

Методическое пособие по теории вероятностей

Мусатов И.А.

Содержание

1	Основы комбинаторики	3
1.1	Принцип сложения и умножения	3
1.2	Базовые комбинаторные конфигурации	4
2	Случайные события и вероятность	9
2.1	Понятие случайного события	9
2.2	Классическое определение вероятности	9
2.3	Типовая задача о выборке	10
2.4	Геометрическое определение вероятности	11
2.5	Типовая задача о встрече	13
3	Условная вероятность и ее приложения	15
3.1	Операции над вероятностями	15
3.2	Типовая задача на надежность схем	16
3.3	Формула полной вероятности	18
3.4	Формула Байеса	19
4	Повторные независимые испытания	20
4.1	Формула Бернулли	20
4.2	Наивероятнейшее число появления события	20
4.3	Теоремы Муавра-Лапласа	21
4.4	Отклонение частоты от вероятности	23
4.5	Формула Пуассона	24
5	Случайные величины	26
5.1	Понятие случайной величины	26
5.2	Дискретные случайные величины	26
5.3	Числовые характеристики д.с.в	28
5.4	Непрерывные случайные величины	31
5.5	Числовые характеристики н.с.в	36
6	Распределения случайных величин	39
6.1	Производящие функции	39
6.2	Биномиальное распределение	39
6.3	Распределение Пуассона	40
6.4	Геометрическое распределение	42
6.5	Гипергеометрическое распределение	44
6.6	Равномерное распределение	45
6.7	Показательное распределение	47
6.8	Нормальное распределение	50

7	Двумерные случайные величины	50
7.1	Дискретные двумерные с.в.	50
7.2	Непрерывные двумерные с.в.	50
7.3	Корреляционная зависимость случайных величин	50
8	Основы математической статистики	50
8.1	Основные понятия статистики	50
A	Таблица значений функции Гаусса	51
B	Таблица значений функции Лапласа	52

1. Основы комбинаторики

Комбинаторика — основополагающий раздел математики, изучающий комбинации элементов некоторого конечного множества. Задачи, для решения которых используются комбинаторные методы, встречаются во многих разделах дискретной математики, в том числе и в дискретной вероятности.

В этой главе рассмотрим основные понятия комбинаторики.

1.1. Принцип сложения и умножения

Определение 1.1 (Правило суммы). Если некоторый объект класса A можно выбрать n способами, а другой объект класса B — m способами, то выбор объекта « A или B » можно осуществить $n + m$ способами.

Пример 1.1. В учебной группе состоят 12 девочек и 18 мальчиков. Требуется определить число возможных способов выбора старосты группы.

Решение. Каждая из девочек может стать старостой. Тогда из класса объектов A имеем $n = 12$ способов выбора. Аналогично и с мальчиками: имеем $m = 18$ способов. Тогда, согласно правилу суммы, общее число способов выбора старосты равно $n + m = 30$. \square

Замечание. При использовании правила суммы надо следить, чтобы способы выбора объектов A и B не пересекались. Иначе мы учтем какой-то выбор несколько раз. Поэтому в более общем случае формула выборки имеет вид: $n + m - k$, где k — число совпадающих способов выбора.

Определение 1.2 (Правило произведения). Если объект класса A можно выбрать n способами и если после каждого такого выбора объект класса B можно выбрать m способами, то выбор упорядоченной пары (A, B) можно осуществить $n \cdot m$ способами.

Для удобства восприятия данного правила можно воспользоваться мнемоникой «корытца». Пусть в первое «корытце» можно положить объект n способами, во второе m , в третье k и так далее. Тогда общее число различных способов разложения равно произведению количеств способов для каждого «корытца».



Рисунок 1. Мнемоника «корытца»

Пример 1.2. Определить количество трехзначных чисел, в записи которых не встречаются цифры «2» и «6».

Решение. Рассмотрим трехзначное число как совокупность трех «корытца» для расположения цифр. Тогда на позицию первой цифры имеем 7 вариантов (всего цифр

10, но «0» не подходит по определению трехзначного числа, а «2» и «6» не подходят по условию). На позицию второй и третьей цифр подходят по 8 вариантов. Тогда по правилу произведения имеем: $7 \cdot 8 \cdot 8 = 448$ возможных трехзначных чисел, удовлетворяющих условию. \square

В качестве более сложного примера рассмотрим задачу, где применяются правила как сложения, так и умножения.

Пример 1.3. Определить количество трехзначных чисел, в записи которых цифра «2» встречается ровно один раз.

Решение. Для начала из всех вариантов выделим взаимоисключающие случаи: когда цифра «2» располагается на первой позиции, когда располагается на второй и когда на третьей. Посчитаем для каждого случая число возможных способов в нем по правилу произведения. «2» на первой позиции: $1 \cdot 9 \cdot 9 = 81$ вариант, «2» на второй позиции: $8 \cdot 1 \cdot 9 = 72$ варианта, «2» на третьей позиции: $8 \cdot 9 \cdot 1 = 72$ варианта. Тогда, учитывая, что случаи взаимно исключают друг друга, общее число вариантов можем найти по правилу суммы: $81 + 72 + 72 = 225$ трехзначных чисел. \square

Частным случаем правила произведения является размещение с повторением.

Определение 1.3 (Размещения с повторениями). Числом размещений с повторениями из n по k (также известно как «выборка с возвращением») называется число:

$$\overline{A}_n^k = n^k$$

Обосновать данную формулу нетрудно. Полагая, что для каждого из k «корытец» у нас доступно n возможных предметов (причем предметы могут участвовать в размещении много раз), по правилу произведения действительно получаем $n \cdot n \cdot \dots \cdot n = n^k$

1.2. Базовые комбинаторные конфигурации

Рассмотрим центральные понятия комбинаторики и их различия.

Определение 1.4 (Перестановки). Числом перестановок из n называется число:

$$P_n = n!$$

Число перестановок — это количество всех различных способов переставить данные n объектов между собой. Понимать перестановки можно и через правило произведения. Пусть имеется конечное число объектов n , и требуется посчитать число их различных упорядочиваний друг за другом. Тогда на первую позицию можно поставить n объектов, на вторую $n-1$ (один объект уже поставлен), на третью $n-2$, ..., на n -ую 1 объект. Тогда общее число вариантов равно $n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 1 = n!$.

Графическая иллюстрация принципа перестановок представлена на рисунке 2.

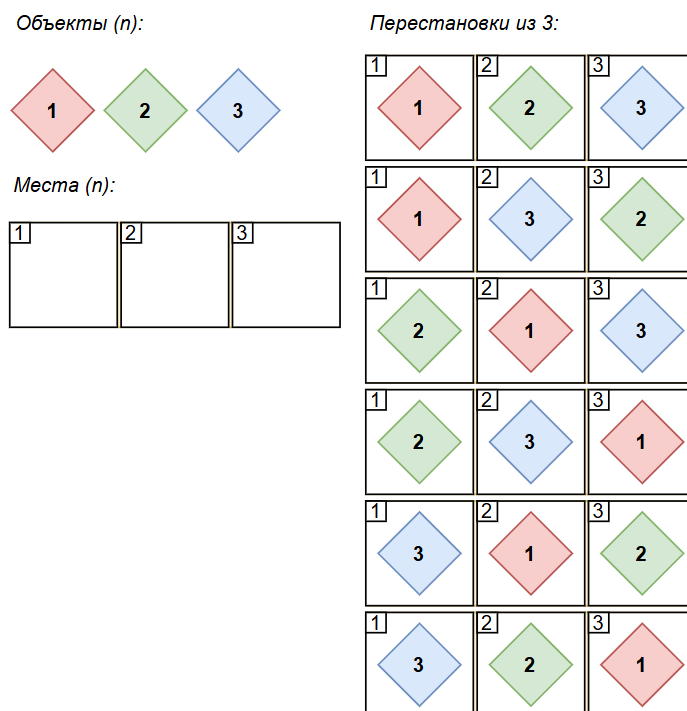


Рисунок 2. Иллюстрация перестановок

Пример 1.4. На полке стоят 10 книг. Найти число таких расстановок, при которых конкретные 4 книги располагаются рядом.

Решение. В данной задаче имеем ограничение на расстановку книг: конкретные 4 книги всегда должны стоять рядом. В таком случае временно «свяжем» эти 4 книги и будем рассматривать их как один объект. Тогда имеем $10 - 4 + 1 = 7$ объектов для свободной расстановки. Число их перестановок равно $7!$. Теперь вспомним, что внутри нашей «связки» книги также могут переставляться произвольно. Поэтому для каждого из $7!$ случаев можем рассматривать $4!$ перестановок внутри связки. Тогда по правилу произведения итоговое количество расстановок книг равно $P_7 \cdot P_4 = 7! \cdot 4! = 120960$. \square

Определение 1.5 (Размещения). Числом размещений из n по k называется число:

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

Число размещений — это количество всех различных способов расположить n элементов по k местам. При этом порядок расположения элементов **важен**. Т.е. расположения $(A B)$ и $(B A)$ в контексте размещений — это **разные** случаи.

Можно заметить, что при $k = n$ формула A_n^k обретает вид $A_n^n = n!$ и становится равной P_n . Действительно, если требуется разместить n элементов на n позиций, то мы имеем дело с их перестановками. Т.е. перестановки — это частный случай размещений.

Формулу размещений также несложно вывести через правило произведения. Сначала преобразуем ее в удобную форму: $\frac{n!}{(n-k)!} = n \cdot (n-1) \cdot \dots \cdot (n-k+1)$. Получили

произведение из k сомножителей, каждый из которых на единицу меньше предыдущего. Т.е. мы постепенно из конечного набора из n элементов k раз достаем по одному произвольному элементу.

Графическая иллюстрация принципа размещений представлена на рисунке 3.

