранее не известно, какое из этих событий наступит, их называют *гипотезами.* Вероятность появления события *A* определяется по формуле полной вероятности:

$$P(A)=P(B_1)P_{B_1}(A)+P(B_2)P_{B_2}(A)+...$$
$$...+P(B_n)P_{B_n}(A). \qquad (2.15)$$

Допустим, что произведено испытание, в результате которого появилось событие *A*. Поставим своей задачей определить, как изменились (в связи с тем, что событие *A* уже наступило) вероятности гипотез. Другими словами, будем искать условные вероятности

$$P_A(B_1),\ P_A(B_2),\ \ldots,\ P_A(B_n).$$

Найдем сначала условную вероятность $P_A(B_1)$. По теореме умножения имеем

$$P(AB_1)=P(A)P_A(B_1)=P(B_1)P_{B_1}(A).$$

Отсюда

$$P_A(B_1)=\frac{P(B_1)P_{B_1}(A)}{P(A)}.$$

Заменив здесь *P (A)* по формуле (*), получим

$$P_A(B_1)=\frac{P(B_1)P_{B_1}(A)}{P(B_1)P_{B_1}(A)+P(B_2)P_{B_2}(A)+P(B_n)P_{B_n}(A)}. \qquad (2.16)$$

Аналогично выводятся формулы, определяющие условные вероятности остальных гипотез, т. е. условная вероятность любой гипотезы B_i (i = l, 2, …, *n)* может быть вычислена по формуле

$$P_A(B_i)=\frac{P(B_i)P_{B_i}(A)}{P(B_1)P_{B_1}(A)+P(B_2)P_{B_2}(A)+P(B_n)P_{B_n}(A)}.$$

Полученные формулы называют *формулами Бейеса* (по имени английского математика, который их вывел; опубликованы в 1764 г.). *Формулы Бейеса позволяют переоценить вероятности гипотез после того, как становится известным результат испытания, в итоге которого появилось событие А.*

Пример 2.22. Детали, изготовляемые цехом завода, попадают для проверки их на стандартность к одному из двух контролеров. Вероятность того, что деталь попадает к первому контролеру, равна 0,6, а ко второму — 0,4. Вероятность того, что годная деталь будет признана стандартной первым контролером, равна 0,94, а вторым — 0,98. Годная деталь при проверке была признана стандартной. Найти вероятность того, что эту деталь проверил первый контролер.

Решение. Обозначим через *А* событие, состоящее в том, что годная деталь признана стандартной. Можно сделать два предположения:

1) деталь проверил первый контролер (гипотеза B_1);
2) деталь проверил второй контролер (гипотеза B_2).

Искомую вероятность того, что деталь проверил первый контролер, найдем по формуле Бейеса:

$$P_A(B_1)=\frac{P(B_1)P_{B_1}(A)}{P(B_1)P_{B_1}(A)+P(B_2)P_{B_2}(A)}.$$

По условию задачи имеем:

$P(B_1)$ = 0,6 (вероятность того, что деталь попадает к первому контролеру);

$P(B_2)$==0,4 (вероятность того, что деталь попадет ко второму контролеру);

$P_{B_1}(A)$= 0,94 (вероятность того, что годная деталь будет признана первым контролером стандартной);

$P_{B_2}(A)$ = 0,98 (вероятность того, что годная деталь будет признана вторым контролером стандартной).

Искомая вероятность

$$P_A(B_1) = (0,6 \cdot 0,94)/(0,6 \cdot 0,94 + 0,4 \cdot 0,98) \approx 0,59\,.$$

Как видно, до испытания вероятность гипотезы B_1 равнялась 0,6, а после того, как стал известен результат испытания, вероятность этой гипотезы (точнее, условная вероятность) изменилась и стала равной 0,59. Таким образом, использование формулы Бейеса позволило переоценить вероятность рассматриваемой гипотезы.

3. ПОВТОРЕНИЕ ИСПЫТАНИЙ

3.1. Формула Бернулли

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

В разных независимых испытаниях событие A может иметь либо различные вероятности, либо одну и ту же вероятность. Будем далее рассматривать лишь такие независимые испытания, в которых событие A имеет одну и ту же вероятность.

Ниже воспользуемся понятием *сложного события,* понимая под ним совмещение нескольких отдельных событий, которые называют *простыми.*

Пусть производится n независимых испытаний, в каждом из которых событие A может появиться либо не появиться. Условимся считать, что вероятность события A в каждом испытании одна и та же, а именно равна p. Следовательно, вероятность ненаступления события A в каждом испытании также постоянна и равна q=1 — p.

Поставим перед собой задачу вычислить вероятность того, что при n испытаниях событие A осуществится ровно k раз и, следовательно, не осуществится n — k раз. Важно подчеркнуть, что не требуется, чтобы событие A повторилось ровно k раз в определенной последовательности. Например, если речь идет о появлении события A три раза в четырех испытаниях, то возможны следующие сложные события: $AAA\bar{A}$, $AA\bar{A}A$, $A\bar{A}AA$, $\bar{A}AAA$. Запись $AAA\bar{A}$ означает, что в первом, втором и третьем испытаниях событие A наступило, а в четвертом испытании оно не появилось, т. е. наступило противоположное событие A; соответственный смысл имеют и другие записи.

Искомую вероятность обозначим $P_n(k)$. Например, символ $P_5(3)$ означает вероятность того, что в пяти испытаниях событие появится ровно 3 раза и, следовательно, не наступит 2 раза.

Поставленную задачу можно решить с помощью так называемой формулы Бернулли.

Вывод формулы Бернулли. Вероятность одного сложного события, состоящего в том, что в n испытаниях событие A наступит k раз и не наступит $n—k$ раз, по теореме умножения вероятностей независимых событий равна $p^k q^{n-k}$. Таких сложных событий может быть столько, сколько можно составить сочетаний из n элементов по k элементов, т. е. C_n^k. Так как эти сложные события несовместны, то по теореме сложения вероятностей несовместных событий искомая вероятность равна сумме вероятностей всех возможных сложных событий. Поскольку же вероятности всех этих сложных событий одинаковы, то искомая вероятность (появления k раз события A в n испытаниях) равна вероятности одного сложного события, умноженной на их число:

$$P_n(k) = C_n^k p^k q^{n-k} \tag{3.1}$$

или

Полученную формулу называют *формулой Бернулли.*

Пример 3.1. Вероятность того, что расход электроэнергии в продолжение одних суток не превысит установленной нормы, равна p=0,75. Найти вероятность того, что в ближайшие 6 суток расход электроэнергии в течение 4 суток не превысит нормы.

Решение. Вероятность нормального расхода электроэнергии в продолжение каждых из 6 суток постоянна и равна p = 0,75. Следовательно, вероятность перерасхода электроэнергии в каждые сутки также постоянна и равна q=1—p=*1*—0,75 = 0,25.

Искомая вероятность по формуле Бернулли равна

$$P_6(4) = C_6^4 p^4 q^2 = C_6^2 p^4 q^2 = \frac{6 \cdot 5}{1 \cdot 2}(0,75)^4 \cdot (0,25)^2 = 0,30 .$$

3.2. Локальная теорема Лапласа

Если вероятность р появления события А в каждом испытании постоянна и отлична от нуля и единицы, то вероятность $P_n(k)$ *того, что событие А появится в* я *испытаниях ровно* k *раз, приближенно равна (тем точнее, чем больше n) значению функции*

$$y = \frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} = \frac{1}{\sqrt{npq}} \cdot \varphi(x)$$

при $x = (k - np)\big/\sqrt{npq}$.

Имеются таблицы, в которых помещены значения функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, соответствующие положительным значениям аргумента x. Для отрицательных значений аргумента пользуются теми же таблицами, так как функция $\varphi(x)$ четна, т. е. $\varphi(-x) = \varphi(x)$. Итак, вероятность того, что событие A появится в n независимых испытаниях ровно k раз, приближенно равна

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \cdot \varphi(x), \tag{3.2}$$

где $x=(k-np)/\sqrt{npq}$.

Пример 3.2. Найти вероятность того, что событие A наступит ровно 80 раз в 400 испытаниях, если вероятность появления этого события в каждом испытании равна 0,2.

Решение. По условию, n = 400; k = 80; p=0,2; q = 0,8. Воспользуемся асимптотической формулой Лапласа:

$$P_{400}(80)\approx\frac{1}{\sqrt{400\cdot 0,2\cdot 0,8}}\cdot\varphi(x)=\frac{1}{8}\cdot\varphi(x).$$

Вычислим определяемое данными задачи значение x:

$$x=(k-np)/\sqrt{npq}=(80-400\cdot 0,2)/8=0.$$

По таблице приложения 1 находим $\varphi(0)=0,3989$.

Искомая вероятность

$$P_{400}(80)=(1/8)\cdot 0,3989=0,04986.$$

Формула Бернулли приводит примерно к такому же результату (выкладки ввиду их громоздкости опущены):

$$P_{400}(80)=0,0498.$$

Пример 3.3. Вероятность поражения мишени стрелком при одном выстреле p = 0,75. Найти вероятность того, что при 10 выстрелах стрелок поразит мишень 8 раз.

Решение. По условию, n = 10; k = 8; p = 0,75; q = 0,25. Воспользуемся асимптотической формулой Лапласа:

$$P_{10}(8)\approx\frac{1}{\sqrt{10\cdot 0,75\cdot 0,25}}\cdot\varphi(x)=0,7301\cdot\varphi(x).$$

Вычислим определяемое данными задачи значение

$$x=\frac{k-np}{\sqrt{npq}}=\frac{8-10\cdot 0,75}{\sqrt{10\cdot 0,75\cdot 0,25}}\approx 0,36.$$

По таблице приложения 1 находим $\varphi(0,36)$=0,3739.

Искомая вероятность

$P_{10}(8)$ =0,7301 - 0,3739 = 0,273.

Формула Бернулли приводит к иному результату, а именно $P_{10}(8)$=0,282. Столь значительное расхождение ответов объясняется тем, что в настоящем примере n имеет малое значение (формула Лапласа дает достаточно хорошие приближения лишь при достаточно больших значениях n).

3.3. Формула Пуассона

Предположим, что вероятность p появления события A в отдельном испытании близка к нулю. Такие события A называются редкими. Тогда даже при большом числе испытаний, но небольшой величине произведения np (меньше 10) вероятность, полученная по формуле (3.2), будет недостаточно близка к своему истинному значению. В таких случаях применяют другую приближенную формулу, называемую формулой Пуассона.

Теорема Пуассона. Если вероятность p появления события A в каждом испытании постоянна, но близка к нулю, число независимых испытаний Бернулли достаточно велико, то вероятность того, что в этих n испытаниях событие A наступит ровно m раз, приближенно равна:

$$P_n\left(A=m\right)\frac{\lambda^m \cdot \ell^{-\lambda}}{m!}, \qquad (3.3)$$

Доказательство. Для вычисления интересующей нас вероятности воспользуемся формулой Бернулли и получим:

$$P_n\left(A=m\right)=C_n^m p^m q^{n-m}=\frac{n(n-1)(n-2)\ldots\left(n-(m-1)\right)}{m!}p^m\left(1-p\right)^{n-m}.$$

Из условия $np=\lambda$ следует, что $p=\frac{\lambda}{n}$, следовательно,

$$P_n\left(A=m\right)=\frac{n(n-1)(n-2)\ldots\left(n-(m-1)\right)}{m!}\left(\frac{\lambda}{n}\right)^m\left(1-\frac{\lambda}{n}\right)^{n-m}=$$

$$=\frac{\lambda^m}{m!}\;\frac{n(n-1)(n-2)\ldots\left(n-(m-1)\right)}{n\cdot n\cdot n\ldots n}\left(1-\frac{\lambda}{n}\right)^{n-m}=$$

$$=\frac{\lambda^m}{m!}\cdot 1\cdot\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{m-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n-m}=$$

$$=\frac{\lambda^m}{m!}\left(1-\frac{\lambda}{n}\right)^{n-m}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{m-1}{n}\right)=$$

$$=\frac{\lambda^m}{m!}\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-m}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{m-1}{n}\right).$$

В последнем выражении при больших n второй сомножитель приближенно равен $\ell^{-\lambda}$ (второй замечательный предел), а все остальные сомножители, начиная с третьего, приближенно равны нулю, поэтому

$$P_n\left(A=m\right)\frac{\lambda^m}{m!}\ell^{-\lambda}.$$

Теорема доказана.

Отметим, что имеются специальные таблицы, пользуясь которыми можно найти вероятность $P_n\left(A=m\right)$ для конкретных значений $\lambda=n\cdot p$ и m.

Формула (3.3) называется *формулой Пуассона.*

Пример 3.4. Завод отправил в магазин 5000 исправных телевизоров. Вероятность того, что во время пути произойдет повреждение телевизора, равна 0,0002. Какова вероятность того, что во время пути произойдут повреждения у трех телевизоров? Найти вероятность по разным формулам и сравнить полученные результаты.

Решение. Событие A означает, что во время пути один телевизор получил повреждение. Тогда по условию примера n=5000; m=3; p=0,0002. Вероятность появления события A — поломка в пути одного телевизора — очень мала. Произведение $n{\cdot}p$=5000 · 0,0002 = 1 также мало. Применим формулу (3.1), получим $P_{5000}(A=3)\approx\frac{1^3\cdot\ell^{-1}}{3!}=\frac{1}{6\cdot\ell}\approx 0,061$.

Найдем ту же вероятность по формуле (3.2), используя локальную теорему Муавра—Лапласа, получим:

$$P_{5000}(A=3)\approx\frac{1}{\sqrt{5000\cdot 0,0002\cdot 0,9998}}\cdot\varphi\left(\frac{3-5000\cdot 0,0002}{\sqrt{5000\cdot 0,0002\cdot 0,9998}}\right)\approx$$
$$\approx 1,0001\cdot\varphi(2,0002)\approx 0,4772\,.$$

Нетрудно видеть, что расхождение в ответах очень большое. Проверим, какая из формул дает более точный результат. Для этого найдем туже вероятность по формуле Бернулли (3.1), получим:

$$P_{5000}(A=3)=C_{5000}^{3}\cdot(0,0002)^3\cdot(0,9998)^{4997}\approx$$
$$\approx 20820835000\cdot 8\cdot 10^{-12}\cdot 0,368\approx 0,061297\,.$$

Очевидно, что формула Пуассона дает практически такой же результат, что и формула Бернулли, а формулу Муавра—Лапласа в этих условиях применять не рекомендуется.

3.4. Интегральная теорема Лапласа

Предположим, что производится n испытаний, в каждом из которых вероятность появления события A ностоянна и равна р ($0<p<1$). Как вычислить вероятность $P_n(k_1,\,k_2)$ того, что событие A появится в n испытаниях не менее k_1 и не более k_2 раз (для краткости будем говорить «от k_1 до k_2 раз»)? На этот вопрос отвечает интегральная теорема Лапласа, которую мы приводим ниже, опустив доказательство.

Теорема. *Если вероятность р наступления события А в каждом испытании постоянна и отлична от нулр и единицы, то вероятность* $P_n(k_1,\,k_2)$ *того, что событие А появится в n испытаниях от* k_1 *до* k_2 *раз, приближенно фавна определенному интегралу*

$$P_n(k_1,\,k_2)\cong\frac{1}{\sqrt{2\pi}}\int_{x'}^{x''}e^{-z^2/2}dz\,, \qquad (3.4)$$

где $x'=(k_1-np)\big/\sqrt{npq}$ и $x''=(k_2-np)\big/\sqrt{npq}$.