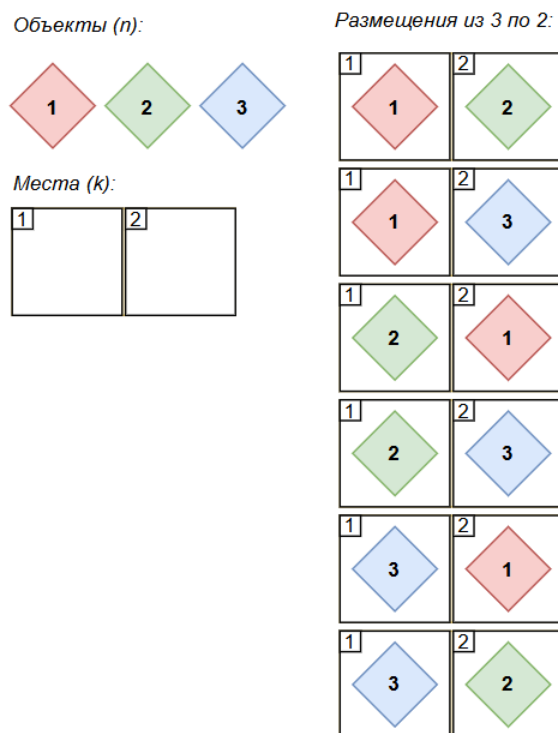


Рисунок 3. Иллюстрация размещений

Пример 1.5. Сколько существует чисел, в записи которых есть только нечетные цифры, причем никакая из цифр не встречается в числе дважды?

Решение. Выпишем все нечетные цифры: $\{1, 3, 5, 7, 9\}$. Их 5. Значит, числа, удовлетворяющие условию, не могут иметь длину больше пяти. Тогда рассмотрим все длины от 1 до 5. В каждом из таких случаев (длина числа равна i) количество возможных способов составления числа равно A_5^i . Действительно: имеется конечный набор цифр, которые нужно расставить на позиции. Т.к. разные длины у одного числа — это случаи несовместные, можем воспользоваться правилом суммы. Тогда итоговое количество чисел, удовлетворяющих условию, равно $A_5^1 + A_5^2 + \dots + A_5^5$ или же:

$$\sum_{i=1}^5 A_5^i = 325$$

□

Определение 1.6 (Сочетания). Числом сочетаний из n по k называется число:

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

Число сочетаний — это количество всех возможных способов выбора k элементов из n предложенных. При этом порядок расположения элементов в выборке **не важен**. Т.е. выборки $(A B)$ и $(B A)$ в контексте сочетаний — это **одинаковые** случаи.

Сочетания можно себе представлять как размещения с намеренным исключением повторений. Пусть имеем размещения A_n^k . Тогда для каждого уникального набора конкретных k объектов (скажем, набора $(A B C)$) в размещениях встречаются их все возможные перестановки $((A B C), (A C B), \dots, (C B A))$. Число таких перестановок нам известно. Это $P_k = k!$. Тогда, если мы хотим исключить такие повторения и для каждого уникального набора оставлять только одну «неупорядоченную» выборку, то достаточно общее число размещений поделить на число таких перестановок. В общем виде это выглядит как: $C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!}$.

Графическая иллюстрация принципа сочетаний представлена на рисунке 4.

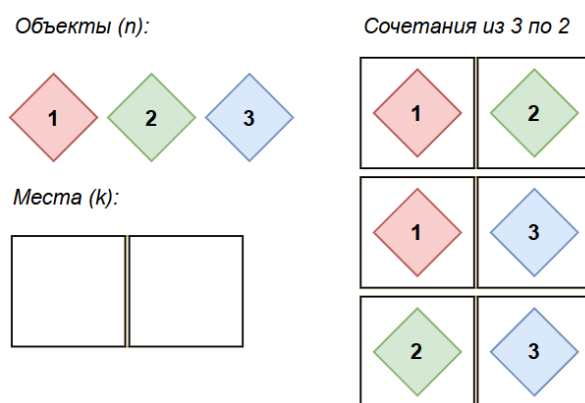


Рисунок 4. Иллюстрация сочетаний

Свойства сочетаний:

- 1°. $C_n^k = C_n^{n-k}$
- 2°. $C_n^0 = C_n^n = 1$
- 3°. $C_n^1 = n$
- 4°. $C_n^2 = \frac{n \cdot (n-1)}{2}$ — также известно как «число рукопожатий»
- 5°. $C_n^k + C_n^{k-1} = C_{n+1}^k$
- 6°. $\sum_{k=0}^n C_n^k = 2^n$

Пример 1.6. В корзине лежат 5 яблок и 3 груши. Найти число способов достать четыре фрукта так, чтобы среди них было не менее двух груш.

Решение. Т.к. в задаче присутствует неравенство («не менее двух груш»), сразу разделим ее на несколько несовместных ситуаций. Груш не менее двух, значит, больше либо равно две. Но всего их три, а значит, представляются две ситуации: 1) берем две груши и два яблока; 2) берем три груши и одно яблоко. В первом случае по правилу произведения считаем число способов выбора: $C_3^2 \cdot C_5^2 = 3 \cdot 10 = 30$. Во втором случае аналогично: $C_3^3 \cdot C_5^1 = 1 \cdot 5 = 5$. Тогда по правилу суммы для несовместных выборов имеем итоговое число способов:

$$C_3^2 \cdot C_5^2 + C_3^3 \cdot C_5^1 = 30 + 5 = 35$$

□

2. Случайные события и вероятность

2.1. Понятие случайного события

Определение 2.1. Случайным событием называется явление, которое либо произойдет, либо не произойдет при наступлении некоторых указанных условий.

«Наступление указанных условий» обычно кем-то производится намеренно. Поэтому будем называть «наступление» *экспериментом* или *испытанием*.

Замечание. Случайные события обозначаются заглавными латинскими буквами

Определение 2.2. Событие называется **достоверным**, если оно обязательно произойдет в данном испытании.

Определение 2.3. Событие называется **невозможным**, если оно никогда не произойдет в данном испытании.

Определение 2.4. Два события A и B называются **несовместными**, если они не могут одновременно появиться в одном испытании. Т.е. появление одного из них исключает появление другого.

Определение 2.5. Противоположным к событию A называется событие \bar{A} , которое заключается в отрицании появления A . Отличие от понятия «несовместности» в том, что в любом испытании всегда произойдет ровно одно из противоположных событий. В то время как несовместные события могут не возникнуть вовсе.

Определение 2.6. Суммой двух событий A и B называется событие $A+B$, которое возникает в случае наступления хотя бы одного из этих событий.

Определение 2.7. Произведением двух событий A и B называется событие $A \cdot B$, которое возникает только в случае, когда оба события наступили.

Нетрудно заметить, что случайные события имеют теоретико-множественную природу. Т.е. операции над случайными событиями схожи с операциями над множествами. Соответственно, к арифметике случайных событий можно применять привычные законы коммутативности, ассоциативности, дистрибутивности, исключенного третьего, идемпотентности, де Моргана.

Определение 2.8. События A_1, A_2, \dots, A_n образуют **полную группу**, если в результате испытания обязательно появится одно из них. Т.е. их сумма является достоверным событием.

Например, события A и \bar{A} противоположны, следовательно, несовместны и образуют полную группу.

2.2. Классическое определение вероятности

Определение 2.9. Элементарные исходы (результаты испытания), в которых интересующее нас случайное событие наступает, называются **благоприятствующими** данному событию.

Определение 2.10 (Классическое определение вероятности). Вероятность события A равна отношению числа благоприятствующих этому событию исходов к общему числу всех исходов данного испытания. Обозначается как:

$$P(A) = \frac{M}{N}$$

Из определения следует, что значение $P(A)$ лежит в промежутке $[0; 1]$. Где значение 0 свидетельствует о невозможности события A , а значение 1 — о достоверности.

Теорема 2.1. Вероятность суммы двух несовместных событий равна сумме их вероятностей:

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

Теорема 2.2. Сумма противоположных событий равна 1. Т.е. вероятность противоположного к A события равна:

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

2.3. Типовая задача о выборке

Нередко в статистике и математическом моделировании встречается задача, которую в общем виде можно описать следующим образом:

Определение 2.11 (Задача о выборке). Имеется N однородных предметов. K из них «особые». Произвольным образом из N выбираются n предметов. Определить вероятность того, что среди выбранных n предметов ровно k будут «особыми».

Решение. Решим задачу в общем виде. Для начала определим общее число возможных исходов. Имеется N элементов, из них выбирается n , причем порядок не имеет значения. Значит, общее число исходов — это число сочетаний C_N^n .

Теперь найдем число благоприятствующих исходов. Т.е. таких случаев, когда среди выбранных n предметов ровно k «особых». Раз число фиксировано, можно разбить задачу на две подзадачи: выбрать k «особых» предмета и $n - k$ не «особых». Первая подзадача решается формулой выбора: C_K^k , вторая — формулой выбора из оставшихся предметов: C_{N-K}^{n-k} . Понятно, что для каждого выбора в первой подзадаче может выпасть любой выбор из второй подзадачи. Значит, для вычисления общего числа благоприятствующих исходов можем воспользоваться правилом произведения, получив при этом: $C_K^k \cdot C_{N-K}^{n-k}$

Тогда итоговая обобщенная формула для решения задачи о выборке имеет вид:

$$P = \frac{C_K^k \cdot C_{N-K}^{n-k}}{C_N^n}$$

□

Пример 2.1. В партии из 10 деталей 5 имеют отличное качество, 3 — хорошее, а оставшиеся 2 — бракованные. Случайным образом из партии выбираются 4 детали. Найти вероятность того, что среди выбранных деталей есть 2 отличного качества и одна бракованная.

Решение. Общее число способов выбора 4 деталей из 10 равно C_{10}^4 .

Число благоприятных ситуаций — это выборка 2х отличных деталей из 5 возможных, 1ой бракованной из 2 возможных, ну и оставшейся 1ой детали из 3 хороших. Тогда по правилу произведения можем представить это число как: $C_5^2 \cdot C_2^1 \cdot C_3^1$.

Следуя классическому определению вероятности, итоговая формула будет иметь вид:

$$P = \frac{C_5^2 \cdot C_2^1 \cdot C_3^1}{C_{10}^4} = \frac{2}{7}$$

□

2.4. Геометрическое определение вероятности

Геометрические методы в математике всегда привлекают своей наглядностью. В теории вероятностей некоторые задачи также могут быть представлены в геометрической форме. Для этого достаточно ввести понятие *меры*. Учитывая, что количество точек на любом отрезке и на любой площади бесконечно, мы прибегаем к таким конечным понятиям, как длина, площадь и объем. Для их обобщения и вводится понятие *меры* как пространственного измерения отрезка, фигуры и тела в пространствах соответствующей размерности.

Определение 2.12 (Геометрическое определение вероятности). Пусть U — область, соответствующая всем возможным элементарным исходам некоторого испытания. A — подобласть (не обязательно непрерывная) области U , соответствующая всем благоприятствующим исходам. Тогда вероятность наступления случайного события соответствует вероятности попадания точки в область A , т.е. отношению мер областей:

$$P(A) = \frac{\mu(A)}{\mu(U)}$$

Пример 2.2. Дан отрезок AB длины L . На нем произвольным образом отмечаются две точки: M и N . Найти вероятность того, что длина получившегося отрезка MN будет меньше, чем четверть длины исходного отрезка.

Решение. Имеем две точки, каждая из которых отмечается на отрезке независимо от другой. Тогда можем перейти к описанию каждой точки через определение переменных. Пусть x отвечает за положение точки M , а y — за N . Тогда справедливо, что $0 \leq x, y \leq L$.

Сформулируем теперь вопрос задачи, используя введенные переменные. Длина отрезка MN через переменные определяется как $|y - x|$. Т.е. требуется найти $P\{|y - x| < L/4\}$

Имея две переменные величины, можем перейти к геометрии на прямоугольной декартовой системе координат. Тогда пространством всех исходов станет квадрат с $0 \leq x, y \leq L$, а пространством благоприятствующих исходов — $|y - x| < L/4 \iff y < x + L/4 \wedge y > x - L/4$. Изобразим это графически (см. рис. 5).

Нетрудно посчитать площади указанных областей. Тогда итоговая формула вероятности имеет вид:

$$P\{[MN] < L/4\} = \frac{\mu(|y - x| < L/4)}{\mu(0 \leq x, y \leq L)} = \frac{7}{16}$$

□

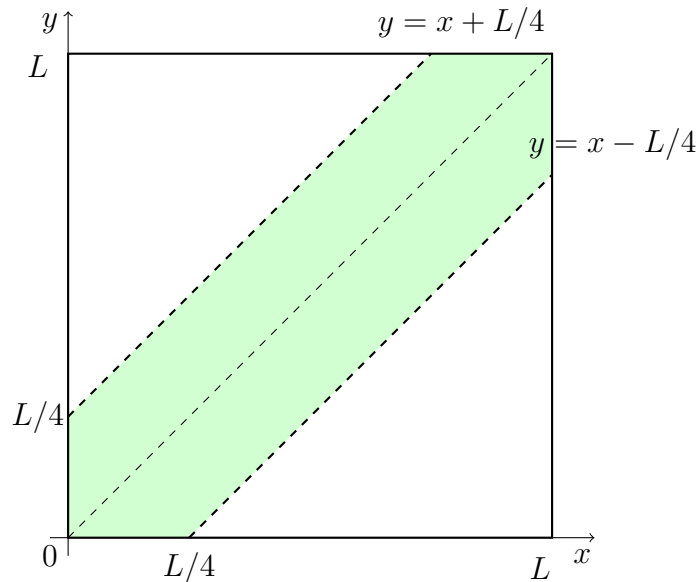


Рисунок 5. Область $|y - x| < \frac{L}{4}$

Пример 2.3. Даны два случайных действительных числа из отрезка $[0; 2]$. Найти вероятность того, что произведение этих чисел будет больше 1.

Решение. Обозначим числа за x и y . Тогда условие задачи можно переписать в виде: $x, y \in \mathbb{R}; 0 \leq x, y \leq 2; P\{x \cdot y > 1\} - ?$.

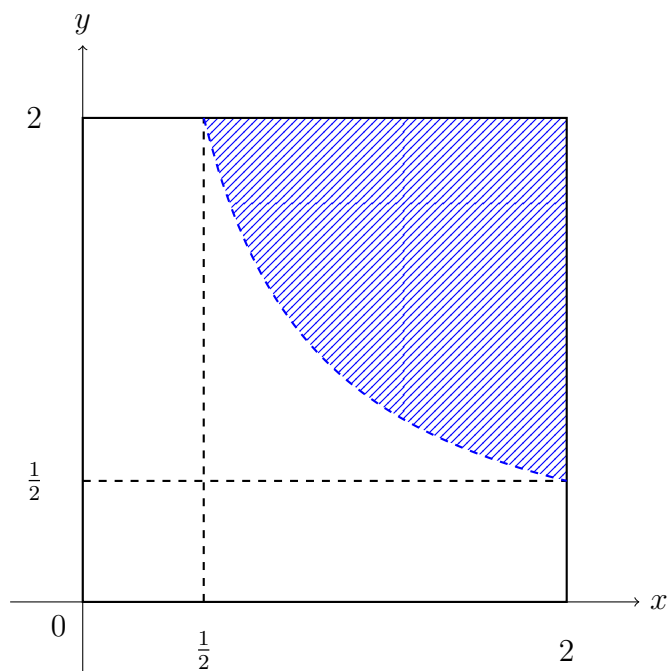
Перейдем к декартовой системе координат и изобразим области всех исходов и благоприятствующих исходов (см. рис. 6). Тогда вероятность может быть найдена как отношение площадей:

$$\mu(x \cdot y > 1) = \frac{3}{2} \cdot 2 - \int_{0.5}^2 \frac{1}{x} dx = 3 - \ln x \Big|_{0.5}^2 = 3 - \ln 4$$

$$\mu(0 \leq x, y \leq 2) = 4$$

$$P\{x \cdot y > 1\} = \frac{\mu(x \cdot y > 1)}{\mu(0 \leq x, y \leq 2)} = \frac{3 - \ln 4}{4}$$

□

Рисунок 6. Область $x \cdot y > 1$

2.5. Типовая задача о встрече

Говоря о геометрическом определении вероятности, стоит рассмотреть отдельный класс задач, называемых задачами о встрече. Рассмотрим пример такой задачи.

Пример 2.4 (Задача о встрече). Два друга договорились о встрече в промежуток времени с 18:00 до 18:50. Причем условились, что каждый, придя на место встречи, будет ожидать другого не более, чем 20 минут. Найти вероятность того, что друзья **не** смогут встретиться.

Решение. Анализируя условие, заметим, что нам не важна привязка начала промежутка времени к конкретному времени дня. Т.е. задача была эквивалентной, если бы они договорились встретиться в промежуток, скажем, с 12:40 до 13:30. Поэтому в задаче важен именно сам *промежуток* времени, т.е. его продолжительность. В данном случае это 50 минут.

На самом деле, подобную задачу мы уже рассматривали в примере 2.2. По сути, у нас есть отрезок (в данном случае времени), и две случайные точки (времена прибытия друзей). Можем так же ввести две переменные $0 \leq x, y \leq 50$ — случайное время прибытия от начала промежутка времени.

Условие встречи друзей можно переформулировать как $|y - x| \leq 20$. Т.е. разница между временами прихода друзей не превышает 20 минут. Тогда условие того, что они не встретятся — $|y - x| > 20$.

Получили условие, переведенное в удобную для построения на плоскости форму. Графическое решение представлено на рис. 7.

А итоговая формула имеет вид:

$$P\{|y - x| > 20\} = \frac{\mu(|y - x| > 20)}{\mu(0 \leq x, y \leq 50)} = \frac{9}{25}$$

□

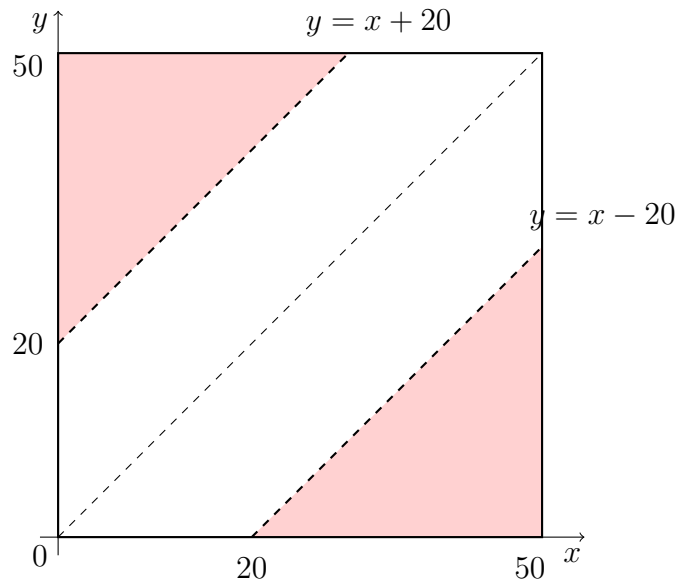


Рисунок 7. Область $|y - x| > 20$

3. Условная вероятность и ее приложения

3.1. Операции над вероятностями

Теорема 3.1 (Теорема о сложении вероятностей). Вероятность суммы двух случайных событий равна сумме их вероятностей за вычетом вероятности их произведения:

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

Замечание. В частности, если A и B — события несовместные, имеем: $P(A \cdot B) = 0$. В таком случае вероятность их суммы: $P(A + B) = P(A) + P(B)$.

Определение 3.1 (Условная вероятность). Условной вероятностью $P(A|B)$ называют вероятность возникновения события A при условии того, что событие B уже произошло.

Условная вероятность события A также обозначается как: $P_B(A)$

Теорема 3.2 (Теорема о произведении вероятностей). Вероятность произведения двух случайных событий равна произведению вероятности первого события на условную вероятность второго:

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

Определение 3.2. Случайное событие B называют **независимым** от события A , если факт возникновения события A не изменяет вероятности возникновения B :

$$P(B|A) = P(B)$$

Следствие. Вероятность произведения независимых событий равна произведению их вероятностей:

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

Определение 3.3. Вероятность появления хотя бы одного из независимых событий A_1, A_2, \dots, A_n противоположна произведению вероятностей не появления каждого из событий:

$$P(A_1 \vee A_2 \vee \dots \vee A_n) = 1 - P(\bar{A}_1) \cdot P(\bar{A}_2) \cdot \dots \cdot P(\bar{A}_n)$$

Пример 3.1. В первом мешке 5 белых и 9 черных шаров, а во втором — 7 белых и 6 черных. Из каждого мешка вынули по одному шару. Найти вероятность того, что вынутые шары будут разного цвета.

Решение. Событие A = «Вынутые шары разного цвета» есть сумма двух несовместных событий: A_1 = «Первый вынутый белый, второй — черный» и A_2 = «Первый вынутый черный, второй — белый».

Событие A_1 — это произведение двух независимых событий: A_{11} = «Вынут белый из первого мешка» A_{12} = «Вынут черный из второго мешка». $P(A_{11}) = 5/14$, $P(A_{12}) = 6/13$. Т.о. $P(A_1) = P(A_{11}) \cdot P(A_{12}) = \frac{5}{14} \cdot \frac{6}{13} = \frac{30}{182}$.