При решении задач, требующих применения интегральной теоремы Лапласа, пользуются специальными таблицами, так как неопределенный интеграл $\int e^{-z^2/2}dz$ не выражается через элементарные функции. Таблица для интеграла $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-z^2/2}dz$ приведена в конце пособия (приложение 2). В таблице даны значения функции $\Phi(x)$ для положительных значений x и для

$x = 0$; для $x < 0$ пользуются той же таблицей [функция $\Phi(x)$ нечетна, т. е. $\Phi(—x) = — \Phi(x)$]. В таблице приведены значения интеграла лишь до $x = 5$, так как для x>5 можно принять $\Phi(x) = 0{,}5$. Функцию $\Phi(x)$ часто называют *функцией Лапласа.*

Для того чтобы можно было пользоваться таблицей функции Лапласа, преобразуем соотношение (3.4) так:

$$P_n(k_1,\ k_2) \cong \frac{1}{\sqrt{2\pi}} \int_{x'}^{0} e^{-z^2/2} dz + \frac{1}{\sqrt{2\pi}} \int_{0}^{x''} e^{-z^2/2} dz =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{0}^{x''} e^{-z^2/2} dz - \frac{1}{\sqrt{2\pi}} \int_{0}^{x'} e^{-z^2/2} dz = \Phi(x'') - \Phi(x').$$

Итак, вероятность того, что событие A появится в n независимых испытаниях от k_1 до k_2 раз,

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

где $x' = (k_1 - np)\big/\sqrt{npq}$ и $x'' = (k_2 - np)\big/\sqrt{npq}$.

Приведем примеры, иллюстрирующие применение интегральной теоремы Лапласа.

Пример 3.4. Вероятность того, что деталь не прошла проверку ОТК, равна p=0,2. Найти вероятность того, что среди 400 случайно отобранных деталей окажется непроверенных от 70 до 100 деталей.

Решение. По условию, p=0,2; q=0,8; n=400; k_1=70; k_2=100. Воспользуемся интегральной теоремой Лапласа:

$$P_{400}(70,\ 100) \approx \Phi(x'') - \Phi(x').$$

Вычислим нижний и верхний пределы интегрирования:

$$x' = \frac{k_1 - np}{\sqrt{npq}} = \frac{70 - 400 \cdot 0{,}2}{\sqrt{400 \cdot 0{,}2 \cdot 0{,}8}} = -1{,}25;$$

$$x'' = \frac{k_2 - np}{\sqrt{npq}} = \frac{100 - 400 \cdot 0{,}2}{\sqrt{400 \cdot 0{,}2 \cdot 0{,}8}} = 2{,}5.$$

Таким образом, имеем

$$P_{400}(70,\ 100) = \Phi(2{,}5) — \Phi(—1{,}25) = \Phi(2{,}5) + \Phi(1{,}25).$$

По таблице приложения 2 находим:

$$\Phi(2{,}5)=0{,}4938;\ \Phi(1{,}25)=0{,}3944.$$

Искомая вероятность

$$P_{400}(70,\ 100) = 0{,}4938 + 0{,}3944 = 0{,}8882.$$

З а м е ч а н и е. Обозначим через m число появлений события A при n независимых испытаниях, в каждом из которых вероятность наступления события A постоянна и равна p. Если число m. изменяется от k_1 до k_2, то дробь $(n - np)\big/\sqrt{npq}$ изменяется от $(k_1 - np)\big/\sqrt{npq} = x'$ до $(k_2 - np)\big/\sqrt{npq} = x''$. Следовательно, интегральную теорему Лапласа можно записать и так:

$$P\left(x' \le \frac{m - np}{\sqrt{npq}} \le x'' \right) \cong \frac{1}{\sqrt{2\pi}} \int_{x'}^{x''} e^{-z^2/2} dz.$$

3.5. Вероятность отклонения относительной частоты от постоянной вероятности в независимых испытаниях

Вновь будем считать, что производится n независимых испытаний, в каждом из которых вероятность появления события A постоянна и равна $p(0<p<1)$. Поставим перед собой задачу найти вероятность того, что отклонение относительной частоты m/n от постоянной вероятности p по абсолютной величине не превышает заданного числа $\varepsilon > 0$. Другими словами, найдем вероятность осуществления неравенства

$$|m/n - p| \le \varepsilon, \tag{3.5}$$

Эту вероятность будем обозначать так: $P(|m/n - p| \le \varepsilon)$. Заменим неравенство (3.5) ему равносильными:

$$-\varepsilon \le m/n - p \le \varepsilon \quad \text{или} \quad -\varepsilon \le (m - np)/n \le \varepsilon.$$

Умножая эти неравенства на положительный множитель $\sqrt{n/(pq)}$, получим неравенства, равносильные исходному:

$$-\varepsilon\sqrt{n/(pq)} \le (m - np)/\sqrt{npq} \le \varepsilon\sqrt{n/(pq)}.$$

Воспользуемся интегральной теоремой Лапласа. Положив $x' = -\varepsilon\sqrt{n/(pq)}$ и $x'' = \varepsilon\sqrt{n/(pq)}$, имеем

$$P\left(-\varepsilon\sqrt{n/(pq)} \le (m - np)/\sqrt{npq} \le \varepsilon\sqrt{n/(pq)}\right) \approx$$

$$\approx \frac{1}{\sqrt{2\pi}} \int_{-\varepsilon\sqrt{n/(pq)}}^{\varepsilon\sqrt{n/(pq)}} e^{-z^2/2} dz = \frac{2}{\sqrt{2\pi}} \int_{0}^{\varepsilon\sqrt{n/(pq)}} e^{-z^2/2} dz =$$

$$= 2\Phi\left(\varepsilon\sqrt{n/(pq)}\right).$$

Наконец, заменив неравенства, заключенные в скобках, равносильным им исходным неравенством, окончательно получим

$$P(|m/n - p| \le \varepsilon) \cong 2\Phi\left(\varepsilon\sqrt{n/(pq)}\right). \tag{3.6}$$

Итак, вероятность осуществления неравенства $|m/n - p| \le \varepsilon$ приближенно равна значению удвоенной функции Лапласа $2\Phi(x)$ при $x = \varepsilon\sqrt{n/(pq)}$.

Пример 3.5. Вероятность того, что деталь не стандартна, $p = 0{,}1$. Найти вероятность того, что среди случайно отобранных 400 деталей относительная частота появления нестандартных деталей отклонится от вероятности $p = 0{,}1$ по абсолютной величине не более чем на 0,03.

Решение. По условию, n=400; p=0,1; q=0,9; ε=0,03. Требуется найти вероятность $P(|m/400 - 0{,}1| \le 0{,}03)$. Пользуясь формулой $P(|m/n - p| \le \varepsilon) \approx 2\Phi\left(\varepsilon\sqrt{n/(pq)}\right)$, имеем

$$P(|m/400 - 0{,}1| \le 0{,}03) \approx 2\Phi\left(0{,}03\sqrt{400/(0{,}1 \cdot 0{,}9)}\right) = 2\Phi(2).$$

По таблице приложения 2 находим Ф(2)= 0,4772. Следовательно, 2Ф(2)=0,9544.

Итак, искомая вероятность приближенно равна 0,9544.

Смысл полученного результата таков: если взять достаточно большое число проб по 400 деталей в каждой, то примерно в 95,44% этих проб отклонение относительной частоты от постоянной вероятности р=0,1 по абсолютной величине не превысит 0,03.

Пример 3.6. Вероятность того, что деталь не стандартна, p=0,1. Найти, сколько деталей надо отобрать, чтобы с вероятностью, равной 0,9544, можно было утверждать, что относительная частота появления нестандартных деталей (среди отобранных) отклонится от постоянной вероятности p по абсолютной величине не более чем на 0,03.

Решение. По условию, p=0,1 q=0,9; ε=0,03; $P\left(\left|m/n-0{,}1\right|\leq 0{,}03\right)=0{,}9544$. Требуется найти n.

Воспользуемся формулой (3.6)

$$P\left(\left|m/n-p\right|\leq\varepsilon\right)\approx 2\Phi\left(\varepsilon\sqrt{n/(pq)}\right).$$

В силу условия

$$2\Phi\left(0{,}03\sqrt{n/(0{,}1\cdot 0{,}9)}\right)=2\Phi\left(0{,}1\sqrt{n}\right)=0{,}9544.$$

Следовательно, $\Phi\left(0{,}1\sqrt{n}\right)=0{,}4772$.

По таблице приложения 2 находим Ф (2) = 0,4772.

Для отыскания числа я получаем уравнение $0{,}1\sqrt{n}=2$. Отсюда искомое число деталей n = 400.

Смысл полученного результата, таков: если взять достаточно большое число проб по 400 деталей, то 95,44% этих проб относительная частота появления нестандартных деталей будет отличаться от постоянной вероятности p=0,1 по абсолютной величине не более чем на 0,03, т. е. относительная частота заключена в границах от 0,07(0,1—0,03=0,07) до 0,13(0,1+0,03 = 0,13). Другими словами, число нестандартных деталей в 95,44% проб будет заключено между 28(7% от 400) и 52(13% от 400).

Если взять лишь одну пробу из 400 деталей, то с большой уверенностью можно ожидать, что в этой пробе будет нестандартных детялей не менее 28 и не более 52. Возможно, хотя и маловероятно, что нестандартных деталей окажется меньше 28 либо больше 52.

4. СЛУЧАЙНАЯ ВЕЛИЧИНА, ЕЕ ВИДЫ И ЗАКОН РАСПРЕДЕЛЕНИЯ

4.1. Понятия случайной величины и закона распределения вероятностей

В научных исследованиях, да и просто в жизни часто приходится сталкиваться с величинами, которые принимают различные значения в зависимости от случайных обстоятельств. Например, количество звонков, поступивших на телефонную станцию в течение ближайшего часа, величина спроса на некий товар в течение определенного промежутка времени, число очков, выпавших при одном броске игрального кубика, и т.д. При попытке хотя бы приближенно оценить значения этих величин или подобных им можно столкнуться с некоторыми, а иногда и с достаточно большими трудностями. Величины, значения которых зависят от некоторых случайностей, естественно считать случайными.

Случайной величиной называется величина, которая в результате испытания случайно принимает одно значение из множества возможных значений. Например, студент идет сдавать экзамен. Оценка, которую он получит, — случайная величина, множество возможных значений которой числа 2, 3, 4, 5.

Случайные величины будем обозначать большими буквами латинского алфавита: X, $У$, Z и т.д. их возможные значения - соответствующими маленькими буквами с индексами. Например, для случайной величины X ее возможные значения — x_1, x_2, ...

Приведем несколько примеров.

Пример 1. X — число очков, набранных при одном выстреле в мишень.

Пример *2*. X — число очков, выпавших при одном броске игрального кубика.

Пример 3. X — число вызовов, поступивших от абонентов на телефонную станцию в течение определенного промежутка времени.

Пример 4. X — время работы компьютера до первого сбоя.

Пример 5. X — расстояние, которое пролетит снаряд при выстреле из некоторого орудия.

Пример 6. X — вес яблока, сорванного со случайно выбранной яблони.

В первых трех примерах случайная величина X может принимать отдельные изолированные значения, которые можно заранее перечислить. Так, для случайной величины X из примера 2 ее возможными значениями будут числа 1, 2, 3, 4, 5, 6. Для случайной величины из примера 3 возможными значениями будут числа 1, 2, 3, 4, ... и т.д. Таким образом, значения этих величин можно перечислить или записать в виде некоторой последовательности чисел. Такие величины будем называть дискретными.

Дискретной случайной величиной называется случайная величина, множество возможных значений которой конечно или счетно, или, иначе, представимо в виде конечной или бесконечной последовательности чисел.

Заметим, что значения дискретной случайной величины отделены друг от друга промежутками, в которых нет других возможных ее значений, поэтому дискретную случайную величину иногда называют прерывной.

В примерах 4, 5 и 6 значения случайной величины не отделены друг от друга, а заполняют некоторый интервал. Границы этого интервала могут быть приблизительно найдены из смыслового содержания величины (пример 6) или неопределенны (пример 4).

Непрерывной случайной величиной называется случайная величина, множество значений которой сплошь заполняет некоторый конечный или бесконечный промежуток.

Очевидно, что для непрерывной случайной величины, в отличие от дискретной, одно значение нельзя отделить от другого некоторым промежутком.

Отметим, что данное определение непрерывной случайной величины является не совсем точным. Более строгое определение будет дано позже, после рассмотрения понятия так называемой функции распределения.

Случайная величина непосредственно связана со случайными событиями. Появление тех или иных значений случайной величины можно рассматривать как случайные события одного испытания. Предположим, что производится некоторое испыто значение или что случайная величина попадет в какой-нибудь интервал.

Существуют стандартные формы задания закона распределения случайной величины, причем эти формы, кроме одной, являются различными для дискретных и непрерывных случайных величин. Наиболее простые формы закона распределения имеет дискретная случайная величина.

4.2. Формы законов распределения дискретной случайной величины

Рассмотрим дискретную случайную величину X с конечным набором ее возможных значений $x_1, x_2, ..., x_n$. В результате некоторого испытания случайная величина X примет одно из этих значений. Иначе говоря, произойдет одно из следующих несовместных событий, образующих полную группу: $X = x_1$, $X = x_2$, ..., $X = x_n$. Обозначим вероятности этих событий через $p_1 = P(X = x_1)$, $p_2 = P(X = x_2)$, $\ldots$, $p_n = P(X = x_n)$. Рассматриваемые события образуют полную группу, в связи с чем сумма вероятностей этих событий, или, иначе, сумма вероятностей появления возможных значений случайной величины, равна единице, т.е.

$$\sum_{i=1}^{n} P(X = x_i) = \sum_{i=1}^{n} p_i = 1. \qquad (4.1)$$

Таким образом, каждому возможному значению случайной величины $x_к$ поставлено в соответствие число $p_к$ — вероятность появления этого значения, т.е. получена зависимость между возможными значениями случайной величины и вероятностями появления этих значений. Эта зависимость определяет основные формы закона распределения дискретной случайной величины. Она может быть представлена таблично, графически и аналитически. Простейшее представление такой зависимости в виде таблицы называется рядом распределения.

Рядом распределения вероятностей дискретной случайной величины называется ее закон распределения, записанный в виде таблицы, в первой строке которой приведены все возможные значения случайной величины (обычно в порядке возрастания), а во второй — соответствующие этим значениям вероятности.

Ряд распределения для дискретной случайной величины с конечным множеством значений имеет вид:

X	x_1	x_2	$\dots$	x_n
P	p_1	p_2	$\dots$	p_n

Для значений второй строки данной таблицы должна быть выполнена формула (4.1). Таким образом, само название этой таблицы — ряд распределения вероятностей — свидетельствует о том, что единица как бы распределена между всеми значениями случайной величины согласно имеющимся вероятностям.

Если множество значений дискретной случайной величины бесконечно (счетно): $x_1, x_2, ..., x_n$, то ряд распределения будет представлен бесконечной таблицей, а формула (4.1) примет вид

$$\sum_{i=1}^{n} p_i = 1.$$

В дальнейшем в основном будут рассматриваться дискретные случайные величины с конечным множеством значений.