Аналогично событие A_2 — это произведение двух независимых событий: A_{21} = «Вынут черный из первого мешка» A_{22} = «Вынут белый из второго мешка». $P(A_{21}) = 9/14$, $P(A_{22}) = 7/13$. Т.о. $P(A_2) = P(A_{21}) \cdot P(A_{22}) = \frac{9}{14} \cdot \frac{7}{13} = \frac{63}{182}$.

Событие A — сумма несовместных событий, значит,

$$P(A) = P(A_1) + P(A_2) = P(A_{11}) \cdot P(A_{12}) + P(A_{21}) \cdot P(A_{22}) = \frac{30}{182} + \frac{63}{182} = \frac{93}{182}$$

□

3.2. Типовая задача на надежность схем

При создании электротехнических приборов в таких критических областях, как медицина, авионика, энергетика, жизненно важно учитывать все риски, которые могут нести в себе устройства. В качестве числовых характеристик для оценки надежности можно применять аппарат теории вероятностей. Ставится задача по вероятности надежности отдельных элементов вычислить надежность схемы в целом.

Рассмотрим основные соединения элементов в цепи:

Последовательное соединение работает исправно только если все элементы, в нем состоящие, работают исправно. На языке вероятности можем выразить это как:

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

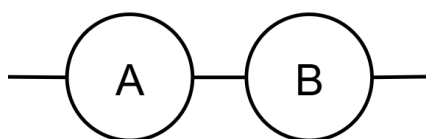


Рисунок 8. Последовательное соединение

Параллельное соединение работает исправно, если хотя бы один элемент такого соединения работает исправно. Иными словами, параллельное соединение не работает только в случае, когда не работают все элементы в нем. В таком случае на языке вероятности удобнее рассматривать надежность параллельного соединения от противоположного:

$$P(A \vee B) = 1 - P(\bar{A}) \cdot P(\bar{B})$$

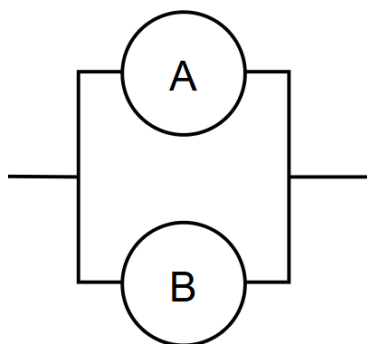


Рисунок 9. Параллельное соединение

Замечание. Определения надежности последовательного и параллельного соединений можно распространить и на произвольное число элементов.

Пример 3.2. Определить надежность схемы, если известны вероятности безотказной работы ее элементов: $P(A) = 0.6$, $P(B) = 0.8$.

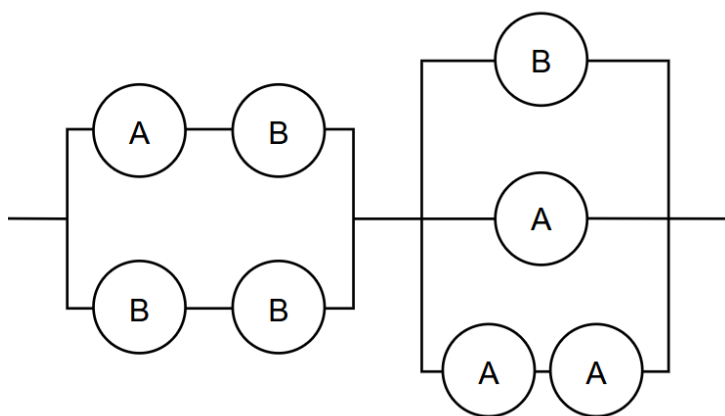


Рисунок 10. Схема к примеру 3.2

Решение. Для начала вычислим надежности всех последовательных соединений на схеме. В нашем случае это: AB , BB , AA .

$$P(AA) = P(A) \cdot P(A) = 0.6 \cdot 0.6 = 0.36$$

$$P(AB) = P(A) \cdot P(B) = 0.6 \cdot 0.8 = 0.48$$

$$P(BB) = P(B) \cdot P(B) = 0.8 \cdot 0.8 = 0.64$$

Затем вычислим надежности параллельных соединений. В нашем случае имеется два параллельных соединения: $AB \vee BB$ и $B \vee A \vee AA$.

$$P(AB \vee BB) = 1 - P(\overline{AB}) \cdot P(\overline{BB}) = 1 - (1 - P(AB)) \cdot (1 - P(BB)) = 0.813$$

$$P(B \vee A \vee AA) = 1 - P(\overline{B}) \cdot P(\overline{A}) \cdot P(\overline{AA}) = 1 - (1 - P(B))(1 - P(A))(1 - P(AA)) = 0.949$$

В арифметике событий всю схему можно представить как:

$$S = (AB \vee BB) \cdot (B \vee A \vee AA)$$

Тогда вероятность ее надежности:

$$P(S) = P((AB \vee BB) \cdot (B \vee A \vee AA)) = P(AB \vee BB) \cdot P(B \vee A \vee AA) = 0.813 \cdot 0.949 = 0.772$$

□

Замечание. Зная надежность схемы, можем определить и вероятность ее отказа. Эти события противоположны, а значит, $P(\bar{S}) = 1 - P(S)$.

3.3. Формула полной вероятности

Теорема 3.3 (Формула полной вероятности). Вероятность события A , которое может наступить только при условии появления одного из несовместных событий H_1, H_2, \dots, H_n , образующих *полную группу*, равна сумме произведений вероятности каждого из событий на условную вероятность события A :

$$P(A) = P(A|H_1) \cdot P(H_1) + P(A|H_2) \cdot P(H_2) + \dots + P(A|H_n) \cdot P(H_n)$$

Или:

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

Замечание. События H_1, H_2, \dots, H_n принято также называть **гипотезами**.

Полной такая вероятность называется потому, что рассматривает все возможные основания появления события A . Т.е. предполагает по отдельности каждую из гипотез, которые в совокупности составляют *полную группу* (см. опр. 2.8).

Пример 3.3. В соревнованиях участвовали 10 лыжников, 6 велосипедистов и 4 бегуна. Вероятность того, что лыжник выполнит свою норму, равняется 0.9, велосипедист — 0.8, бегун — 0.75. Определить вероятность события A = «Случайно выбранный спортсмен выполнит свою норму».

Решение. Составим небольшую табличку по данному условию:

	Лыж	Вел	Бег
Кол-во	10	6	4
$P(A H_i)$	0.9	0.8	0.75

Определим гипотезы, т.е. все возможные несовместные предпосылки события A :

H_1 — «Выбранный спортсмен - лыжник»

H_2 — «Выбранный спортсмен - велосипедист»

H_3 — «Выбранный спортсмен - бегун»

$$P(H_1) = 10/20 = 0.5$$

$$P(H_2) = 6/20 = 0.3$$

$$P(H_3) = 4/20 = 0.2$$

Тогда по формуле полной вероятности:

$$\begin{aligned} P(A) &= P(A|H_1) \cdot P(H_1) + P(A|H_2) \cdot P(H_2) + P(A|H_3) \cdot P(H_3) = \\ &= 0.9 \cdot 0.5 + 0.8 \cdot 0.3 + 0.75 \cdot 0.2 = 0.84 \end{aligned}$$

□

3.4. Формула Байеса

Теорема 3.4 (Формула Байеса). Пусть A — событие, которое может наступить только при осуществлении одной из гипотез полной группы (H_1, \dots, H_n) . Тогда вероятность того, что событие A осуществилось благодаря условию H_i , равна:

$$P(H_i|A) = \frac{P(A|H_i) \cdot P(H_i)}{P(A)}$$

Замечание. Формулу Байеса также называют **формулой переоценки гипотез**.

Грубо говоря, формулу Байеса можно интерпретировать как взвешенную долю, которую вложила данная гипотеза в осуществление события A .

Пример 3.4. Для примера 3.3 найти вероятность того, что «Спортсмен, выполнивший норму, является бегуном»

Решение. Значения $P(A)$, $P(H_3)$, $P(A|H_3)$ уже известны. Требуется только по формуле определить $P(H_3|A)$:

$$P(H_3|A) = \frac{P(A|H_3) \cdot P(H_3)}{P(A)} = \frac{0.75 \cdot 0.2}{0.84} = 0.1786$$

□

4. Повторные независимые испытания

Общую постановку задач на повторные независимые испытания можно сформулировать следующим образом. Проводится n однородных независимых испытаний, в результате каждого из которых событие A может наступить или не наступить. Требуется определить вероятность того, что в n испытаниях событие A возникнет ровно m раз: $P_n(m) = ?$.

4.1. Формула Бернулли

Теорема 4.1 (Формула Бернулли). Пусть проводится конечное число n испытаний, в каждом из которых событие A может возникнуть с вероятностью p и не возникнуть с вероятностью $q = 1 - p$. Тогда вероятность возникновения события A ровно m раз равна:

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

Формулу Бернулли нетрудно обосновать. Общая вероятность независимых событий равна произведению их вероятностей. Тогда раз мы имеем ровно m событий с вероятностью p и $n - m$ событий с вероятностью $q = 1 - p$, достаточно перемножить их соответствующее число раз. Но при этом не стоит также забывать, что мы не учли различные порядки следования успешности испытаний. Поэтому общую вероятность нужно домножить на число возможных конфигураций успешных событий. Оно равно числу выбора m испытаний из n , т.е. числу сочетаний.

Пример 4.1. Вероятность попадания в мишень при одном выстреле равна 0.6. Было произведено 10 выстрелов.

- а) Найти вероятность того, что успешных выстрелов было ровно три;
- б) Найти вероятность того, что успешных выстрелов было не больше восьми;

Решение.

$$\text{а) } P_{10}(3) = C_{10}^3 \cdot 0.6^3 \cdot 0.4^7 = 0.042$$

б) Можем пойти от вероятности противоположного события. «Успешных выстрелов было больше 8». Тогда искомая вероятность: $P_{10}(\leq 8) = 1 - P_{10}(9) - P_{10}(10) = 1 - (C_{10}^9 \cdot 0.6^9 \cdot 0.4 + C_{10}^{10} \cdot 0.6^{10}) = 1 - 0.046 = 0.954$ \square

4.2. Наивероятнейшее число появления события

Определение 4.1 (Наивероятнейшее число появления события). Наивероятнейшим числом появления события A в n испытаниях называется такое число m^* , вероятность $P_n(m^*)$ которого максимальна. Определяется через неравенство:

$$(n + 1) \cdot p - 1 \leq m^* \leq (n + 1) \cdot p$$

То же, что и:

$$np - q \leq m^* \leq np + p$$

Замечание. Если обе границы неравенства являются целыми числами, то существует два наивероятнейших числа появления данного события.