Ряд распределения является табличной формой задания закона распределения. Однако он может быть задан, как и функция в математическом анализе, графическим способом. Если по оси абсцисс отложить возможные значения случайной величины, а по оси ординат — вероятности этих значений, то, соединив точки $(x_к, p_к)$ последовательно отрезками прямой линии, получим ломаную, которая является графической формой закона распределения дискретной случайной величины и называется ***многоугольником распределения вероятностей***. Иногда многоугольником распределения вероятностей называют, по аналогии с геометрическими фигурами, саму фигуру (многоугольник), расположенную под ломаной линией.

Покажем на конкретном примере, как, используя формулы первой главы, найти закон распределения случайной величины.

Пример 4.1. В денежной лотерее разыгрывается 10 выигрышей в 50 руб. и 20 выигрышей по 10 руб. при общем количестве билетов 100. Найти закон распределения в виде ряда распределения и многоугольника распределения выигрыша X для владельца одного лотерейного билета.

Решение. Возможные значения для случайной величины X- это 0 руб., 10 руб. и 50 руб. Вероятности этих значений могут быть найдены по классической формуле, т.е.

$$p_3 = P(X = 50) = \frac{10}{100} = 0,1;$$

$$p_2 = P(X = 10) = \frac{20}{100} = 0,2;$$

$$p_1 = P(X = 0) = 1 - (p_2 + p_3) = 1 - 0,3 = 0,7.$$

Таким образом, ряд распределения имеет вид:

X	0	10	50
p	0,7	0,2	0,1

Построим многоугольник распределения (рис. 4.1).

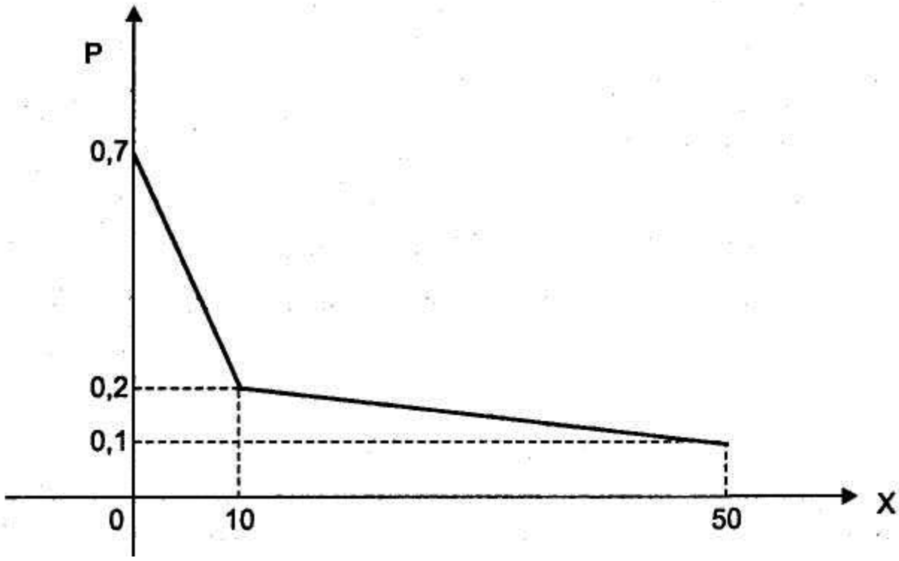

Рис. 4.1

В данном примере многоугольник характеризует невыпуклое множество.

Закон распределения дискретной случайной величины может быть задан и аналитически, в виде функции $p_k = \varphi(x_k)$, k= 1,2, ..., n, позволяющей находить вероятности какого-либо значения по определенной формуле, зная само это значение случайной величины. В зависимости от вида этой функции некоторые законы распределения получили свое название.

Ряд распределения для непрерывной случайной величины построен быть не может, так как невозможно даже записать первую строку этого ряда, т.е. перечислить все возможные значения случайной величины. Более того, в дальнейшем будет показано, что вероятность отдельного значения непрерывной случайной величины просто равна нулю. Однако, несмотря на нулевые вероятности отдельных значений случайной величины, нахождение ее возможных значений в различных, иногда очень маленьких, интервалах обладает различными и отличными от нуля вероятностями. Поэтому распределение непрерывной случайной величины задают, используя не отдельные значения этой величины, а интервалы, в которые могут попадать эти значения. С помощью интервалов может быть задано и распределение дискретной случайной величины. Таким образом, существует особая форма закона распределения, которая относится и к дискретной, и к непрерывной случайным величинам, т.е. является универсальным законом распределения. Этой формой является функция распределения.

4.3. Биномиальное распределение

Пусть производится n независимых испытаний, в каждом из которых событие A может появиться либо не появиться. Вероятность наступления события во всех испытаниях постоянна и равна p (следовательно, вероятность непоявления $q=l—p$). Рассмотрим в качестве дискретной случайной величины X число появлений события A в этих испытаниях.

Поставим перед собой задачу: найти закон распределения величины X. Для ее решения требуется определить возможные значения X и их вероятности. Очевидно, событие A в n испытаниях может либо не появиться, либо появиться 1 раз, либо 2 раза, ..., либо n раз. Таким образом, возможные значения X таковы: $x_1=0$, $x_2=1$, $x_3=2$, ..., $x_{n+1}=n$. Остается найти вероятности этих возможных значений, для чего достаточно воспользоваться формулой Бернулли (3.1), рассмотренную в предыдущей главе:

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

где $k = 0, 1, 2, ..., n$.

Формула (3.1) и является аналитическим выражением искомого закона распределения.

Биномиальным называют распределение вероятностей, определяемое формулой Бернулли, Закон назван «биномиальным» потому, что правую часть равенства (4.2) можно рассматривать как общий член разложения бинома Ньютона:

$$(p+q)^n = C_n^n p^n + C_n^{n-1} p^{n-1} q + ... + C_n^k p^k q^{n-k} + ... + C_n^0 q^n .$$

Таким образом, первый член разложения p^n определяет вероятность наступления рассматриваемого события n раз в n независимых испытаниях; второй член $np^{n-1}q$ определяет вероятность наступления события n — 1 раз; ...; последний член q^n определяет вероятность того, что событие не появится ни разу.

Напишем биномиальный закон в виде таблицы:

X	n	n-1	…	K	…	0
P	p^n	$np^{n-1}q$	…	$C_n^k p^k q^{n-k}$	…	q^n

Пример 4.2. Монета брошена 2 раза. Написать в виде таблицы закон распределения случайной величины X — числа выпадений «герба».

Решение. Вероятность появления «герба» в каждом бросании монеты $p=1/2$, следовательно, вероятность непоявления «герба» $q=1-1/2=1/2$.

При двух бросаниях монеты «герб» может появиться либо 2 раза, либо 1 раз, либо совсем не появиться. Таким образом, возможные значения X таковы: $x_1=2$, $x_2=1$, $x_3=0$. Найдем вероятности этих возможных значений по формуле Бернулли:

$$P_2(2)=C_2^2 p^2 = 0,25,$$

$P_2(1) = C_2^1 pq = 2 \cdot (1/2) \cdot (1/2) = 0{,}5$,

$P_2(0) = C_2^0 q^2 = (1/2)^2 = 0{,}25$.

Напишем искомый закон распределения:

X	2	1	0
P	0,25	0,5	0,25

Контроль: 0,25+0,5 + 0,25=1.

4.4. Геометрическое распределение

Пусть производятся независимые испытания, в каждом из которых вероятность появления события A равна p ($0<p<1$) и, следовательно, вероятность его непоявления $q=1—p$. Испытания заканчиваются, как только появится событие A. Таким образом, если событие A появилось в k-м испытании, то-в предшествующих $k—1$ испытаниях оно не появлялось.

Обозначим через X дискретную случайную величину — число испытаний, которые нужно провести до первого появления события A. Очевидно, возможными значениями X являются натуральные числа: $x_1 = 1$, $x_2=2$, ...

Пусть в первых $k—1$ испытаниях событие A не наступило, а в k-м испытании появилось. Вероятность этого «сложного события», по теореме умножения вероятностей независимых событий,

$$P(X = k) = q^{k-1} p. \qquad (4.2)$$

Полагая $k=1$, 2, ... в формуле (*), получим геометрическую прогрессию с первым членом p и знаменателем q ($0<q<1$):

$$p,\ pq,\ q^2 p,\ \ldots,\ q^{k-1} p,\ \ldots . \qquad (4.3)$$

По этой причине распределение (4.2) называют *геометрическим.*

Легко убедиться, что ряд (4.3) сходится и сумма его равна единице. Действительно, сумма ряда (4.3)

$$p/(1-q) = p/p = 1.$$

Пример 4.4. Из орудия производится стрельба по цели до первого попадания. Вероятность попадания в цель $p = 0{,}6$. Найти вероятность того, что попадание произойдет при третьем выстреле.

Решение. По условию, $p=0{,}6$, $q=0{,}4$, $k=3$. Искомая вероятность по формуле (4.2)

$$P = q^{k-1} \cdot p = 0{,}4^2 \cdot 0{,}6 = 0{,}096.$$

4.5. Гипергеометрическое распределение

Прежде чем дать определение гипергеометрического распределения, рассмотрим задачу. Пусть в партии из N изделий имеется M стандартных ($M<N$). Из партии случайно отбирают n изделий (каждое изделие может быть

извлечено с одинаковой вероятностью), причем отобранное изделие теред отбором следующего не возвращается в партию (поэтому формула Бернулли здесь неприменима). Обозначим через X случайную величину — число m стандартных изделий среди n отобранных. Очевидно, возможные значения X таковы: 0, 1, 2, ..., min (M, n).

Найдем вероятность того, что $X = m$, т. е. что среди n отобранных изделий ровно m стандартных. Используем для этого классическое определение вероятности.

Общее число возможных элементарных исходов испытания равно числу способов, которыми можно извлечь n изделий из N изделий, т. е. числу сочетаний C_N^n.

Найдем число исходов, благоприятствующих событию $X=m$ (среди взятых n изделий ровно m стандартных); m стандартных изделий можна извлечь из M стандартных изделий C_M^m способами; при этом остальные $n—m$ изделий должны быть нестандартными; взять же $n—m$ нестандартных изделий из $N—m$ нестандартных изделий можно C_{N-M}^{n-m} способами. Следовательно, число благоприятствующих исходов равно $C_M^m C_{N-M}^{n-m}$.

Искомая вероятность равна отношению числа исходов, благоприятствующих событию $X — m$, к числу всех элементарных исходов

$$P\left(X = m\right) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}. \qquad (4.4)$$

Формула (4.4) определяет распределение вероятностей, которое называют *гипергеометрическим.*

Учитывая, что m — случайная величина, заключаем, что гипергеометрическое распределение определяется тремя параметрами: N, M, n. Иногда в качестве параметров этого распределения рассматривают N, n и $p=M/N$, где p — вероятность того, что первое извлеченное изделие стандартное.

Заметим, что если n значительно меньше N (практически если $n<0,1N$), то гипергеометрическое распределение дает вероятности, близкие к вероятностям, найденным по биномиальному закону.

Пример 4.5. Среди 50 изделий 20 окрашенных. Найти вероятность того, что среди наудачу извлеченных 5 изделий окажется ровно 3 окрашенных.

Решение. По условию, $N=50$, $M=20$, $n=5$, $m=3$. Искомая вероятность

$$P\left(X = 3\right) = C_{20}^3 C_{30}^2 / C_{50}^5 = 0,234.$$

5. ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН

5.1. Математическое ожидание и его свойства

Для решения многих практических задач не всегда требуется знание всех возможных значений случайной величины и их вероятностей. Более того, иногда закон распределения исследуемой случайной величины просто неизвестен. Однако требуется выделить какие-то особенности этой случайной величины, иначе говоря, числовые характеристики.

Числовые характеристики — это некоторые числа, характеризующие те или иные свойства, отличительные признаки случайной величины, например, среднее значение случайной величины, средний разброс всех значений случайной величины вокруг своего среднего и т.д. Главное назначение числовых характеристик состоит в том, чтобы в сжатой форме выразить наиболее важные особенности распределения исследуемой случайной величины. Числовые характеристики в теории вероятностей играют огромную роль. Они помогают решать, даже без знания законов распределения, очень многие важные практические задачи.

Среди всех числовых характеристик в первую очередь выделим *характеристики положения*. Это характеристики, которые фиксируют положение случайной величины на числовой оси, т.е. некое среднее значение, около которого группируются остальные значения случайной величины.

Из характеристик положения наибольшую роль в теории вероятностей играет математическое ожидание.

Математическое ожидание иногда называют просто средним значением случайной величины. Оно является неким центром распределения.

5.1.1. Математическое ожидание дискретной случайной величины

Рассмотрим понятие математического ожидания вначале для дискретной случайной величины.

Прежде чем вводить формальное определение, решим следующую простую задачу.

Пример 5.1. Пусть некий стрелок производит 100 выстрелов по мишени. В результате получена следующая картина: 50 выстрелов — попадание в «восьмерку», 20 выстрелов — попадание в «девятку» и 30 — в «десятку». Какова средняя сумма очков при одном выстреле?

Решение Очевидно, что решение данной задачи сводится 1 к нахождению среднего значения 100 чисел:

$$x_{\text{ср}} = \frac{8\cdot 50 + 9\cdot 20 + 10\cdot 30}{100} = \frac{880}{100} = 8{,}8 \text{ очков.}$$

Преобразуем дробь, почленно поделив числитель на знаменатель, и представим среднее значение в виде следующей формулы:

$$x_{\text{ср}} = 8\cdot\frac{50}{100} + 9\cdot\frac{20}{100} + 10\cdot\frac{30}{100} = 8\cdot 0{,}5 + 9\cdot 0{,}2 + 10\cdot 0{,}3 = 8{,}8\,.$$

Предположим теперь, что число очков при одном выстреле — это значения некоторой дискретной случайной величины *X*. Из условия задачи ясно, что $x_1 = 8$; x_2=9; x_3=10. Известны относительные частоты появления этих значений, которые, как известно, при большом числе испытаний приближенно равны вероятностям соответствующих значений, т.е. $p_1 \approx 0{,}5$; $p_2 \approx 0{,}2$; $p_3 \approx 0{,}3$. Итак,

$x_{ср} \approx x_1 \cdot p_1 + x_2 \cdot p_2 + x_3 \cdot p_3$. Величина в правой части — это математическое ожидание дискретной случайной величины.

Математическим ожиданием дискретной случайной величины X называется сумма произведений всех ее возможных значений на вероятности этих значений.

Пусть дискретная случайная величина X задана своим рядом распределения:

X	x_1	x_2	$\dots$	x_n
P	p_1	p_2	$\dots$	p_n

Тогда математическое ожидание $M(X)$ дискретной случайной величины определяется по следующей формуле:

$$M(X) = x_1 \cdot p_1 + x_2 \cdot p_2 + \dots + x_n \cdot p_n = \sum_{i=1}^{n} x_i \cdot p_i . \qquad (5.1)$$

Если дискретная случайная величина принимает бесконечное счетное множество значений, то математическое ожидание выражается формулой

$$M(X) = \sum_{i=1}^{\infty} x_i \cdot p_i ,$$

причем математическое ожидание существует, если ряд в правой части равенства абсолютно сходится.

Пример 5.2. Найти математическое ожидание выигрыша X условиях примера 4.1.

Решение. Напомним, что ряд распределения Xимеет следующий вид:

X	0	10	50
P	0,7	0,2	0,1

Получим $M(X) = 0 \cdot 0{,}7 + 10 \cdot 0{,}2 + 50 \cdot 0{,}1 = 7$. Очевидно, что 7 руб. — справедливая цена билета в данной лотерее, без различных затрат, например, связанных с распространением или изготовлением билетов.