Пример 4.2. Имеется 8 станков. Вероятность исправной работы каждого из них равна 0.6. Определить, какое количество станков наиболее вероятно окажется исправным.

Решение. Имеем: $n = 8$, $p = 0.6$. Тогда:

$$9 \cdot 0.6 - 1 \leq m^* \leq 9 \cdot 0.6$$

$$4.4 \leq m^* \leq 5.4$$

$$m^* = 5$$

□

Пример 4.3. Вероятность выпуска прибора высшего качества составляет 0.75. Было выпущено 103 прибора. Определить наиболее вероятное число приборов высшего качества среди выпущенных.

Решение. Имеем: $n = 103$, $p = 0.75$. Тогда:

$$104 \cdot 0.75 - 1 \leq m^* \leq 104 \cdot 0.75$$

$$77 \leq m^* \leq 78$$

$$m^* = 77, m^* = 78$$

□

4.3. Теоремы Муавра-Лапласа

При большом числе повторных независимых испытаний формула Бернулли вынуждает оперировать с числами значительно разных порядков (Так, скажем, уже при $n = 10$, $m = 7$: $10! = 3\,628\,800$, а $0.2^7 = 0.0000128$. Что может стать причиной неустранимой погрешности вычислений. В целях минимизации данной погрешности используются локальная теорема Муавра-Лапласа и интегральная теорема Лапласа, дающие асимптотическую формулу для нахождения приближенных значений вероятности.

Замечание. Условие хорошего приближения данных теорем — достаточно большое n . Чтобы дать этому численную характеристику, зададим условие: $n \cdot p \cdot q \geq 10$.

Теорема 4.2 (Локальная теорема Муавра-Лапласа). Если в каждом из n независимых испытаний вероятность p появления события A постоянна, не равна 0 и 1, то вероятность возникновения события A ровно m раз приближенно равна:

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

где $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2}$ — функция Гаусса. $\varphi(-x) = \varphi(x)$; $\varphi(\infty) = 0$.

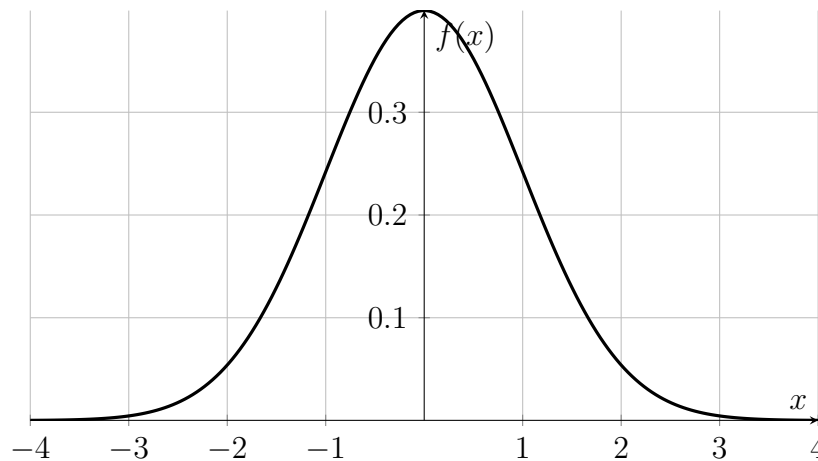


Рисунок 11. График функции Гаусса

Замечание. Значения функции Гаусса приведены в конце данного пособия в виде таблицы.

Пример 4.4. Монетка подбрасывается 324 раза. Определить вероятность того, что «Орел» выпадет ровно 172 раза.

Решение. Вероятность выпадения «Орла» $p = 0.5$. Проверим применимость локальной теоремы Муавра-Лапласа к данной задаче: $npq = 324 \cdot 0.5 \cdot 0.5 = 81 \geq 10$. Тогда:

$$P_{324}(172) = \frac{1}{\sqrt{npq}} \cdot \varphi\left(\frac{m - np}{\sqrt{npq}}\right) = \frac{1}{9} \cdot \varphi\left(\frac{172 - 162}{9}\right) = \frac{1}{9} \cdot \varphi(1.11) = 0.024$$

□

Теорема 4.3 (Интегральная теорема Лапласа). Если в каждом из n независимых испытаний вероятность p появления события A постоянна, не равна 0 и 1, то вероятность возникновения события A не менее m_1 , но не более m_2 раз приблизительно равна:

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

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^x e^{-t^2/2} dt$ — функция Лапласа. $\Phi(-x) = -\Phi(x)$; $\Phi(\infty) = 0.5$; $\Phi(0) = 0$.

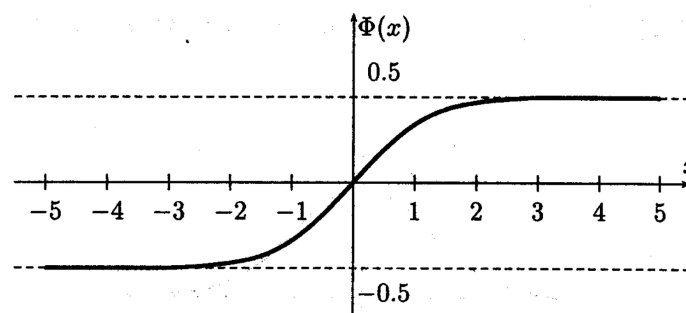


Рисунок 12. График функции Лапласа

Замечание. Значения функции Лапласа приведены в конце данного пособия в виде таблицы.

Пример 4.5. На потоке учатся 150 студентов. Вероятность того, что случайно взятый студент придет на лекцию по теории вероятностей — 0.2. Найти вероятность того, что на лекции по теории вероятностей будет от 25 до 35 студентов.

Решение. Проверим применимость интегральной теоремы Лапласа: $npq = 150 \cdot 0.2 \cdot 0.8 = 24 \geq 10$. Тогда:

$$\begin{aligned} P_{150}(25 \leq m \leq 35) &= \Phi\left(\frac{35 - 150 \cdot 0.2}{\sqrt{24}}\right) - \Phi\left(\frac{25 - 30}{\sqrt{24}}\right) = \Phi(1.02) - \Phi(-1.02) = \\ &= \Phi(1.02) + \Phi(1.02) = 2 \cdot \Phi(1.02) = 0.69 \end{aligned}$$

□

Пример 4.6. Вероятность того, что деталь из новой партии соответствует ГОСТ, составляет 0.8. Найти количество деталей в этой партии, если известно, что больше 78 деталей соответствуют ГОСТ с вероятностью 0.6915.

Решение. По условию имеем: $p = 0.8$, $q = 0.2$, $P(78 \leq m) = 0.6915$.

$$P(78 \leq m) = \Phi(\infty) - \Phi\left(\frac{78 - 0.8 \cdot n}{\sqrt{0.8 \cdot 0.2n}}\right) = 0.6915$$

Знаем, что $\Phi(\infty) = 0.5$. Тогда:

$$\Phi\left(\frac{78 - 0.8 \cdot n}{\sqrt{0.8 \cdot 0.2n}}\right) = -0.1915 \iff \Phi\left(\frac{0.8 \cdot n - 78}{\sqrt{0.8 \cdot 0.2n}}\right) = 0.1915$$

Смотрим значение аргумента по таблице:

$$\frac{0.8 \cdot n - 78}{\sqrt{0.8 \cdot 0.2n}} = 0.5$$

$$0.8n - 78 = 0.2\sqrt{n}$$

$$4n - \sqrt{n} - 390 = 0$$

$$n = 100$$

□

4.4. Отклонение частоты от вероятности

Определение 4.2. Проводятся независимые испытания с постоянной вероятностью p возникновения события A в каждом из них. При этом в рамках данного эксперимента событие A появилось m раз в n испытаниях. Тогда вероятность того, что отклонение относительной частоты m/n от вероятности p по модулю не превзойдет значения ε , приблизительно составляет:

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

Пример 4.7. Проводятся измерения механического станка, изготавливающего качественные детали с вероятностью $p = 0.8$. Вероятность того, что отклонение частоты от вероятности не превзойдет $\varepsilon = 0.05$ составляет 0.95. Найти число проведенных испытаний n .

Решение. Имеем: $p = 0.8$, $P\left\{\left|\frac{m}{n} - 0.8\right| \leq 0.05\right\} = 0.95$

Определим n исходя из формулы:

$$P\{\dots\} = 0.95 \approx 2 \cdot \Phi\left(0.05 \cdot \sqrt{\frac{n}{0.8 \cdot 0.2}}\right)$$

$$\Phi\left(0.05 \cdot \sqrt{\frac{n}{0.8 \cdot 0.2}}\right) = 0.475$$

Значение аргумента из таблицы функции Лапласа:

$$0.05 \cdot \sqrt{\frac{n}{0.8 \cdot 0.2}} = 1.96$$

$$n = 245.86 \approx 246$$

□

4.5. Формула Пуассона

Хоть локальная и интегральная формулы Муавра-Лапласа и хорошо приближают значение вероятности при большом числе испытаний, они, все же, дают большие погрешности при значениях вероятности, стремящихся к 1 или 0. В подобных ситуациях применяют асимптотическую формулу Пуассона.

Теорема 4.4 (Формула Пуассона). Если число n испытаний достаточно велико, а вероятность p появления события A в каждом испытании постоянна и близка к 0, то вероятность появления события A ровно m раз приближенно равна:

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

где $\lambda = n \cdot p$.

Замечание. Формула Пуассона дает хорошее приближение при небольших λ . Приблизительно $2 \leq \lambda \leq 9$.

Пример 4.8. Высокоточный фрезерный ЧПУ-станок выпускает 1000 шестерней за смену. Вероятность того, что шестерня окажется бракованной, составляет 0.008. Определить вероятность того, что за смену ровно 5 деталей, выпущенных станком, окажутся бракованными.