Пример 5.3. Пусть случайная величина X— это число появлений некоторого события A в одном испытании. Вероятность этого события равна p. Найти $M(X)$.

Решение. Очевидно, что возможные значения случайной величины $x_1 = 0$ - событие A не появилось и $x_2 = 1$ - событие A появилось. Ряд распределения имеет вид

X	0	1
P	$1-p$	p

Тогда $M(X) = 0 \cdot (1-p) + 1 \cdot p = p$.

Итак, математическое ожидание числа появлений события в одном испытании равно вероятности этого события.

В начале данного подраздела была приведена конкретная задача, где указывалась связь между математическим ожиданием и средним значением случайной величины. Поясним это в общем виде.

Пусть произведено k испытаний, в которых случайная величина X приняла k_1 раз значение x_1; k_2 раз значение x_2 и т.д. и, наконец, k_n раз значение x_n. Очевидно, что $k_1+k_2+\ldots+k_n=k$.

Найдем среднее арифметическое $\bar{X}$ всех этих значений, имеем

$$\bar{X}=\frac{x_1k_1+x_2k_2+\ldots+x_nk_n}{k_1+k_2+\ldots+k_n}=\frac{x_1k_1+x_2k_2+\ldots+x_nk_n}{k}=\sum_{i=1}^{n}x_i\cdot\frac{k_i}{k}.$$

Заметим, что дробь $\dfrac{k_i}{k}$ — это относительная частота появления значения x_i в k испытаниях. При большом числе испытаний относительная частота приближенно равна вероятности, т.е. $p_i=\dfrac{k_i}{k}$. Отсюда следует, что

$$\bar{X}=\sum_{i=1}^{n}x_i\cdot\frac{k_i}{k}\approx\sum_{i=1}^{n}x_i\cdot p_i=M\left(X\right).$$

Таким образом, математическое ожидание приближенно равно среднему арифметическому наблюдаемых значений случайной величины, причем тем точнее, чем больше число испытаний, — в этом состоит *вероятностный смысл математического ожидания.*

Математическое ожидание иногда называют *центром распределения случайной величины,* так как очевидно, что возможные значения случайной величины расположены на числовой оси слева и справа от ее математического ожидания.

Перейдем теперь к понятию математического ожидания для непрерывной случайной величины.

5.1.2. Свойства математического ожидания

Математическое ожидание случайной величины обладает следующими *свойствами:*

1. Математическое ожидание постоянной величины равно самой постоянной, т.е. $M(C)=C$.

Доказательство. Постоянную величину можно рассматривать как дискретную с одним значением $x_1 = C$ и вероятностью этого значения $p_1=1$. По формуле (5.1) получим $M(C) = C\cdot 1 = C$.

Свойство доказано.

2. Математическое ожидание алгебраической суммы двух случайных величин равно алгебраической сумме их математических ожиданий, т.е. $M(X\pm Y) = M(X)\pm M(Y)$.

Доказательство. Доказательство проведем только для дискретных случайных величин. Пусть случайные величины X и Y заданы своими рядами распределения:

X	x_1	x_2	$\dots$	x_n
P	p_1	p_2	$\dots$	p_n

и

X	y_1	y_2	$\dots$	y_m
Q	q_1	q_2	$\dots$	q_m

Возможными значениями суммы $X\pm$ Yявляются числа $x_i \pm y_j$. Обозначим через p_{ij} вероятность того, что величина X примет значение x_i а величина Y — значение y_j. По определению математического ожидания имеем

$$M\left(X \pm Y\right) = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(x_i \pm y_j\right)\cdot p_{ij} = \sum_{i=1}^{n} x_i\left(\sum_{j=1}^{m} p_{ij}\right) \pm \sum_{j=1}^{m} y_j\left(\sum_{i=1}^{n} p_{ij}\right).$$

Нетрудно понять, что по теореме о полной вероятности имеют место равенства $\sum\limits_{j=1}^{m} p_{ij}$ и $\sum\limits_{i=1}^{n} p_{ij}$. Следовательно

$$M\left(X \pm Y\right) = \sum_{i=1}^{n} x_i p_i \pm \sum_{j=1}^{m} y_j p_j = M\left(X\right) \pm M\left(Y\right).$$

Свойство доказано.

С л е д с т в и е. Математическое ожидание алгебраической суммы конечного числа случайных величин равно алгебраической сумме математических ожиданий этих величин.

Доказательство данного следствия можно провести методом математической индукции.

3. Математическое ожидание произведения двух независимых случайных величин равно произведению их математических ожиданий, т.е. если X и Y — независимые случайные величины, то $M(X \cdot Y) = M(X) \cdot M(Y)$.

Д о к а з а т е л ь с т в о. Доказательство проведем также только для дискретных случайных величин. Пусть случайные величины X и Y заданы, как и при доказательстве свойства 2, своими рядами распределения. Очевидно, что с учетом независимости случайных величин ряд распределения случайной величины $Z= X\cdot Y$ имеет вид

XY	x_1y_1	x_1y_1	$\dots$	x_1y_m	x_2y_1	x_2y_2	$\dots$	x_2y_m	$\dots$	x_ny_1	x_ny_1	$\dots$	x_ny_m
P	p_1q_1	p_1q_2	$\dots$	p_1q_m	p_2q_1	p_2q_2	$\dots$	p_2q_m	$\dots$	p_nq_1	p_nq_2	$\dots$	p_nq_m

Согласно определению математического ожидания получим:

$$M\left(X \cdot Y\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i \cdot y_j \cdot p_i \cdot q_j = \sum_{i=1}^{n} x_i p_i \cdot \sum_{j=1}^{m} y_j q_j = M\left(X\right)\cdot M\left(Y\right).$$

Свойство доказано.

С л е д с т в и е. Математическое ожидание произведения конечного числа независимых случайных величин равно произведению математических ожиданий этих величин.

Доказательство данного следствия можно провести методом математической индукции.

4. Постоянный множитель можно выносить за знак математического ожидания, т.е. $M(C\cdot X)=C\cdot M(X)$.

Д о к а з а т е л ь с т в о. Применив свойства 3 и 1, получим:

$$M(C\cdot X)=M(C)\cdot M(X)=C\cdot M(X).$$

Свойство доказано.

Несмотря на то, что доказательство свойств приведено для дискретных случайных величин, они все справедливы и для непрерывных случайных величин.

Пример 5.4. Найти математическое ожидание числа появлений события *A* в *n* независимых испытаниях, если вероятность появления *A* в каждом испытании постоянна и равна *p*.

Решение. Пусть случайная величина *X* — число появлений события *A в n* испытаниях. Введем в рассмотрение еще и случайных величин:

X_1 — число появлений события *A* в 1-м испытании;

X_2 — число появлений события *A* во 2-м испытании;

…………………………………………………………

X_n — число появлений события *A* в n-м испытании.

Очевидно, что $X = X_1 + X_2+\ldots+X_n$. Используя свойство 2, получим:

$$M(X)=M(X_1+X_2+\ldots+X_n)=M(X_1)+M(X_2)+\ldots+M(X_n)=$$
$$=p+p+\ldots+p=n\cdot p.$$

В последнем равенстве использовались результаты примера 5.3.

Математическое ожидание — это не единственная характеристика положения случайной величины. К таким характеристикам относятся также мода и медиана.

Модой случайной величины называется ее наиболее вероятное значение.

Очевидно, что для дискретной случайной величины модой является то значение x_i, для которого вероятность p_i является самой большой. Для непрерывной случайной величины модойявляется то значение *x,* при котором функция плотности $f(x)$ достигает максимального значения.

Если вероятность или плотность вероятности достигают максимума не в одной, а в нескольких точках, распределение называется *полимодальным* (многомодальным); если в одной точке - то *унимодальным* (одномодальным).

Медианой случайной величины *X* называется такое значение x_m, *для* которого одинаково вероятными оказываются следующие события: "$X<x_m$" и "$X>x_m$".

Как правило, медиана применяется в основном для непрерывных случайных величин. Если x_m — медиана некоторой непрерывной случайной величины, то для нее выполнены равенства

$$P\left(X < x_m\right) = P\left(X > x_m\right) = \frac{1}{2}.$$

Геометрически медиана — это точка на оси абсцисс, такая, что прямая, проходящая через эту точку параллельно оси ординат, делит фигуру под графиком функции плотности на две равные по площади фигуры. Очевидно, что площади этих фигур равны $\frac{1}{2}$.

Кроме характеристик положения распределение случайной величины могут определять характеристики разброса.

5.2. Дисперсия и среднеквадратическое отклонение

При решении практических задач могут встретиться случайные величины, имеющие разные распределения, но одинаковые математические ожидания. При этом у одних из этих величин отклонения значений от математического ожидания небольшие, у других, наоборот, могут быть значительными. Иначе говоря, у величин может быть разный разброс значений вокруг математического ожидания. Например, для двух дискретных случайных величин, заданных следующими законами:

X	-1	0	1
P	0,3	0,4	0,3

и

Y	-100	0	100
P	0,2	0,6	0,2

математические ожидания равны, т.е. $M(X) = M(Y) = 0$. Однако понятно, что это разные случайные величины и они отличаются прежде всего разбросом значений по оси абсцисс слева и справа от точки 0 — своего математического ожидания.

Приведенные рассуждения свидетельствуют о том, что было бы целесообразно ввести в рассмотрение некоторую числовую характеристику, связанную с разбросом. На первый взгляд может показаться, что такой характеристикой может быть среднее значение всех отклонений возможных значений случайной величины от математического ожидания.

Отклонением случайной величины X от своего математического ожидания M{X) называется разность между случайной величиной и ее математическим ожиданием.

Очевидно, что отклонение также является случайной величиной. Найдем среднее значение отклонения, т.е. математическое ожидание отклонения:

$$M(X - M(X)) = M(X) - M(MX)) = M(X) - M(X) = 0.$$

Итак, математическое ожидание отклонения случайной величины равно нулю. Этот факт можно объяснить также тем, что возможные значения отклонения имеют как положительные, так и отрицательные знаки, поэтому при нахождении среднего значения (математического ожидания) слагаемые взаимно уничтожаются. Избежать этого можно, убрав отрицательные знаки значений отклонения. Для этого эти значения либо берут по абсолютной величине, либо возводят в квадрат. Первый путь используется крайне редко, так как работа с абсолютными величинами вызывает, как правило, серьезные трудности, например, при дифференцировании. Поэтому в качестве характеристики разброса используют математическое ожидание квадрата отклонения.

5.2.1. Определение дисперсии и среднеквадратического отклонения

Дисперсией $D(X)$[1] случайной величины X называется математическое ожидание квадрата отклонения данной случайной величины от своего математического ожидания, т.е.

$$D(X) = M\left[\left(X - M(X)\right)^2\right]. \tag{5.2}$$

Нетрудно понять, что вероятности значений случайных величин X и $\left(X - M(X)\right)^2$ одинаковы. Для того чтобы величина $\left(X - M(X)\right)^2$ приняла значение, например, $\left(x_1 - M(X)\right)^2$, достаточно, чтобы случайная величина X приняла значение x_1. Вероятность этого события равна *p.,* следовательно, и вероятность того, что величина $\left(X - M(X)\right)^2$ примет значение $\left(x_1 - M(X)\right)^2$, также равна p_1. Аналогично обстоит дело и с остальными возможными значениями. Поэтому формула (5.2) с учетом определения математического ожидания случайной величины примет вид:

для дискретной случайной величины с конечным множеством значений —

$$D(X) = \sum_{i=1}^{n}\left(x_i - M(X)\right)^2 \cdot p_i\,; \tag{5.3}$$

Среднеквадратическим отклонением $\sigma(X)$ случайной величины X называется квадратный корень из дисперсии, т.е.

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

Пример 5.5. Найти дисперсию дискретной случайной величины, заданной следующим рядом распреджеления:

X	1	2	5
P	0,3	0,5	0,2

Решение. Получим вначале матиматеическое ожилание данной случайной величины: $M(X) = 1\cdot 0{,}3 + 2\cdot 0{,}5 + 5\cdot 0{,}2 = 2{,}3$.

Найдем закон распределения величины $\left(X - M(X)\right)^2$:

[1] Само слово «дисперсия» в переводе с латинского языка означает «рассеивание».

$\left(X-M\left(X\right)\right)^{2}$	$\left(1-2,3\right)^{2}$	$\left(2-2,3\right)^{2}$	$\left(5-2,3\right)^{2}$
P	0,3	0,5	0,2

После вычислений получим:

$\left(X-M\left(X\right)\right)^{2}$	1,69	0,09	7,29
P	0,3	0,5	0,2

Найдем математическое ожидание полученной случайной величины: $D\left(X\right)=M\left[\left(X-M\left(X\right)\right)^{2}\right]=1{,}69\cdot 0{,}3+0{,}09\cdot 0{,}5+7{,}29\cdot 0{,}2=2{,}01$.

Теорема 6.1. Дисперсия случайной величины равна разности между математическим ожиданием квадрата этой случайной величины и квадратом математического ожидания:

$$D\left(X\right)=M\left(X^{2}\right)-M^{2}\left(X\right). \tag{5.4}$$

Доказательство. Преобразуем формулу (5.2), используя свойства математического ожидания:

$$D\left(X\right)=M\left[\left(X-M\left(X\right)\right)^{2}\right]=M\left[X^{2}-2XM\left(X\right)+M^{2}\left(X\right)\right]=M\left(X^{2}\right)-$$
$$-M\left(2XM\left(X\right)\right)+M\left(M^{2}\left(X\right)\right)=M\left(X^{2}\right)-2M\left(X\right)M\left(X\right)+M^{2}\left(X\right)=$$
$$=M\left(X^{2}\right)-2M^{2}\left(X\right)+M^{2}\left(X\right)=M\left(X^{2}\right)-M^{2}\left(X\right).$$

Теорема доказана.

Пример 5.6. Решить пример 5.5, используя формулу (5.4).

Решение. Математическое ожидание уже получено, оно равно $M(X)=2{,}3$. Теперь найдем закон распределения величины X^2:

X^4	1	4	25
P	0,3	0,5	0,2

Определим $M\left(X^{2}\right)=1\cdot 0{,}3+4\cdot 0{,}5+25\cdot 0{,}2=7{,}3$. Тогда дисперсия равна

$$D\left(X\right)=7{,}3-\left(2{,}3\right)^{2}=2{,}01.$$

Очевидно, что применение формулы (5.4) значительно упрощает процесс нахождения дисперсии. Понятно, что эту же формулу можно применять и для определения дисперсии непрерывной случайной величины.

5.2.2. Свойства дисперсии

Дисперсия случайной величины обладает следующими свойствами.

1. Дисперсия постоянной величины равна нулю, т.е. $D\left(C\right)=0$, где C - постоянная величина.

Доказательство. Для нахождения дисперсии постоянной величины С используем формулу (5.2), а затем применим свойства математического ожидания:

$$D(C)=M\left[(C-M(C))^2\right]=M\left[(C-C)^2\right]=M(0)=0.$$

Этот результат достаточно очевиден, так как постоянная величина принимает всего одно значение, поэтому разброс значений отсутствует.

Свойство доказано.

2. Постоянный множитель можно выносить за знак дисперсии, возводя его в квадрат, т.е. $D(CX)=C^2D(X)$.

Доказательство. По определению дисперсии с использованием свойств математического ожидания получим:

$$D(CX)=M\left[(CX-M(CX))^2\right]=M\left[(CX-CM(X))^2\right]=$$

$$=M\left[C^2(X-M(X))^2\right]=C^2M\left[(X-M(X))^2\right]=C^2D(X).$$

Свойство доказано.

3. Дисперсия суммы двух независимых случайных величин равна сумме дисперсий этих величин, т.е. если величины X и Y независимы, то

$$D(X+Y)=D(X)+D(Y).$$

Доказательство. Для доказательства применим формулу (5.4) и свойства математического ожидания, получим:

$$D(X+Y)=M\left((X+Y)^2\right)-M^2(X+Y)=M\left(X^2+2XY+Y^2\right)-\left(M(X+Y)\right)^2=$$

$$=M\left(X^2\right)+M(2XY)+M\left(Y^2\right)-\left(M(X)+M(Y)\right)^2=M\left(X^2\right)+$$

$$+2M(X)M(Y)+M\left(Y^2\right)-M^2(X)-2M(X)M(Y)-M^2(Y)=$$

$$=M\left(X^2\right)-M^2(X)+M\left(Y^2\right)-M^2(Y)=D(X)+D(Y).$$

Свойство доказано.

Следствие. Дисперсия суммы нескольких независимых величин равна сумме дисперсий этих величин.

Доказательство можно провести методом математической индукции

4. Дисперсия разности двух независимых случайных величин равна сумме их дисперсий:

$$D(X-Y)=D(X)+D(Y).$$

Доказательство. Применяя свойства 2 и 3 дисперсии, получим:

$$D(X-Y)=D(X)+D(-Y)=D(X)+(-1)^2D(Y)=D(X)+D(Y).$$

Свойство доказано.

Доказанное свойство также легко распространить на любое конечное число независимых случайных величин.

Пример 5.7. Найти дисперсию дискретной случайной величины *X,* равной числу появлений события *A* в *n* независимых испытаниях, если вероятность появления *A* в каждом испытании постоянна и равна *p.*

Решение. Пусть случайная величина *X* — число появлений события *A* в *n* испытаниях.

Введем в рассмотрение еще *n* случайных величин:

X_1 - число появлений события A в 1-м испытании;

X_2 - число появлений события A во 2-м испытании;

X_n — число появлений события A в n-м испытании.

Очевидно, что $X=X_1+X_2+...+X_n$. Величины X_1, X_2, ..., X_n взаимно независимы, так как исход каждого испытания не зависит от исходов остальных. Воспользуемся следствием свойства 4 дисперсии, получим:

$$D\left(X\right)=D\left(X_1\right)+D\left(X_2\right)+...+D\left(X_n\right).$$

Найдем дисперсию величины X_1 Ряд распределения этой величины имеет вид

X_1	0	1
P	$1-p$	p

Тогда

$$M\left(X_1\right)=p;\ M\left(X_1^{\ 2}\right)=p;\qquad D\left(X_1\right)=p-p^2=p\left(1-p\right)=pq.$$

Очевидно, что дисперсия каждой из остальных случайных величин также равна pq, поэтому

$$D(X) = D(X_1)+D(X_2)+...+D(X_n) = npq.$$

6. ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Определение функции распределения

Вспомним, что дискретная случайная величина может быть задана перечнем всех ее возможных значений и их вероятностей. Такой способ задания не является общим: он неприменим, например, для непрерывных случайных величин.

Действительно, рассмотрим случайную величину X, возможные значения которой сплошь заполняют интервал (a, b). Можно ли составить перечень всех возможных значений X? Очевидно, что этого сделать нельзя. Этот пример указывает на целесообразность дать общий способ задания любых лилов случайных величин. С этой целью и вводят функции распределения вероятностей случайной величины.

Пусть x — действительное число. Вероятность события, состоящего в том, что X примет значение, меньшее x, т. е. вероятность события $X < x$, обозначим через $F(x)$. Разумеется, если x изменяется, то, вообще говоря, изменяется и $F(x)$, т. е. $F(x)$ — функция от x.

Функцией распределения называют функцию $F(x)$, определяющую вероятность того, что случайная величина X в результате испытания примет значение, меньшее x, т. е.

$$F\left(x\right)=P\ \left(X<x\right).$$

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

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

Теперь можно дать более точное определение непрерывной случайной величиной случайную величину называют *непрерывной,* если ее функция распределения есть непрерывная, кусочно-дифференцируемая функция с непрерывной производной.

Свойства функции распределения

Свойство 1. *Значения функции распределения принадлежат отрезку* [0,1]:

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

Доказательство. Свойство вытекает из определения функции распределения как вероятности: вероятность всегда есть неотрицательное число, не превышающее единицы.

Свойство 2. $F(x)$ — *неубывающая функция, т. е.*

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

Доказательство. Пусть $x_2 > x_1$. Событие, состоящее в том, что X примет значение, меньшее x_2, можно подразделить на следующие два несовместных события: 1) X примет значение, меньшее x_1, с вероятностью $P(X < x_1)$; 2) X примет значение, удовлетворяющее неравенству $x_1 \le X < x_2$, с вероятностью $P(x_1 \le X < x_2)$. По теореме сложения имеем

$$P(X < x_2) = P(X < x_1) + P(x_1 \le X < x_2).$$

Отсюда

$$P(X < x_2) - P(X < x_1) = P(x_1 \le X < x_2),$$

или

$$F(x_2) - F(x_1) = P(x_1 \le X < x_2). \qquad (6.1)$$

Так как любая вероятность есть число неотрицательное, то $F(x_2) - F(x_1) \ge 0$, или $F(x_2) \ge F(x_1)$, что и требовалось доказать.

Следствие 1. *Вероятность того, что случайная величина примет значение, заключенное в интервале* (a, b)*, равна приращению функции распределения на этом интервале:*

$$P(a \le X < b) = F(b) - F(a). \qquad (6.2)$$

Это важное следствие вытекает из формулы (6.2), если положить $x_2 = b$ и $x_1 = a$.

Пример 6.1. Случайная величина X задана функцией распределения

$$F(x) = \begin{cases} 0 & \text{при} \quad x \le -1; \\ x/4 + 1/4 & \text{при} \quad -1 < x \le 3; \\ 1 & \text{при} \quad x > 3. \end{cases}$$

Найти вероятность того, что в результате испытания X примет значение, принадлежащее интервалу (0, 2):

$$P(0<X<2) = F(2) — F(0).$$

Решение. Так как на интервале (0, 2), по условию,

$$F(x) = x/4 + 1/4,$$

то

$$F(2) - F(0) = (2/4+ 1/4) - (0/4 + 1/4) = 1/2.$$

Итак,

$$P(0 < X < 2) = 1/2.$$

Следствие 2. *Вероятность того, что непрерывная случайная величина X примет одно определенное значение, равна нулю.*

Действительно, положив в формуле (6.2) $a = x_1$, $b = x_1 + \Delta x$, имеем

$$P(x_1 \le X < x_1 + \Delta x) = F(x_1 + \Delta x) - F(x_1).$$

Устремим Δx к нулю. Так как X — непрерывная случайная величина, то функция $F(x)$ непрерывна. В силу непрерывности $F(x)$ в точке x_1 разность $F(x_1 + \Delta x) - F(x_1)$ также стремится к нулю; следовательно, $P(X = x_1) = 0$. Используя это положение, легко убедиться в справедливости равенств

$$P(a \le X < b) = P(a < X < b) =$$
$$= P(a < X \le b) = P(a \le X \le b). \qquad (6.3)$$

Например, равенство $P(a < X \le b) = P(a < X < b)$ доказывается так:

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

Таким образом, не представляет интереса говорить о вероятности того, что непрерывная случайная величина примет одно определенное значение, но имеет смысл рассматривать вероятность попадания ее в интервал, пусть даже сколь угодно малый. Этот факт полностью соответствует требованиям практических задач. Например, интересуются вероятностью того, что размеры деталей не выходят за дозволенные границы, но не ставят вопроса о вероятности их совпадения с проектным размером.

Заметим, что было бы неправильным думать, что равенство нулю вероятности $P(X = x_1)$ означает, что событие $X = x_1$ невозможно (если, конечно, не ограничиваться классическим определением вероятности). Действительно, в результате испытания случайная величина обязательно примет одно из возможных значений; в частности, это значение может оказаться равным x_t.

Свойство 3. *Если возможные значения случайной величины принадлежат интервалу* (a, b), *то:* 1) $F(x)=0$ *при* $x \le a$; 2) $F(x) = 1$ *при* $x \ge b$.

Доказательство. 1) Пусть $x_1 \le a$. Тогда событие $X = x_1$ невозможно (так как значений, меньших x_1, величина X по условию не принимает) и, следовательно, вероятность его равна нулю.

2) Пусть $x_2 \ge b$. Тогда событие $X < x_2$ достоверно (так как все возможные значения X меньше x_2) и, следовательно, вероятность его равна единице.

Следствие. *Если возможные значения непрерывной случайной величины расположены на всей оси x, то справедливы следующие предельные соотношения:*

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

График функции распределения

Доказанные свойства позволяют представить, как выглядит график функции распределения непрерывной случайной величины.

График расположен в полосе, ограниченной прямыми $y=0$, $y=1$ (первое свойство).

При возрастании x в интервале (a, b), в котором заключены все возможные значения случайной величины, график «подымается вверх» (второе свойство).

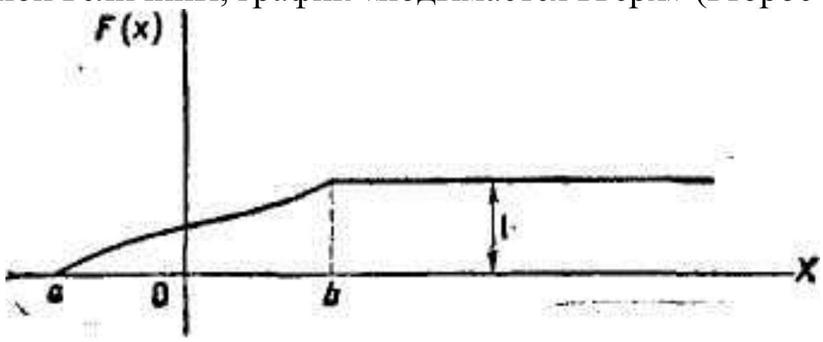

Рис. 6.1.

При $x \le a$ ординаты графика равны нулю; при $x \ge b$ ординаты графика равны единице (третье свойство).

График функции распределения непрерывной случайной величины изображен на рис. 6.1.

Замечание. График функции распределения дискретной случайной величины имеет ступенчатый вид. Убедимся в этом на примере.

Пример 6.2. Дискретная случайная величина X задана таблицей распределения

X	1	4	8
P	0,3	0,1	0,6

Найти функцию распределения и вычертить ее график.

Решение. Если $x \le 1$, то $F(x)=0$ (третье свойство).

Если $1 < x \le 4$, то $F(x)=0{,}3$. Действительно, X может принять значение 1 с вероятностью 0,3.

Если $4 < x \le 8$, то $F(x)=0{,}4$. Действительно, если x_1 удовлетворяет неравенству $4 < x_1 \le 8$, то $F(x_1)$ равно вероятности события $X < x_1$ которое может быть осуществлено, когда X примет значение 1 (вероятность этого события равна 0,3) или значение 4 (вероятность этого события равна 0,1). Поскольку эти два события несовместны, то по теореме сложения вероятность события $X < x_1$ равна сумме вероятностей 0,3+0,1=0,4.

Если $x>8$, то $F(x)=1$. Действительно, событие $X \le 8$ достоверно, следовательно, его вероятность равна единице.

Итак, функция распределения аналитически может быть записана так:

$$F(x)=\begin{cases} 0 & \text{при} \quad x \le -1, \\ 0,3 & \text{при} \quad 1 < x \le 4, \\ 0,4 & \text{при} \quad 4 < x \le 8, \\ 1 & \text{при} \quad x > 8. \end{cases}$$

График этой функции приведен на рисунке 6.2.

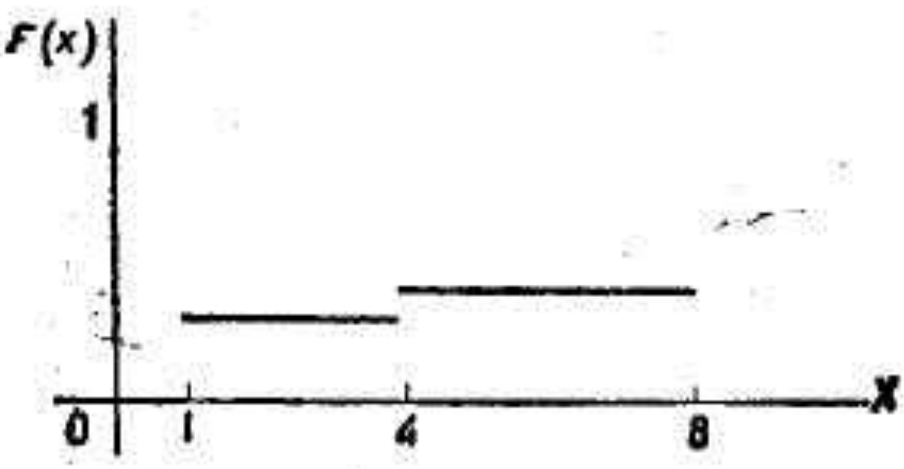

Рис. 6.2.

7. ПЛОТНОСТЬ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ НЕПРЕРЫВНОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Определение плотности распределения

Выше непрерывная случайная величина задавалась с помощью функции распределения. Этот способ задания не является единственным. Непрерывную случайную величину можно также задать, используя другую функцию, которую называют плотностью распределения или плотностью вероятности (иногда ее называют дифференциальной функцией).

Плотностью распределения вероятностей непрерывной случайной величины X называют функцию $f(x)$ — первую производную от функции распределения $F(x)$:

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

Из этого определения следует, что функция распределения является первообразной для плотности распределения.

Заметим, что для описания распределения вероятностей дискретной случайной величины плотность распределения неприменима.

Вероятность попадания непрерывной случайной величины в заданный интервал

Зная плотность распределения, можно вычислить вероятность того, что непрерывная случайная величина примет значение, принадлежащее заданному интервалу. Вычисление основано на следующей теореме.

Теорема. *Вероятность того, что непрерывная случайная величина X примет значение, принадлежащее интервалу* (*a*, *b*), *равна определенному интегралу от плотности распределения, взятому в пределах от a до b*:

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

Доказательство.

По формуле Ньютона—Лейбница,

$$F(b) - F(a) = \int_a^b F'(x)dx = \int_a^b f(x)dx.$$

Таким образом,

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

Так как $P(a \le X < b) = P(a < X < b)$, то окончательно получим

$$P(a < X < b) = \int_a^b f(x)dx. \tag{7.1}$$

Геометрически полученный результат можно истолковать так: вероятность того, что непрерывная случайная величина примет значение, принадлежащее интервалу $(a,\ b)$, равна площади криволинейной трапеции, ограниченной осью *Ox*, кривой распределения $f(x)$ и прямыми *x*=*a* и *x*=*b*.

Замечание. В частности, если $f(x)$ — четная функция и концы интервала симметричны относительно начала координат, то

$$P(-a < X < a) = P(|X| < a) = 2\int_0^a f(x)dx.$$

Пример 7.1. Задана плотность вероятности случайной величины *X*

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

Найти вероятность того, что в результате испытания *X* примет значение, принадлежащее интервалу (0,5; 1).