Решение. По условию: $n = 1000$, $p = 0.008$. $\lambda = np = 8$.

$$P_{1000}(5) = \frac{\lambda^m}{m!} \cdot e^{-\lambda} = \frac{8^5}{5!} \cdot e^{-8} = 0.092$$

□

Пример 4.9. В городе есть 200 книжных магазинов. Вероятность того, что студент в случайно выбранном магазине найдет книгу по функциональному анализу, составляет 0.01. Определить вероятность того, что студент сможет отыскать книгу хотя бы в двух магазинах.

Решение. Имеем: $n = 200$, $p = 0.01$. $\lambda = np = 2$. Требуется найти $P_n(m \geq 2)$. Рассмотрим противоположное событие: студент сможет найти книгу меньше, чем в двух магазинах: $P_n(m < 2) = P_{200}(0) + P_{200}(1)$.

$$P_{200}(0) = \frac{\lambda^m}{m!} \cdot e^{-\lambda} = \frac{2^0}{0!} \cdot e^{-2} = 0.135$$

$$P_{200}(1) = \frac{\lambda^m}{m!} \cdot e^{-\lambda} = \frac{2^1}{1!} \cdot e^{-2} = 0.27$$

Тогда вероятность исходного события:

$$P_{200}(m \geq 2) = 1 - P_{200}(m < 2) = 1 - (P_{200}(0) + P_{200}(1)) = 1 - 0.405 = 0.595$$

□

5. Случайные величины

5.1. Понятие случайной величины

Для придания формальной характеристики результату испытания вводится такое понятие, как случайная величина. В общем виде это величина, которая в ходе испытания может принять одно значение из известного множества вероятных исходов.

Замечание. Случайные величины традиционно обозначаются греческими буквами $(\xi, \eta, \zeta, \dots)$, а их значения — латинскими буквами (x_i, y_i, z_i, \dots) .

По характеру множества допустимых значений случайные величины делят на дискретные и непрерывные.

Определение 5.1. **Дискретной** называют случайную величину, которая принимает отдельные значения из конечного или бесконечного, но счетного множества.

Определение 5.2. **Непрерывной** называют случайную величину, которая принимает значения из числового промежутка: интервала, полуинтервала или отрезка (т.е. значения из несчетного множества).

Замечание. В дальнейшем повествовании именование «дискретные случайные величины» будем сокращать до «д.с.в.», а «непрерывные случайные величины» — до «н.с.в.».

5.2. Дискретные случайные величины

Определение 5.3. **Законом распределения** дискретной случайной величины ξ называется соотношение между возможными значениями x_i этой величины и соответствующими им вероятностями $p_i = P\{\xi = x_i\}$. Такое соотношение обычно выражается в виде таблицы:

ξ	x_1	x_2	\dots	x_n
p	p_1	p_2	\dots	p_n

Замечание. Закон распределения также называют рядом распределения или же просто распределением случайной величины.

Замечание. Таблица является законом распределения случайной величины только если выполняется условие «нормировки»: сумма вероятностей должна быть равна 1.

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

Определение 5.4. **Многоугольником распределения** называется график закона распределения. Т.е. откладывание значений случайной величины по оси Ох, а вероятностей — по оси Оу.

Пример 5.1. Имеется закон распределения д.с.в. ξ . Требуется построить многоугольник ее распределения.

ξ	-1	0	2	5
p	0.2	0.5	0.1	0.2

Решение. Перенесем данные из таблицы на график.

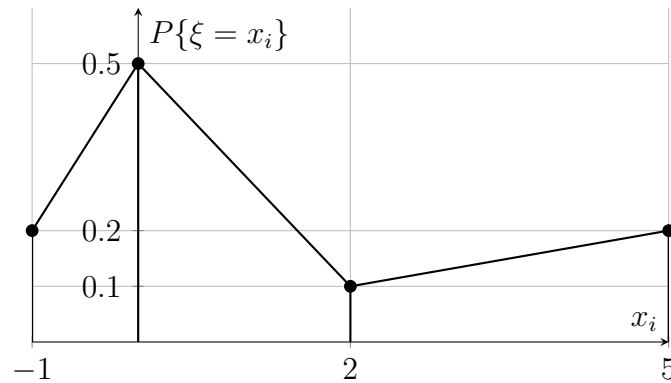


Рисунок 13. Многоугольник распределения д.с.в. ξ

□

Определение 5.5. **Функцией распределения** $F(x)$ случайной величины ξ называется вероятность того, что случайная величина будет меньше указанного аргумента:

$$F(x) = P\{\xi < x\}$$

Замечание. Т.к. функция распределения в каждой точке равна соответствующей вероятности, то естественно ожидать, что ее значения будут лежать в диапазоне от 0 до 1.

Замечание. Функция распределения обладает свойством *кумулятивности*. Т.е. значение функции от большего аргумента будет включать в себя значение от меньшего аргумента. Или же можно сказать, что значение функции постепенно «накапливается», представляя собой префиксную сумму. Такое свойство особенно полезно для работы с диапазонами значений случайной величины, т.к. вероятность попадания на промежуток равна разности значений функции распределения на его концах.

Свойства функции распределения:

1°. $0 \leq F(x) \leq 1$

2°. $F(-\infty) = 0; F(+\infty) = 1$

3°. $P\{x_1 \leq \xi < x_2\} = F(x_2) - F(x_1)$

4°. $F(x)$ — неубывающая функция

5°. $F(x)$ — непрерывна слева

Пример 5.2. Для д.с.в. ξ , заданной следующим законом распределения, определить функцию распределения, построить ее график, а также найти вероятность попадания ξ на интервал $(1; 7)$.

ξ	-1	0	2	5
p	0.2	0.5	0.1	0.2

Решение. Пользуясь определением 5.5, составим кусочную функцию распределения д.с.в. ξ :

$$F(x) = \begin{cases} 0, & x \leq -1 \\ 0.2, & -1 < x \leq 0 \\ 0.7, & 0 < x \leq 2 \\ 0.8, & 2 < x \leq 5 \\ 1, & x > 5 \end{cases}$$

По аналитическому заданию функции нетрудно построить ее график:

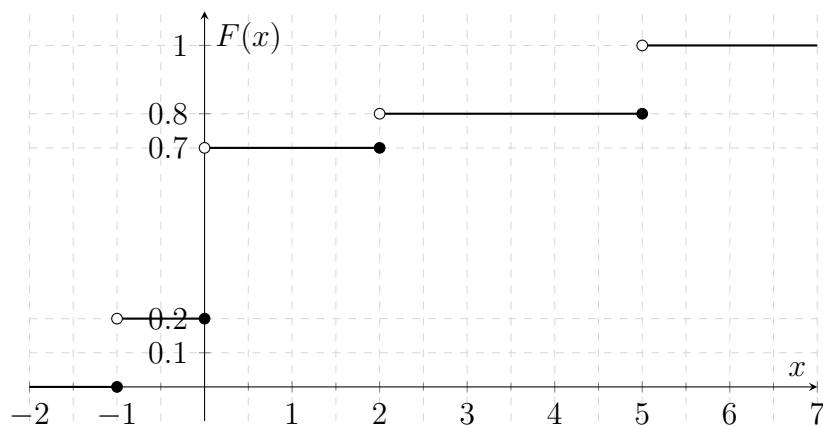


Рисунок 14. График функции распределения д.с.в. ξ

Теперь найдем вероятность попадания ξ на интервал $(1; 7)$. Для этого достаточно найти разность значений функции распределения:

$$P\{\xi \in (1; 7)\} = P\{1 < \xi < 7\} = P\{\xi < 7\} - P\{\xi < 1\} = F(7) - F(1) = 1 - 0.7 = 0.3$$

□

5.3. Числовые характеристики д.с.в

Для извлечения определенной информации о случайных величинах традиционно используют несколько числовых характеристик: математическое ожидание, дисперсия, среднее квадратическое отклонение, корреляционный момент (ковариация).

Определение 5.6. Математическим ожиданием $M(\xi)$ дискретной случайной величины ξ называют взвешенную сумму ее значений (сумму произведений значений на соответствующие вероятности):

$$M(\xi) = \sum_{i=1}^n x_i \cdot p_i$$

Замечание. В некоторых источниках можно встретить и другие обозначения для математического ожидания: $M_\xi, E[\xi]$.

Замечание. Математическое ожидание стоит понимать как средневзвешенное значение случайной величины. Т.е. ее наиболее «ожидаемое» значение.

Свойства математического ожидания:

1°. $M(C) = C$, где C - константа

2°. $M(C \cdot \xi) = C \cdot M(\xi)$ — вынос константы

3°. $M(\xi + \eta) = M(\xi) + M(\eta)$ — свойство линейности

4°. $M(\xi \cdot \eta) = M(\xi) \cdot M(\eta)$ — только для независимых с.в. ξ и η

Определение 5.7. Дисперсией случайной величины называется математическое ожидание квадрата ее отклонения от математического ожидания:

$$D(\xi) = M[(\xi - M(\xi))^2]$$

, или же:

$$D(\xi) = M(\xi^2) - (M(\xi))^2$$

Для дискретной с.в. справедлива также формула:

$$D(\xi) = \sum_{i=1}^n (x_i - M(\xi))^2 \cdot p_i$$

Замечание. Для оценки степени «разброса» с.в. относительно ее мат.ожидания нецелесообразно рассматривать разницу с.в. и мат.ожидания, т.к. в среднем такая разница в силу случайности равна нулю. Чтобы предотвратить такое обнуление, изучают мат.ожидание квадрата отклонения. Таким образом, дисперсия представляет собой ожидаемое значение квадрата отклонения. Т.е. это величина квадратичная.

Свойства дисперсии:

1°. $D(\xi) \geq 0$

2°. $D(C) = 0$, где C — константа

3°. $D(C \cdot \xi) = C^2 \cdot D(\xi)$

4°. $D(\xi + \eta) = D(\xi) + D(\eta) + 2 \cdot K(\xi, \eta)$, где K — ковариация с.в. ξ и η

5°. $D(\xi + \eta) = D(\xi) + D(\eta)$ — для независимых с.в. ξ и η

Определение 5.8. Средним квадратическим отклонением случайной величины ξ называют квадратный корень из ее дисперсии:

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

Замечание. Среднее квадратическое отклонение — это прямая характеристика среднего отклонения величины от ее математического ожидания. Основным ожидаемым разброс значений с.в. будет располагаться на интервале $(M_\xi - \sigma_\xi; M_\xi + \sigma_\xi)$.