Решение. Искомая вероятность

$$P(0,5 < X < 1) = 2\int_{0,5}^{1} xdx = x^2\Big|_{0,5}^{1} = 1 - 0,25 = 0,75.$$

Нахождение функции распределения по известной плотности распределения

Зная плотность распределения $f(x)$, можно найти функцию распределения $F(x)$ по формуле

$$F(x)=\int_{-\infty}^{x} f(x)dx. \qquad (7.2)$$

Действительно, мы обозначили через $F(x)$ вероятность того, что случайная величина примет значение, меньшее x, т. е.

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

Очевидно, неравенство $X<x$ можно записать в виде двойного неравенства $-\infty < X < x$, следовательно,

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

Полагая в формуле (7.1) $a=-\infty$, $b=x$, имеем

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

Наконец, заменив $P(-\infty < X < x)$ на $F(x)$, в силу (7.1), окончательно получим

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

Таким образом, зная плотность распределения, можно найти функцию распределения. Разумеется, по известной функции распределения может быть найдена плотность распределения, а именно:

$$f(x)=\Gamma'(x).$$

Пример 7.2. Найти функцию распределения по данной плотности распределения:

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

Построить график найденной функции.

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

Если $x\le a$, то $f(x)=0$, следовательно, $F(x)=0$. Если $a<x\le b$, то $f(x)=1/(b-a)$, следовательно,

$$F(x)=\int_{-\infty}^{\infty} f(x)dx=\int_{-\infty}^{\infty} 0dx+\int_{a}^{x}\frac{1}{b-a}dx=\frac{x-a}{b-a}.$$

Если $x>b$, то

$$F(x)=\int_{-\infty}^{a} 0dx+\int_{a}^{b}\frac{dx}{b-a}+\int_{b}^{x}0dx=\frac{b-a}{b-a}=1.$$

Итак, искомая функция распределения

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

График этой функции изображен на рис. 7.1.

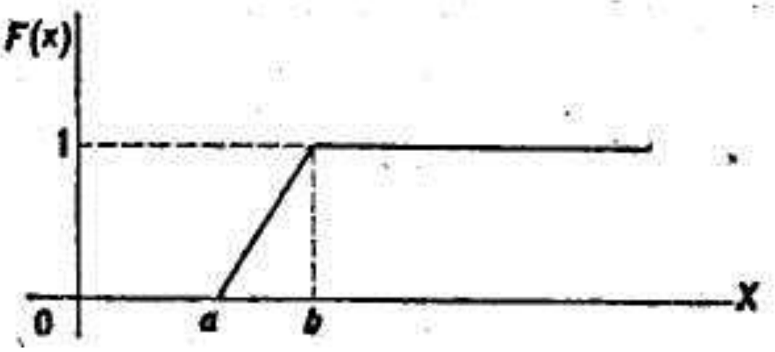

Рис. 7.1.

Свойства плотности распределения

Свойство 1. *Плотность распределения—неотрицательная функция:*

$$f(x) > 0.$$

Доказательство. Функция распределения — неубывающая функция, следовательно, ее производная $F'(x)=f(x)$ — функция неотрицательная.

Геометрически это свойство означает, что точки, принадлежащие графику плотности распределения, расположены либо над осью *Ox,* либо на этой оси.

График плотности распределения называют *кривой распределения.*

Свойство 2. *Несобственный интервал от плотности распределения в пределах от $-\infty$ до ∞ равен единице:*

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

Доказательство. Несобственный интеграл $\int_{-\infty}^{\infty} f(x)dx$ выражает вероятность события, состоящего в том, что случайная величина примет значение, принадлежащее интервалу $(-\infty,\ \infty)$. Очевидно, такое событие достоверно, следовательно, вероятность его равна единице.

Геометрически это означает, что вся площадь криволинейной трапеции, ограниченной осью *Ox* и кривой распределения, равна единице.

В частности, если все возможные значения случайной величины принадлежат интервалу $(a,\ b)$, то

$$\int_{a}^{b} f(x)dx=1.$$

Пример 7.3. Плотность распределения случайной величины *X* задана:

$$f(x)=\frac{a}{e^{-x}+e^{x}}.$$

Найти постоянный параметр *a*.

Решение. Плотность распределения должна удовлетворять условию $\int\limits_{-\infty}^{\infty} f(x)dx=1$, поэтому потребуем, чтобы выполнялось равенство

$$a\int\limits_{-\infty}^{\infty}\frac{dx}{e^{-x}+e^{x}}=1.$$

Отсюда

$$a=\frac{1}{\int\limits_{-\infty}^{\infty}\frac{dx}{e^{-x}+e^{x}}}.$$

Найдем неопределенный инетграл:

$$\int\frac{dx}{e^{-x}+e^{x}}=\int\frac{e^{x}dx}{1+e^{2x}}=arctge^{x}.$$

Вычислим несобственный интеграл:

$$a\int\limits_{-\infty}^{\infty}\frac{dx}{e^{-x}+e^{x}}=\lim_{b\to-\infty}\int\limits_{b}^{0}\frac{dx}{e^{-x}+e^{x}}+\lim_{c\to\infty}\int\limits_{0}^{c}\frac{dx}{e^{-x}+e^{x}}=$$

$$=\lim_{b\to-\infty}\left(-arctge^{b}\right)+\lim_{c\to\infty}\left(arctge^{c}\right)=\pi/2.$$

Таким образом, искомый параметр

$$a=\frac{1}{\pi/2}=\frac{2}{\pi}.$$

Вероятностный смысл плотности распределения

Пусть $F(x)$ — функция распределения непрерывной случайной величины X. По определению плотности распределения, $f(x) = F'(x)$, или в иной форме

Как уже известно, разность $F(x+\Delta x)-F(x)$ определяет вероятность того, что X примет значение, принадлежащее интервалу $(x,\ x+\Delta x)$. Таким образом, предел отношения вероятности того, что непрерывная случайная величина примет значение, принадлежащее интервалу $(x,\ x+\Delta x)$, к длине этого интервала (при $\Delta x\to 0$) равен значению плотности распределения в точке x.

По аналогии с определением плотности массы в точке [1] целесообразно рассматривать значение функции $f(x)$ в точке x как плотность вероятности в этой точке.

Итак, функция $f(x)$ определяет плотность распределения вероятности для каждой точки x.

Из дифференциального исчисления известно, что приращение функции приближенно равно дифференциалу функции, т. е.

$$F\left(x+\Delta x\right)-F\left(x\right)\approx dF\left(x\right),$$

или

$$F\left(x+\Delta x\right)-F\left(x\right)\approx F'\left(x\right)dx.$$

Так как $F'(x) = f(x)$ и $dx=\Delta x$, то

$$F\left(x+\Delta x\right)-F\left(x\right)\approx f\left(x\right)dx.$$

Вероятностный смысл этого равенства таков: вероятность того, что случайная величина примет значение, принадлежащее интервалу $\left(x,\ x+\Delta x\right)$, приближенно равна (с точностью до бесконечно малых высшего порядка относительно Δx) произведению плотности вероятности в точке x на длину интервала Δx.

Геометрически этот результат можно истолковать так: вероятность того, что случайная величина примет значение, принадлежащее интервалу $\left(x,\ x+\Delta x\right)$, приближенно равна площади прямоугольника с основанием Δx и высотой $f\left(x\right)$.

На рис. 7.2 видно, что площадь заштрихованного прямоугольника, равная произведению $f\left(x\right)\Delta x$, лишь приближенно равна площади криволинейной трапеции (истинной вероятности, определяемой определенным интегралом $\int\limits_{x}^{x+\Delta x} f\left(x\right)\Delta x$. Допущенная при этом погрешность равна площади криволинейного треугольника ABC.

[1] Если масса непрерывно распределена вдоль оси x по некоторому закону, например $F(x)$, то плотностью $\rho\left(x\right)$ массы в точке x называют предел отношения массы интервала $(x,\ x + Ax)$ к длине интервала при $\Delta x\to 0$, т.е. $\rho\left(x\right)=\lim\limits_{\Delta x\to 0}\dfrac{F\left(x+\Delta x\right)-F\left(x\right)}{\Delta x}$.

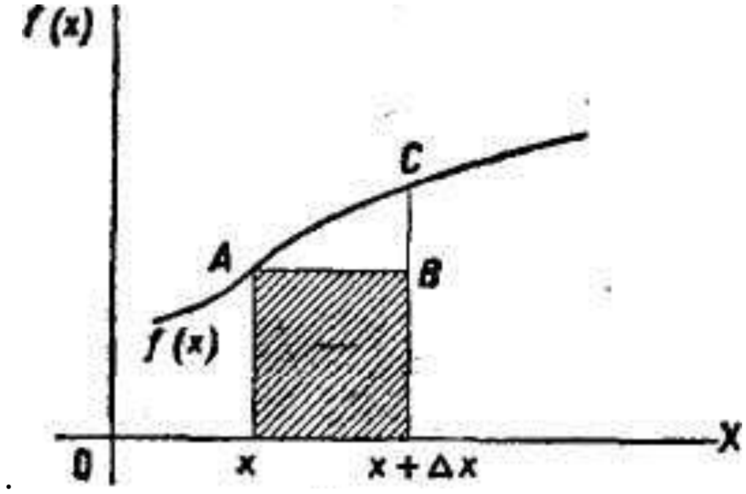

Рис. 7.2.

Закон равномерного распределения вероятностей

При решении задач, которые выдвигает практика, приходится сталкиваться с различными распределениями непрерывных случайных величин. Плотности распределений непрерывных случайных величин называют также *законами распределений.* Часто встречаются, например, законы равномерного, нормального и показательного распределений. В настоящем параграфе рассматривается закон равномерного распределения вероятностей. Нормальному и показательному законам посвящены следующие две главы.

Распределение вероятностей называют *равномерным,* если на интервале, которому принадлежат все возможные значения случайной величины, плотность распределения сохраняет постоянное значение.

Приведем пример равномерно распределенной непрерывной случайной величины.

Пример 7.4. Шкала измерительного прибора проградуирована в некоторых единицах. Ошибку при округлении отсчета до ближайшего целого деления можно рассматривать как случайную величину *X,* которая может принимать с постоянной плотностью вероятности любое значение между двумя соседними целыми делениями. Таким образом, *X* имеет равномерное распределение.

Найдем плотность равномерного распределения $f(x)$, считая, что все возможные значения случайной величины заключены в интервале (a, b), на котором функция $f(x)$ сохраняет постоянные значения.

По условию, *X* не принимает значений вне интервала (a, b), поэтому $f(x)=0$ при $x < a$ и $x > b$.

Найдем постоянную *C*. Так как все возможные значения случайной величины принадлежат интервалу (a, b), то должно выполняться соотношение

$$\int_a^b f(x)dx = 1, \qquad \text{или} \qquad \int_a^b Cdx = 1.$$

Отсюда

$$C = 1 \Big/ \int_a^b C dx = 1/(b-a).$$

Итак, искомая плотность вероятности равномерного распределения

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

График плотности равномерного распределения изображен на рис. 7.3, а график функции распределения — на рис. 7.1.

З а м е ч а н и е. Обозначим через R непрерывную случайную величину, распределенную равномерно в интервале (0, 1), а через r — ее возможные значения. Вероятность попадания величины R (в результате испытания) в интервал (с, d), принадлежащий интервалу (0, 1), равна его длине:

$$P(c < R < d) = d—c.$$

Действительно, плотность рассматриваемого равномерного распределения

$$f(r) = 1/(1-0) = 1.$$

Следовательно, вероятность попадания случайной величины R в интервал (с, d)

$$P(c < R < d) = \int_c^d f(r) dr = \int_c^d 1 \cdot dr = d - c.$$

8. НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Числовые характеристики непрерывных случайных величин

Распространим определения числовых характеристик дискретных величин на величины непрерывные. Начнем с математического ожидания.

Пусть непрерывная случайная величина X задана плотностью распределения f (x). Допустим, что все возможные значения X принадлежат отрезку $[a, b]$. Разобьем этот отрезок на n частичных отрезков длиной $\Delta x_1, \Delta x_2, ..., \Delta x_n$ и выберем в каждом из них произвольную точку x_i $(i = 1, 2, ..., n)$. Нам надо определить математическое ожидание непрерывной величины по аналогии с дискретной; составим сумму произведений возможных значений x_i на вероятности попадания их в интервал Δx_i (напомним, что произведение $f(x)\Delta x$ приближенно равно вероятности попадания X в интервал Δx:

$$\sum x_i f(x_i) \Delta x_i .$$

Перейдя к пределу при стремлении к нулю длины наибольшего из частичных отрезков, получим определенный интеграл $\int_a^b x f(x) dx$.

Математическим ожиданием непрерывной случайной величины X, возможные значения которой принадлежат отрезку $[a,\ b]$, называют определенный интеграл

$$M(X)=\int_a^b xf(x)dx. \tag{8.1}$$

Если возможные значения принадлежат всей оси *Ox,* то

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

Предполагается, что несобственный интеграл сходится абсолютно, т. е. существует интеграл $\int_{-\infty}^{\infty}|x|\,f(x)dx$. Если бы это требование не выполнялось, то значение интеграла зависело бы от скорости стремления (в отдельности) нижнего предела к $-\infty$, а верхнего — к $+\infty$.

По аналогии с дисперсией дискретной величины определяется и дисперсия непрерывной величины.

Дисперсией непрерывной случайной величины называют математическое ожидание квадрата ее отклонения.

Если возможные значения *X* принадлежат отрезку $[a,\ b]$, то

$$D(X)=\int_a^b\left[x-M(X)\right]^2 f(x)dx;$$

если возможные значения принадлежат всей оси *x,*

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

Среднее квадратическое отклонение непрерывной случайной величины определяется, как и для величины дискретной, равенством

Замечание 1. Можно доказать, что свойства математического ожидания и дисперсии дискретных величин сохраняются и для непрерывных величин.

Замечание 2. Легко получить для вычисления дасдерсин более удобные формулы:

$$D(X)=\int_a^b x^2 f(x)dx\left[M(X)\right]^2, \tag{8.2}$$

Пример 8.1. Найти матеиатическое ожидание и дисперсию случайной величины *X*, заданной функцией распределения

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

Решение. Найдем плотность распределения:

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

Найдем математическое ожидание по формуле (8.1):

$$M(X)=\int_0^1 x\cdot 1\cdot dx = x^2/2\Big|_0^1=1/2 .$$

Найдем дисперсию по формуле (8.2):

$$D(X)=\int_0^1 x\cdot 1\cdot dx-[1/2]^2 = x^3/3\Big|_0^1-1/4=1/12 .$$

Пример 8.2. Найти математическое ожидание и дисперсию непрерывной случайной величины *X*, распределенной равномерно в интервале $(a,\ b)$.

Решение. Найдем математическое ожидание *X* по формуле (8.1), учитывая, что плотность равномерного распределения $f(x)=1/(b-a)$

$$D(X)=\int_a^b xf(x)dx=\frac{1}{b-a}\int_a^b xdx .$$

Выполнив элементарные выкладки, получим

$$M(X)=(a+b)/2 .$$

Найдем дисперсию *X* по формуле (8.2):

$$D(X)=\int_a^b x^2 f(x)dx-\left[M(X)\right]^2=\frac{1}{b-a}\int_a^b x^2dx-\left[\frac{a+b}{2}\right]^2 .$$

Выполнив элементарные выкладки, получим

$$D(X)=(a+b)^2/12 .$$

З а м е ч а н и е 3. Математическое ожидание и дисперсия случайной величины *R*, распределенной равномерно в интервале (0, 1), т. е. если a=0, b=1, как следует из примера 8.2, соответственно равны $M(R)=1/2$, $D(R)=1/12$. Этот же результат мы получили в примере 8.1 по заданной функции распределения случайной величины *R*.

Нормальное распределение

Нормальным называют распределение вероятностей непрерывной случайной величины, которое описывается плотностью

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

Мы видим, что *нормальное распределение определяется двумя параметрами: a и σ*. Достаточно знать эти параметры, чтобы задать нормальное распределение. Покажем, что вероятностный смысл этих параметров таков: *a* есть математическое ожидание, σ — среднее квадратическое отклонение нормального распределения.

а) По определению математического ожидания непрерывной случайной величины,

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

Введем новую переменную $z=\left(x-a\right)/\sigma$. Отсюда $x=\sigma z+a$, $dx=\sigma dz$. Приняв во внимание, что новые пределы инегрирования равны старым, получим

$$M\left(X\right)=\frac{\sigma}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\left(\sigma z+a\right)e^{-z^2/2}dz=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sigma ze^{-z^2/2}dz+\frac{a}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-z^2/2}dz.$$

Первое из слагаемых равно нулю (под знаком интеграла нечетная функция; пределы интегрирования симметричны относительно начала координат). Второе из слагаемых равно *a* (интеграл Пуассона $\int_{-\infty}^{\infty}e^{-z^2/2}dz=\sqrt{2\pi}$).

Итак, $M\left(X\right)=a$, т. е. *математическое ожидание нормального распределения равно параметру а.*

б) По определению дисперсии непрерывной случайной величины, учитывая, что $M\left(X\right)=a$, имеем

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

Введем новую переменную $z=\left(x-a\right)/\sigma$. Отсюда $x-a=\sigma z$, $dx=\sigma dz$. Приняв во внимание, что новые пределы интегрирования равны старым, получим

$$D\left(X\right)=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}z\cdot ze^{-z^2/2}dz.$$

Интегрируя по частям, положив $u=z$, $dv=z^{-z^2/2}dz$, найдем

$$D\left(X\right)=\sigma^2.$$

Следовательно,

$$\sigma\left(X\right)=\sqrt{D\left(X\right)}=\sqrt{\sigma^2}=\sigma.$$

Итак, *среднее квадратическое отклонение нормального распределения равно параметру σ.*

Вероятность попадания в заданный интервал нормальной случайной величины

Уже известно, что если случайная величина *X* задана плотностью распределения *f*(*x*), то вероятность того, что *X* примет значение, принадлежащее интервалу $\left(\alpha,\ \beta\right)$, такова:

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x)\,dx.$$

Пусть случайная величина *X* распределена по нормальному закону. Тогда вероятность того, что *X* примет значение, принадлежащее интервалу $(\alpha,\ \beta)$, равна

$$P(\alpha < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-(x-a)^2/(2\sigma^2)} dx.$$

Преобразуем эту формулу так, чтобы можно было пользоваться готовыми таблицами. Введем новую переменную $z = (x-a)/\sigma$. Отсюда $x = \sigma z + a$, $dx = \sigma dz$. Найдем новые пределы интегрирования. Если $x = \alpha$, то $z = (\alpha - a)/\sigma$; если $x = \beta$, то $z = (\beta - a)/\sigma$.

Таким образом, имеем

$$P(\alpha < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{(\alpha-a)/\sigma}^{(\beta-a)/\sigma} e^{-z^2/2} (\alpha dz) =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{(\alpha-a)/\sigma}^{0} e^{-z^2/2} dz + \frac{1}{\sqrt{2\pi}} \int_{0}^{(\beta-a)/\sigma} e^{-z^2/2} dz =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{0}^{(\beta-a)/\sigma} e^{-z^2/2} dz - \frac{1}{\sqrt{2\pi}} \int_{0}^{(\alpha-a)/\sigma} e^{-z^2/2} dz.$$

Пользуясь функцией Лапласа

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-z^2/2} dz,$$

окончательно получим

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right). \qquad (8.3)$$

Пример 8.3. Случайная величина *X* распределена по нормальному закону. Математическое ожидание и среднее квадрахическое отклонение этой величины соответственно равны 30 и 10. Найти вероятность того, что *X* примет значение, принадлежащее интервалу (10, 50).

Решение. Воспользуемся формулой (*). По условию, $\alpha = 10$, $\beta = 50$, *a* = 30, $\sigma = 10$, следовательно,

$$P(10 < X < 50) = \Phi\left(\frac{50-30}{10}\right) - \Phi\left(\frac{10-30}{10}\right) = 2\Phi(2).$$

По таблице приложения 2 находим Ф(2)=0,4772. Отсюда искомая вероятность

$$P(10 < X < 50) = 2 \cdot 0{,}4772 = 0{,}9544.$$

Вычисление вероятности заданного отклонения

Часто требуется вычислить вероятность того, что отклонение нормально распределенной случайной величины X по абсолютной величине меньше заданного положительного числа δ, т. е. требуется найти вероятность осуществления неравенства $|X-a|<\delta$.

Заменим это неравенство равносильным ему двойным неравенством
$-\delta<X-a<\delta$, или $a-\delta<X<a+\delta$.

Пользуясь формулой (8.3), получим

$$P\left(|X-a|<\delta\right)=P\left(a-\delta<X<a+\delta\right)=$$

$$=\Phi\left(\frac{(a+\delta)-a}{\sigma}\right)-\Phi\left(\frac{(a-\delta)-a}{\sigma}\right)=\Phi\left(\frac{\delta}{\sigma}\right)-\Phi\left(-\frac{\delta}{\sigma}\right).$$

Приняв во внимание равенство

$$\Phi\left(-\delta/\sigma\right)=-\Phi\left(\delta/\sigma\right)$$

(функция Лапласа — нечетная), окончательно имеем

$$\Phi\left(|X-a|<\delta\right)=2\Phi\left(\delta/\sigma\right).$$

В частности, при $a=0$

$$\Phi\left(|X|<\delta\right)=2\Phi\left(\delta/\sigma\right).$$

Пример 8.4. Случайная величина X распределена нормально. Математическое ожидание и среднее квадратическое отклонение X соответственно равны 20 и 10. Найти вероятность того, что отклонение по абсолютной величине будет меньше трех.

Решение. Воспользуемся формулой

$$\Phi\left(|X-a|<\delta\right)=2\Phi\left(\delta/\sigma\right).$$

По условию, $\delta=3$, a=20, $\sigma=10$. Следовательно,

$$\Phi\left(|X-20|<3\right)=2\Phi\left(3/10\right)=2\Phi\left(0,3\right).$$

По таблице приложения 2 находим $\Phi\left(0,3\right)=0,1179$.

Искомая вероятность

$$\Phi\left(|X-20|<3\right)=0,2358.$$

Правило трех сигм

Преобразуем формулу $P\left(|X-a|<\delta\right)=2\Phi\ \left(\delta/\sigma\right)$, положив $\delta=\sigma t$. В итоге получим

$$P\left(|X-a|<\sigma t\right)=2\Phi\ \left(t\right).$$

Если $t=3$ и, следовательно, $\sigma t=3\sigma$, то

$$P\left(|X-a|<3\sigma\right)=2\Phi\ \left(3\right)=2\cdot 0,49865=0,9973,$$

т.е. вероятность того, что отклонение по абсолютной величине будет меньше утроенного среднего квадратичского отклонения, равна 0,9973.

Другими словами, вероятность того, что абсолютная величина отклонения превысит утроенное среднее квадратическое отклонение, очень мала, а именно равна 0,0027. Это означает, что лишь в 0,27% случаев так может произойти. Такие события, исходя из принципа невозможности маловероятных событий, можно считать практически невозможными. В этом и состоит сущность правила трех сигм: *если случайная величина распределена нормально, то абсолютная величина ее отклонения от математического ожидания не превосходит утроенного среднего квадратического отклонения.*

На практике правило трех сигм применяют так: если распределение изучаемой случайной величины неизвестно, но условие, указанное в приведенном правиле, выполняется, то есть основание предполагать, что изучаемая величина распределена нормально; в противном случае она не распределена нормально.

9. НЕРАВЕНСТВО ЧЕБЫШЕВА. ТЕОРЕМЫ ЧЕБЫШЕВА И БЕРНУЛЛИ

Теорема Чебышева

Неравенство Чебышева справедливо для дискретных и непрерывных случайных величин. Для простоты ограничимся доказательством этого неравенства для дискретных величин.

Рассмотрим дискретную случайную величину *X*, заданную таблицей распределения:

X	x_1	x_2	$\ldots$	x_n
P	p_1	p_2	$\ldots$	p_n

Поставим перед собой задачу оценить вероятность того, что отклонение случайной величины от ее математического ожидания не превышает по абсолютной величине положительного числа е. Если е достаточно мало, то мы оценим, таким образом, вероятность того, что *X* примет значения, достаточно близкие к своему математическому ожиданию. П. Л. Чебышев доказал неравенство, позволяющее дать интересующую нас оценку. (Приводится без доказательства).

Неравенство Чебышева. *Вероятность того, что отклонение случайной величины X от ее математического ожидания по абсолютной величине меньше положительного числа* ε*, не меньше, чем* $1-D(X)/\varepsilon^2$:

$$P\left(\left|X-M(X)\right|<\varepsilon\right)\geq 1-D(X)/\varepsilon^2 .$$

Теорема Чебышева

Теорема Чебышева. *Если* X_1, X_2, …, X_n, ...*— попарно независимые случайные величины, причем дисперсии их равномерно ограничены (не превышают постоянного числа С), то, как бы мало ни было положительное число* е, *вероятность неравенства*

$$\left|\frac{X_1+X_2+...+X_n}{n}-\frac{M\left(X_1\right)+M\left(X_2\right)+...+M\left(X_n\right)}{n}\right|<\varepsilon$$

будет как угодно близка к единице, если число случайных величин достаточно велико.

Другими словами, в условиях теоремы

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

Таким образом, теорема Чебышева утверждает, что если рассматривается достаточно большое число независимых случайных величин, имеющих ограниченные дисперсии, то почти достоверным можно считать событие, состоящее в том, что отклонение среднего арифметического случайных величин от среднего арифметического их математических ожиданий будет по абсолютной величине сколь угодно малым.

Формулируя теорему Чебышева, мы предполагали, что случайные величины имеют различные математические ожидания. На практике часто бывает, что случайные величины имеют одно и то же математическое ожидание. Очевидно, что если вновь допустить, что дисперсии этих величин ограничены, то к ним будет применима теорема Чебышева.

Обозначим математическое ожидание каждой из случайных величин через *a;* в рассматриваемом случае среднее арифметическое математических ожиданий, как легко видеть, также равно *a*. Мы можем сформулировать теорему Чебышева для рассматриваемого частного случая. *Если* $X_1+X_2+...+X_n$, ... — *попарно независимые случайные величины, имеющие одно и то же математическое ожидание а, и если дисперсии этих величин равномерно ограничены, то, как бы мало ни было число* $\varepsilon>0$, *вероятность неравенства*

$$\left|\frac{X_1+X_2+...+X_n}{n}-a\right|<\varepsilon$$

будет как угодно близка к единице, если число случайных величин достаточно велико.

Другими словами, в условиях теоремы будет иметь место равенство

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

Сущность доказанной теоремы такова: хотя отдельные независимые случайные величины могут принимать значения, далекие от своих математических ожиданий, среднее арифметическое достаточно большого числа случайных величин с большой вероятностью принимает значения, близкие к определенному постоянному числу, а именно к числу $\left(M\left(X_1\right)+M\left(X_2\right)+...+M\left(X_n\right)\right)/n$ (или к числу *a* в частном случае). Иными словами, отдельные случайные величины могут иметь значительный разброс, а их среднее арифметическое рассеянно мало.

Теорема Бернулли

Пусть производится я независимых испытаний, в каждом из которых вероятность появления события *A* равна *p*. Можно ли предвидеть, какова примерно будет относительная частота появлений события? Положительный ответ на этот вопрос дает теорема, доказанная Якобом Бернулли (опубликована в 1713 г.), которая получила название «закона больших чисел» и положила начало теории вероятностей как науке. Доказательство Бернулли было сложным; простое доказательство дано П. Л. Чебышевым в 1846 г.

Теорема Бернулли. *Если в каждом из n независимых испытаний вероятность p появления события A постоянна, то как угодно близка к единице вероятность того, что Отклонение относительной частоты от вероятности p по абсолютной величине будет сколь угодно малым если число испытаний достаточно велико. (Без доказательства).*

Другими словами, если ε —сколь угодно малое положительное число, то при соблюдении условий теоремы имеет место равенство

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

10. ВЫБОРОЧНЫЙ МЕТОД МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ. ХАРАКТЕРИСТИКИ ВЫБОРКИ

Установление закономерностей, которым подчинены массовые случайные явления, основано на изучении методами теории вероятностей статистических данных — результатов наблюдений.

Первая задача математической статистики—указать способы сбора и группировки статистических сведений, полученных в результате наблюдений или в результате специально поставленных экспериментов.

Вторая задача математической статистики — разработать методы анализа статистических данных в зависимости от целей исследования. Сюда относятся:

а) оценка неизвестной вероятности события; оценка неизвестной функции распределения; оценка параметров распределения, вид которого известен; оценка зависимости случайной величины от одной или нескольких случайных величин и др.;

б) проверка статистических гипотез о виде неизвестного распределения или о величине параметров распределения, вид которого известен.

Задача математической статистики состоит в создании методов сбора и обработки статистических данных для получения научных и практических выводов.

Генеральная и выборочная совокупности

Пусть требуется изучить совокупность однородных объектов относительно некоторого качественного или количественного признака, характеризующего эти объекты. Например, если имеется партия деталей, то качественным признаком может служить, стандартность детали, а количественным — контролируемый размер детали.

Выборочной совокупностью или просто *выборкой* называют совокупность случайно отобранных объектов.

Генеральной совокупностью называют совокупность Объектов, из которых производится выборка.

Объемом совокупности (выборочной или генеральной) называют число объектов этой совокупности. Например, если из. 1000 деталей отобрано для обследования 100 деталей, то объем генеральной совокупности N=1000, а объем выборки n=100.

Повторная и бесповторная выборки. Репрезентативная выборка

При составлении выборки можно поступать двумя Способами: после того как объект отобран и над ним произведено наблюдение, он может быть возвращен либо не возвращен в генеральную совокупность. В соответствии со сказанным выборки подразделяют на повторные и бесповторные.

Повторной называют выборку, при которой отобранный объект (перед отбором следующего) возвращается в генеральную совокупность.

Бесповторной называют выборку, при которой отобранный объект в генеральную совокупность не возвращается.

На практике обычно пользуются бесповторным случайным отбором.

Статистическое распределение выборки

Пусть из генеральной совокупности извлечена выборка, причем x_1 наблюдалось n_1 раз, $x_2 - n_2$ раз, $x_k - n_k$ раз и $\sum n_i = n$ — обем выборки. Наблюдаемые значения x_i называют *вариантами,* а последовательность вариант, записанных в возрастающем порядке,— *вариационным рядом.* Числа наблюдений называют *частотами,* а их отношения к объему выборки $n_i / n = W_i$ — *относительными частотами.*

Статистическим распределением выборки называют перечень вариант и соответствующих им частот или относительных частот. Статистическое распределение можно задать также в виде последовательности интервалов и соответствующих им частот (в качестве частоты, соответствующей интервалу, принимают сумму частот, попавших в этот интервал).