Определение 5.9. Ковариацией (корреляционным моментом) двух случайных величин ξ и η называется математическое ожидание произведения «центрированных» случайных величин ξ и η :

$$K(\xi, \eta) = M[(\xi - M(\xi)) \cdot (\eta - M(\eta))]$$

, или же

$$K(\xi, \eta) = M(\xi \cdot \eta) - M(\xi) \cdot M(\eta)$$

Замечание. В некоторых источниках ковариация также может обозначаться как $cov(\xi, \eta)$.

Замечание. Ковариация показывает степень созависимости случайных величин. Т.е. то, насколько значения одной случайной величины влияют на значение другой. Если случайные величины заведомо независимы, то их ковариация равна нулю. Более подробно эту числовую характеристику рассмотрим в главе «двумерные случайные величины».

Пример 5.3. Для д.с.в. ξ , заданной следующим законом распределения, определить математическое ожидание, дисперсию, среднее квадратическое отклонение, а также найти ковариацию ξ с постоянной случайной величиной: $\eta = 5$.

ξ	-1	0	2	5
p	0.2	0.5	0.1	0.2

Решение. Согласно формуле мат.ожидания для д.с.в. имеем:

$$M(\xi) = \sum_{i=1}^n x_i \cdot p_i = -1 \cdot 0.2 + 0 \cdot 0.5 + 2 \cdot 0.1 + 5 \cdot 0.2 = 1$$

Для нахождения дисперсии посчитаем мат.ожидание квадрата случайной величины:

$$M(\xi^2) = \sum_{i=1}^n x_i^2 \cdot p_i = (-1)^2 \cdot 0.2 + 0^2 \cdot 0.5 + 2^2 \cdot 0.1 + 5^2 \cdot 0.2 = 5.6$$

Теперь можем вычислить дисперсию по упрощенной формуле:

$$D(\xi) = M(\xi^2) - (M(\xi))^2 = 5.6 - 1^2 = 4.6$$

Среднее квадратическое отклонение:

$$\sigma(\xi) = \sqrt{D(\xi)} = \sqrt{4.6} \approx 2.1448$$

Для вычисления ковариации сначала вычислим мат.ожидание произведения случайных величин. Для этого рассмотрим все возможные комбинации значений с.в. и их вероятности. Т.к. величина η постоянна и имеет значение 5 с вероятностью 1, достаточно будет рассмотреть только четыре ситуации:

η	5	5	5	5
ξ	-1	0	2	5
p	0.2	0.5	0.1	0.2

Математическое ожидание произведения с.в. вычисляется как:

$$M(\xi \cdot \eta) = \sum_{i=1}^n \sum_{j=1}^m x_i \cdot y_j \cdot p_{ij} = 5 \cdot (-1 \cdot 0.2 + 0 \cdot 0.5 + 2 \cdot 0.1 + 5 \cdot 0.2) = 5$$

Математическое ожидание константной случайной величины η :

$$M(\eta) = M(5) = 5$$

Тогда ковариация:

$$K(\xi, \eta) = M(\xi \cdot \eta) - M(\xi) \cdot M(\eta) = 5 - 1 \cdot 5 = 0$$

Такой же результат можно было бы получить из факта независимости данных с.в. □

5.4. Непрерывные случайные величины

Определение 5.10. Случайная величина ξ называется *непрерывной*, если ее функция распределения $F(x)$ непрерывна и обладает кусочно-непрерывной производной $F'(x)$.

Вероятность попадания непрерывной случайной величины в заданный промежуток зависит от скорости роста функции распределения. Поэтому непрерывную случайную величину задают при помощи производной от функции распределения.

Определение 5.11. **Плотностью распределения** $f(x)$ н.с.в. ξ называют первую производную от ее функции распределения:

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

Замечание. Плотность распределения — основная характеристика непрерывных случайных величин. По своему значению она эквивалентна закону распределения у дискретных случайных величин, обеспечивая меру вероятности для значений случайной величины.

Свойства плотности распределения:

1°. $f(x) \geq 0$ — плотность не может вносить отрицательную вероятность

2°. $f(-\infty) = f(+\infty) = 0$

3°. $f(x)$ — кусочно-непрерывная функция

4°. $F(x) = \int_{-\infty}^x f(t)dt$ — определение функции распределения через плотность

5°. $P\{x_1 \leq \xi < x_2\} = \int_{x_1}^{x_2} f(x)dx$ — вычисление вероятности через плотность

6°. $\int_{-\infty}^{+\infty} f(x)dx = 1$ — условие нормировки плотности

Пример 5.4. Непрерывная случайная величина ξ задана функцией распределения:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 0.25 \cdot x^2, & 0 < x \leq 2 \\ 1, & x > 2 \end{cases}$$

Требуется найти функцию плотности н.с.в. и проверить результат условием нормировки. Определить вероятность того, что случайная величина ξ примет значение из интервала $(1; 3)$

Решение. Функция плотности представляет собой первую производную функции распределения. Поэтому на каждом промежутке кусочного определения найдем производную: для $x \leq 0$ это $0' = 0$, для $0 < x \leq 2$ производная: $(0.25 \cdot x^2)' = 0.5 \cdot x$, для $x > 2$ это $1' = 0$. Т.о.:

$$f(x) = \begin{cases} 0, & x \leq 0 \\ 0.5 \cdot x & 0 < x \leq 2 \\ 0, & x > 2 \end{cases}$$

Проверим корректность нахождения функции плотности при помощи условия нормировки:

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

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x)dx &= \int_{-\infty}^0 f(x)dx + \int_0^2 f(x)dx + \int_2^{+\infty} f(x)dx = \\ &= \int_{-\infty}^0 0 \cdot dx + \int_0^2 0.5 \cdot x dx + \int_2^{+\infty} 0 \cdot dx = 0 + (0.25 \cdot x^2) \Big|_0^2 + 0 = 1 \end{aligned}$$

Т.о., условие нормировки соблюдено.

Вероятность того, что ξ примет значение из интервала $(1; 3)$:

$$\begin{aligned} P\{\xi \in (1; 3)\} &= P\{1 < \xi < 3\} = \int_1^3 f(x)dx = \int_1^2 f(x)dx + \int_2^3 f(x)dx = \\ &= \int_1^2 0.5 \cdot x dx + \int_2^3 0 \cdot dx = (0.25 \cdot x^2) \Big|_1^2 + 0 = 1 - 0.25 = 0.75 \end{aligned}$$

□

Пример 5.5. Задана функция распределения н.с.в. ξ :

$$F(x) = \begin{cases} 0, & x \leq -1 \\ A \cdot (x^3 + 1), & -1 < x \leq 3 \\ 1, & x > 3 \end{cases}$$

Найти значение параметра A и вероятность того, что случайная величина ξ примет значение из интервала $(-2; 2)$. Построить график функции распределения.

Решение. Функция распределения н.с.в. обязательно должна быть непрерывной. Исходя из этого факта, определим параметр A так, чтобы односторонние пределы функции в окрестности $x = 3$ были равны, тем самым обеспечивая непрерывность функции:

$$\begin{aligned} \lim_{x \rightarrow 3-0} A \cdot (x^3 + 1) &= \lim_{x \rightarrow 3+0} 1 \\ A \cdot (27 + 1) &= 1 \\ A &= \frac{1}{28} \end{aligned}$$

Т.о., функция распределения имеет вид:

$$F(x) = \begin{cases} 0, & x \leq -1 \\ \frac{1}{28} \cdot (x^3 + 1), & -1 < x \leq 3 \\ 1, & x > 3 \end{cases}$$

Вероятность попадания случайной величины на интервал можно считать как интеграл от функции плотности, а можно и как разность значений функции распределения (эквивалентно подстановке Ньютона-Лейбница). Поэтому:

$$P\{\xi \in (-2; 2)\} = P\{-2 < \xi < 2\} = F(2) - F(-2) = \frac{1}{28} \cdot (2^3 + 1) - 0 = \frac{9}{28}$$

Напоследок построим график функции распределения $F(x)$:

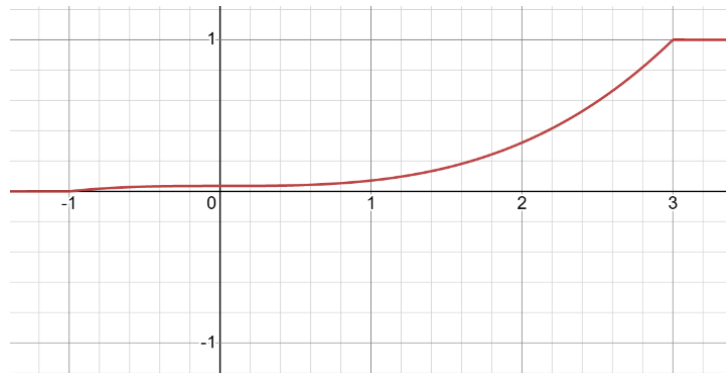


Рисунок 15. График функции распределения $F(x)$ к примеру 5.5

□

Пример 5.6. Задана плотность распределения н.с.в. ξ :

$$f(x) = \begin{cases} 0, & x \leq 0 \\ 0.5x, & 0 < x \leq 1 \\ A \cdot (2x - x^2), & 1 < x \leq 2 \\ 1, & x > 2 \end{cases}$$

Найти значение параметра A и функцию распределения н.с.в. Построить графики функций распределения и плотности.