Заметим, что в теории вероятностей под *распределением* понимают соответствие между возможными значениями случайной величины и их вероятностями, а в математической статистике — соответствие между наблюдаемыми вариантами и их частотами, или относительными частотами.

Пример 10.1. Задано распределение частот выборки объема n = 20:

x_i	2	6	12
n_i	3	10	7

Написать распределение относительных частот.

Решение, Найдем относительные частоты, для чего разделим частоты на объем выборки:

$W_1 = 3/20 = 0{,}15$, $W_2 = 10/20 = 0{,}50$, $W_3 = 7/20 = 0{,}35$.

Напишем распределение относительных частот:

x_i	2	6	12
W_i	0,15	0,50	0,35

Контроль: 0,15 + 0,50+0,35=1.

Эмпирическая функция распределения

Пусть известно статистическое распределение частот количественного признака X. Введем обозначения: n_x—число наблюдений, при которых наблюдалось значение признака, меньшее x; n—общее число наблюдений (объем выборки). Ясно, что относительная частота события $X < x$ равна n_x/n. Если x изменяется, то, вообще говоря, изменяется и относительная частота, т. е. относительная частота n_x/n естаг функция от x. Так как эта функция находится эмпирическим (опытным) путем, то ее называют эмпирической.

Эмпирической функцией распределения (функцией распределения выборки) называют функцию $F^*(x)$, определяющую для каждого значения x относительную частоту события $X < x$.

Итак, по определению,

$$F^*(x) = n_x/n,$$

где n_x — число вариант, меньших x; n — объем выборки. Таким образом, для того чтобы найти, например, $P^*(x_2)$, надо число вариант, меньших x_2, разделить на объем выборки:

$$F^*(x_2) = n_{x_2}/n.$$

В отличие от эмпирической функции распределения выборки функцию распределения $F(x)$ генеральной совокупности называют *теоретической функцией распределения.* Различие между эмпирической и теоретической функциями состоит в том, что теоретическая функция $F(x)$ определяет вероятность события $X < x$, а эмпирическая функция $F^*(x)$ определяет относительную частоту этого же события. Из теоремы Бернулли следует, что относительная частота события $X < x$, т. е. $F^*(x)$ стремится по вероятности к вероятности $F(x)$ этого события. Другими словами, при больших n числа $F^*(x)$ и $F(x)$ мало отличаются одно от другого в том смысле, что $\lim\limits_{n\to\infty} P\left[\left|F(x) - F^*(x)\right| < \varepsilon\right] = 1(\varepsilon > 0)$. Уже отсюда следует целесообразность использования эмпирической функции распределения выборки для приближенного представления теоретической (интегральной) функции распределения генеральной совокупности.

Эмпирическая функция распределения выборки служит для оценки теоретической функции распределения генеральной совокупности.

Пример 10.2. Построить империческую функцию по данному распределению выборки:

варианты x_i	2	6	10
частоты n_i	12	18	30

Решение. Нйдем объем выборки: 12+18+30=60. Наименьшая варианта 2, следовательно,

$$F^*(x) = 0 \text{ при } x \le 2.$$

Значение $X < 6$, а именно $x_1 = 2$, наблюдалось 12 раз, следовательно,

$$F^*(x) = 12/60 = 0{,}2, \text{ при } 2 < x \le 6.$$

Значения $X < 10$, а именно $x_1 = 2$ и $x_2 = 6$, наблюдались 12+18=30 раз, следовательно,

$$F^*(x) = 30/60 = 0{,}5 \text{ при } 6 < x \le 10.$$

Так как $x = 10$ - наибольшая варианта, то

$$F^*(x) = 1 \text{ при } x > 10.$$

Искомая эмпирическая функция

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

График этой функции изображен на рис. 10.1.

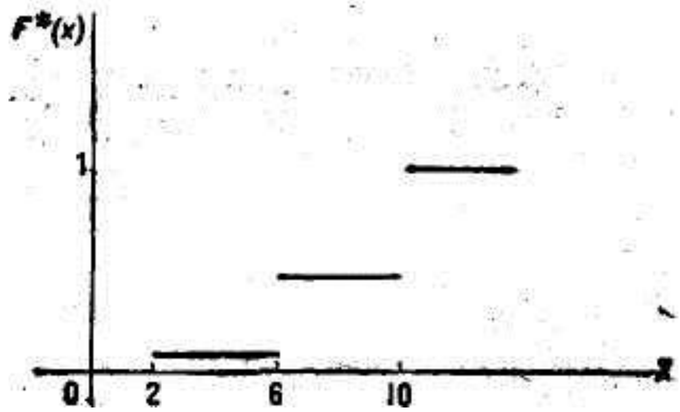

Рис. 10.1.

Полигон и гистограмма

Для наглядности строят различные графики статистического распределения и, в частности, полигон и гистограмму.

Полигоном частот называют ломаную, отрезки которой соединяют точки $(x_1;\, n_1)$, $(x_2;\, n_2)$, ..., $(x_k;\, n_k)$. Для построения полигона частот на оси абсцисс откладывают варианты x_i, а на оси ординат — соответствующие им частоты n_i. Точки $(x_i;\, n_i)$ соединяют отрезками прямых и получают полигон частот.

Полигоном относительных частот называют ломаную, отрезки которой соединяют точки $(x_1;\, W_1)$, $(x_2;\, W_2)$, ..., $(x_k;\, W_k)$. Для построения полигона относительных частот на оси абсцисс откладывают варианты x_i а на оси ординат — соответствующие им относительные частоты W_i. Точки $(x_i;\ W_i)$ соединяют отрезками прямых и получают полигон относительных частот.

На рис. 10.2 изображен полигон относительных частот следующего распределения:

X	1,5	3,5	5,5	7,5
W	0,1	0,2	0,4	0,3

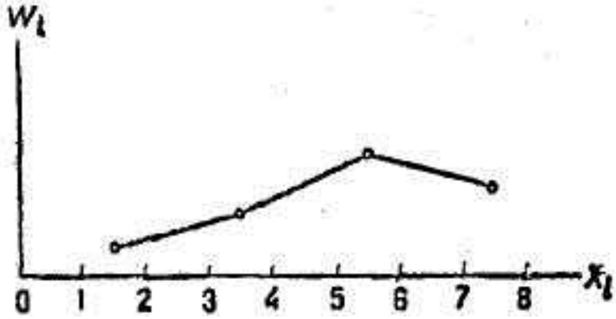

Рис. 10.2

В случае непрерывного признака целесообразно строить гистограмму, для чего интервал, в котором заключены все наблюдаемые значения признака, разбивают на несколько частичных интервалов длиной h и находят для каждого частичного интервала n_i — сумму частот вариант, попавших в i-й интервал.

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

Для построения гистограммы частот на оси абсцисс откладывают частичные интервалы, а над ними проводят отрезки, параллельные оси абсцисс на расстоянии n_i/h.

Площадь i-го частичного прямоугольника равна $hn_i/h = n_i$ — сумме частот вариант i-го интервала; следовательно, *площадь гистограммы частот равна сумме всех частот, т. е. объему выборки.*

На рис. 10.3 изображена гистограмма частот распределения объема n=100, приведенного в табл. 10.1.

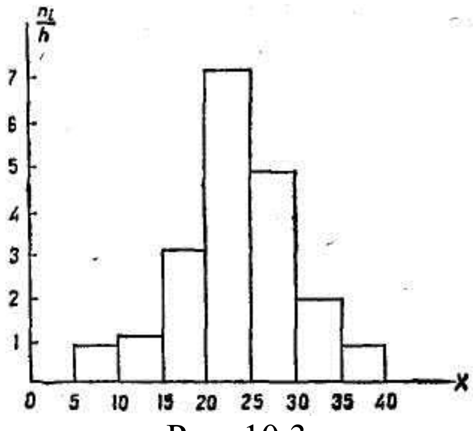

Рис. 10.3.

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

Для построения гистограммы относительных частот на оси абсцисс откладывают частичные интервалы, а над ними проводят отрезки,

Таблица 10.1

Частичный интервал длинною h=5	Сумма частот вариант частичного интервала n_i	Плотность частоты n_i/h
5—1.0	4	0,8
10—15	6	1,2
15-20	16	3,2
20—25	36	7,2
25—30	24	4,8
30—35	10	2,0
35—40	4	0,8

параллельные оси абсцисс на расстоянии W_i/h. Площадь i-го частичного прямоугольника равна $hW_i/h = W_i$ — относительной частоте вариант, попавших в i-й интервал. Следовательно, *площадь гистограммы относительных частот равна сумме всех относительных частот, т. е. единице.*

ЛИТЕРАТУРА

1. Большакова Л.В. Теория вероятностей для экономистов: учеб. Пособие /Л.В. Большакова. – М.: Финансы и статистика, 2009.-208 с.: ил.
2. Бондаренко В.Н. Курс теории вероятностей (задачи и упражнения): учебное пособие. – М.: МГИУ, 2007. – 100с.
3. Гмурман В.Е. Руководство по решению задач по теории вероятностей и математической статистике: Учеб. Пособие. 11-е изд., перераб. – М.: Высшее образование, 2009. – 404 с.

ПРИЛОЖЕНИЯ

Приложение 1

Таблица значений функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-x^2}{2}}$

	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3652	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0.9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127

	0	1	2	3	4	5	6	7	8	9
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0043
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Приложение 2

Таблица значений функции $\Phi\left(x\right)=\dfrac{1}{\sqrt{2\pi}}\cdot\int\limits_0^x e^{\frac{-z^2}{2}}\,dz$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000	0,32	0,1255	0,64	0,2389	0,96	0,3315
0,01	0,0040	0,33	0,1293	0,65	0,2422	0,97	0,3340
0,02	0,0080	0,34	0,1331	0,66	0,2454	0,98	0,3365
0,03	0,0120	0,35	0,1368	0,67	0,2486	0,99	0,3389
0,04	0,0160	0,36	0,1406	0,68	0,2517	1,00	0,3413
0,05	0,0199	0,37	0,1443	0,69	0,2549	1,01	0,3438
0,06	0,0239	0,38	0,1480	0,70	0,2580	1,02	0,3461
0,07	0,0279	0,39	0,1517	0,71	0,2611	1,03	0,3485
0,08	0,0319	0,40	0,1554	0,72	0,2642	1,04	0,3508
0,09	0,0359	0,41	0,1591	0,73	0,2673	1,05	0,3531
0,10	0,0398	0,42	0,1628	0,74	0,2703	1,06	0,3554
0,11	0,0438	0,43	0,1664	0,75	0,2734	1,07	0,3577
0,12	0,0478	0,44	0,1700	0,76	0,2764	1,08	0,3599
0,13	0,0517	0,45	0,1736	0,77	0,2794	1,09	0,3621
0,14	0,0557	0,46	0,1772	0,78	0,2823	1,10	0,3643
0,15	0,0596	0,47	0,1808	0,79	0,2852	1,11	0,3665
0,16	0,0636	0,48	0,1844	0.80	0,2881	1,12	0,3686
0,17	0,0675	0,49	0,1879	0,81	0,2910	1,13	0,3708
0,18	0,0714	0,50	0,1915	0,82	0,2939	1,14	0,3729
0,19	0,0753	0,51	0,1950	0,83	0,2967	1,15	0,3749
0,20	0,0793	0,52	0,1985	0,84	0,2995	1,16	0,3770
0,21	0,0832	0,53	0,2019	0,85	0,3023	1,17	0,3790
0,22	0,0871	0,54	0,2054	0,86	0,3051	1,18	0,3810
0,23	0,0910	0,55	0,2088	0,87	0,3078	1,19	0,3830
0,24	0,0948	0,56	0,2123	0,88	0,3106	1,20	0,3849
0,25	0,0987	0,57	0,2157	0,89	0,3133	1,21	0,3869
0,26	0,1026	0,58	0,2190	0,90	0,3159	1,22	0,3883
0,27	0,1064	0,59	0,2224	0,91	0,3186	1,23	0,3907
0,28	0,1103	0,60	0,2257	0,92	0,3212	1,24	0,3925
0,29	0,1141	0,61	0,2291	0,93	0,3238	1,25	0,3944
0,30	0,1179	0,62	0,2324	0,94	0,3264		
0,31	0,1217	0,63	0,2357	0,95	0,3289		

x	Φ (x)	x	Φ (x)	x	Φ (x)	x	Φ (x)
1,26	0,3962	1,59	0,4441	1,92	0,4726	2,50	0,4938
1,27	0,3980	1,60	0,4452	1,93	0,4732	2,52	0,4941
1,28	0,3997	1,61	0,4463	1,94	0,4738	2,54	0,4945
1,29	0,4015	1,62	0,4474	1,95	0,4744	2,56	0,4948
1,30	0,4032	1,63	0,4484	1,96	0,4750	2,58	0,4951
1,31	0,4049	1,64	0,4495	1,97	0,4756	2,60	0,4953
1,32	0,4066	1,65	0,4505	1,98	0,4761	2,62	0,4956
1,33	0,4082	1,66	0,4515	1,99	0,4767	2,64	0,4959
1,34	0,4099	1,67	0,4525	2,00	0,4772	2,66	0,4961
1,35	0,4115	1,68	0,4535	2,02	0,4783	2,68	0,4963
1,36	0,4131	1,69	0,4545	2,04	0,4793	2,70	0,4965
1,37	0,4147	1,70	0,4554	2,06	0,4803	2,72	0,4967
1,38	0,4162	1,71	0,4564	2,08	0,4812	2,74	0,4969
1,39	0,4177	1,72	0,4573	2,10	0,4821	2,76	0,4971
1,40	0,4192	1,73	0,4582	2,12	0,4830	2,78	0,4973
1,41	0,4207	1,74	0,4591	2,14	0,4838	2,80	0,4974
1,42	0,4222	1,75	0,4599	2,16	0,4846	2,82	0,4976
1,43	0,4236	1,76	0,4608	2,18	0,4854	2,84	0,4977
1,44	0,4251	1,77	0,4616	2,20	0,4861	2,86	0,4979
1,45	0,4265	1,78	0,4625	2,22	0,4868	2,88	0,4980
1,46	0,4279	1,79	0,4633	2,24	0,4875	2,90	0,4981
1,47	0,4292	1,80	0,4641	2,26	0,4881	2,92	0,4982
1,48	0,4306	1,81	0,4649	2,28	0,4887	2,94	0,4984
1,49	0,4319	1,82	0,4656	2,30	0,4893	2,96	0,4985
1,50	0,4332	1,83	0,4664	2,32	0,4898	2,98	0,4986
1,51	0,4345	1,84	0,4671	2,34	0,4904	3,00	0,49865
1,52	0,4357	1,85	0,4678	2,36	0,4909	3,20	0,49931
1,53	0,4370	1,86	0,4686	2,38	0,4913	3,40	0,49966
1,54	0,4382	1,87	0,4693	2,40	0,4918	3,60	0,499841
1,55	0,4394	1,88	0,4699	2,42	0,4922	3,80	0,499928
1,56	0,4406	1,89	0,4706	2,44	0,4927	4,00	0,499968
1,57	0,4418	1,90	0,4713	2,46	0,4931	4,50	0,499997
1,58	0,4429	1,91	0,4719	2,48	0,4934	5,00	0,499999