Решение. Параметр A будем определять исходя из нормировки функции плотности. Т.е. «подгоним» параметр под условие нормировки:

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x) dx &= 1 \\ \int_{-\infty}^{+\infty} f(x) dx &= \int_{-\infty}^0 f(x) dx + \int_0^1 f(x) dx + \int_1^2 f(x) dx + \int_2^{+\infty} f(x) dx = \\ &= \int_{-\infty}^0 0 \cdot dx + \int_0^1 0.5x \cdot dx + \int_1^2 A \cdot (2x - x^2) dx + \int_2^{+\infty} 0 \cdot dx = \\ &= 0.25 \cdot x^2 \Big|_0^1 + A(x^2 - x^3/3) \Big|_1^2 = 0.25 + \frac{2}{3} \cdot A \\ 0.25 + \frac{2}{3} \cdot A &= 1 \implies A = \frac{9}{8} \end{aligned}$$

Тогда функция плотности имеет вид:

$$f(x) = \begin{cases} 0, & x \leq 0 \\ 0.5x, & 0 < x \leq 1 \\ \frac{9}{8} \cdot (2x - x^2), & 1 < x \leq 2 \\ 0, & x > 2 \end{cases}$$

Вычислим функцию распределения н.с.в. Для этого для каждого участка воспользуемся формулой:

$$F(x) = \int_{-\infty}^x f(t)dt$$

Для $x \leq 0$:

$$F(x) = \int_{-\infty}^x 0 \cdot dt = 0$$

Для $0 < x \leq 1$:

$$F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^0 0 \cdot dt + \int_0^x 0.5tdt = 0.25x^2$$

Для $1 < x \leq 2$:

$$F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^0 0 \cdot dt + \int_0^1 0.5tdt + \int_1^x \frac{9}{8} \cdot (2t - t^2)dt = 0.25 + \frac{9}{8}(t^2 - t^3/3) \Big|_1^x = -\frac{3}{8}x^3 + \frac{9}{8}x^2 - 0.5$$

Для $x > 2$:

$$F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^0 0 \cdot dt + \int_0^1 0.5tdt + \int_1^2 \frac{9}{8} \cdot (2t - t^2)dt + \int_2^x 0 \cdot dt = 1$$

Т.о., функция распределения имеет вид:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 0.25x^2, & 0 < x \leq 1 \\ -\frac{3}{8}x^3 + \frac{9}{8}x^2 - 0.5, & 1 < x \leq 2 \\ 1, & x > 2 \end{cases}$$

Построим графики функций распределения и плотности н.с.в. ξ :

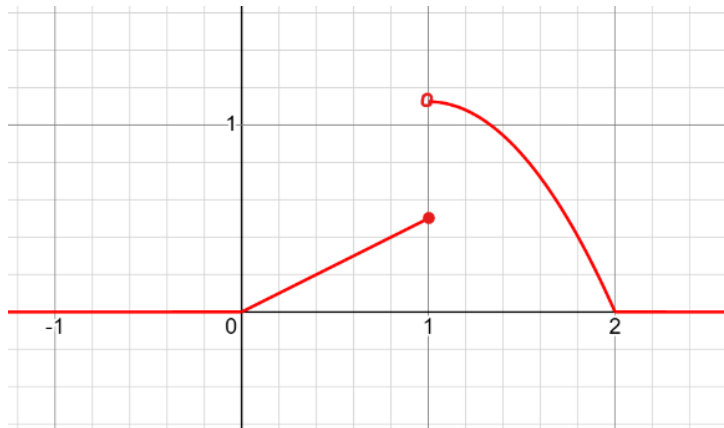


Рисунок 16. График функции плотности распределения $f(x)$ к примеру 5.6

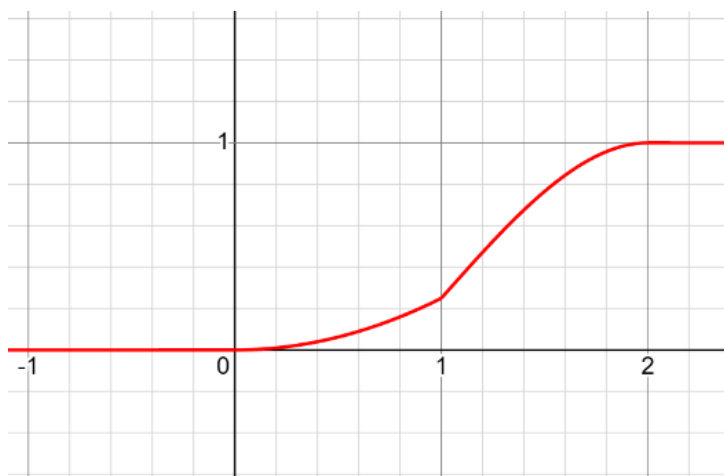


Рисунок 17. График функции распределения $F(x)$ к примеру 5.6

□

5.5. Числовые характеристики н.с.в

Определение 5.12. Математическим ожиданием непрерывной случайной величины ξ с плотностью распределения $f(x)$ называется:

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

В более общем виде данная формула справедлива для нахождения мат.ожидания любой н.с.в., зависящей от данной:

$$\eta = \varphi(\xi)$$

$$M(\eta) = M(\varphi(\xi)) = \int_{-\infty}^{+\infty} \varphi(x) \cdot f(x) dx$$

Определение 5.13. Дисперсией н.с.в. ξ все так же называется мат.ожидание квадрата отклонения ξ от ее мат.ожидания:

$$D(\xi) = M[(\xi - M(\xi))^2]$$

Для н.с.в. справедлива формула:

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

, или же:

$$D(\xi) = M(\xi^2) - (M(\xi))^2$$

Замечание. Среднее квадратическое отклонение все так же определяется как корень из дисперсии: $\sigma(\xi) = \sqrt{D(\xi)}$.

Пример 5.7. Для н.с.в. ξ , заданной плотностью распределения $f(x)$, определить мат.ожидание, дисперсию и среднее квадратическое отклонение.

$$f(x) = \begin{cases} 0, & x \leq 0 \\ 0.5x, & 0 < x \leq 1 \\ \frac{9}{8} \cdot (2x - x^2), & 1 < x \leq 2 \\ 0, & x > 2 \end{cases}$$

Решение. Вычислим математическое ожидание согласно интегральной формуле:

$$\begin{aligned} M(\xi) &= \int_{-\infty}^{+\infty} x \cdot f(x) dx \\ M(\xi) &= 0 + \int_0^1 x \cdot 0.5x \cdot dx + \int_1^2 x \cdot \frac{9}{8} \cdot (2x - x^2) \cdot dx + 0 = \\ &= (0.5 \cdot x^3/3) \Big|_0^1 + \frac{9}{8} \cdot (2x^3/3 - x^4/4) \Big|_1^2 = \frac{1}{6} + \frac{33}{32} = \frac{115}{96} \approx 1.198 \end{aligned}$$

Для нахождения дисперсии вычислим сначала мат.ожидание квадрата н.с.в:

$$\begin{aligned} M(\xi^2) &= \int_{-\infty}^{+\infty} x^2 \cdot f(x) dx \\ M(\xi^2) &= 0 + \int_0^1 x^2 \cdot 0.5x \cdot dx + \int_1^2 x^2 \cdot \frac{9}{8} \cdot (2x - x^2) \cdot dx + 0 = \\ &= (0.5 \cdot x^4/4) \Big|_0^1 + \frac{9}{8} \cdot (x^4/2 - x^5/5) \Big|_1^2 = \frac{1}{8} + \frac{117}{80} = \frac{127}{80} \approx 1.588 \end{aligned}$$

Дисперсия н.с.в:

$$D(\xi) = M(\xi^2) - (M(\xi))^2 = 1.588 - 1.198^2 = 0.1528$$

Среднее квадратическое отклонение:

$$\sigma(\xi) = \sqrt{D(\xi)} = \sqrt{0.1528} = 0.391$$

□

6. Распределения случайных величин

В данной главе рассмотрим основные виды распределений случайных величин, традиционно встречающиеся в теории вероятностей.

6.1. Производящие функции

Определение 6.1. Производящей функцией для д.с.в. ξ называется функция вида:

$$\varphi(z) = \sum_{k=0}^{\infty} p_k \cdot z^k = p_0 + p_1 z + p_2 z^2 + \dots$$

, где z — произвольный параметр, $z \in (0; 1]$

Производящие функции — это аналитический инструмент для работы с д.с.в., принимающими натуральные значения. Они помогают получать упрощенные формулы для численных характеристик д.с.в. Так, математическое ожидание д.с.в. — это значение первой производной производящей функции в точке 1. Они пригодятся нам для вывода кратких формул мат.ожидания и дисперсии различных дискретных распределений.

6.2. Биномиальное распределение

Рассмотрим n повторных независимых испытаний. Пусть в каждом из них некое событие A возникает с вероятностью p . (Такие испытания называются испытаниями Бернулли). Обозначим за случайную величину ξ число появлений такого события A в n испытаниях. Тогда по уже рассмотренной формуле Бернулли имеем:

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

Определение 6.2. Биномиальным называется распределение дискретной случайной величины, вероятности значений которой определяются по формуле Бернулли. Биномиальное распределение задается таблицей:

ξ	0	1	2	...	k	...	n
p	q^n	$n p q^{n-1}$	$C_n^2 p^2 q^{n-2}$...	$C_n^k p^k q^{n-k}$...	p^n

Замечание. Биномиальное распределение характеризуется двумя параметрами: n — числом испытаний и p — вероятностью возникновения ожидаемого события в каждом отдельном испытании.

Замечание. Производящей функцией для биномиального распределения является:

$$\varphi(z) = (q + pz)^n$$

Теорема 6.1. Математическое ожидание для биномиально распределенной случайной величины можно вычислить по формуле:

$$M(\xi) = np$$

А дисперсию:

$$D(\xi) = npq$$

Пример 6.1. Вероятность того, что стрелок при выстреле попадет в мишень, составляет 0.6. Случайная величина ξ — число успешных попаданий в мишень в результате трех выстрелов. Составить закон распределения д.с.в. ξ . Найти мат.ожидание и дисперсию. Нарисовать многоугольник распределения ξ .

Решение. Для величины ξ представимы следующие значения: 0, 1, 2, 3. Вероятность возникновения каждого из них определяется по формуле Бернулли. А значит, распределение является биномиальным. Найдем вероятности возникновения каждого из значений:

$$P\{\xi = 0\} = C_3^0 \cdot (0.6)^0 \cdot (0.4)^3 = 0.064; \quad P\{\xi = 1\} = C_3^1 \cdot (0.6)^1 \cdot (0.4)^2 = 0.288$$

$$P\{\xi = 2\} = C_3^2 \cdot (0.6)^2 \cdot (0.4)^1 = 0.432; \quad P\{\xi = 3\} = C_3^3 \cdot (0.6)^3 \cdot (0.4)^0 = 0.216$$

Составим закон распределения:

ξ	0	1	2	3
p	0.064	0.288	0.432	0.216

Посчитаем мат.ожидание и дисперсию по упрощенным формулам:

$$M(\xi) = np = 3 \cdot 0.6 = 1.8$$

$$D(\xi) = npq = 3 \cdot 0.6 \cdot 0.4 = 0.72$$

Построим многоугольник распределения. Для этого достаточно перенести данные из закона распределения на график:

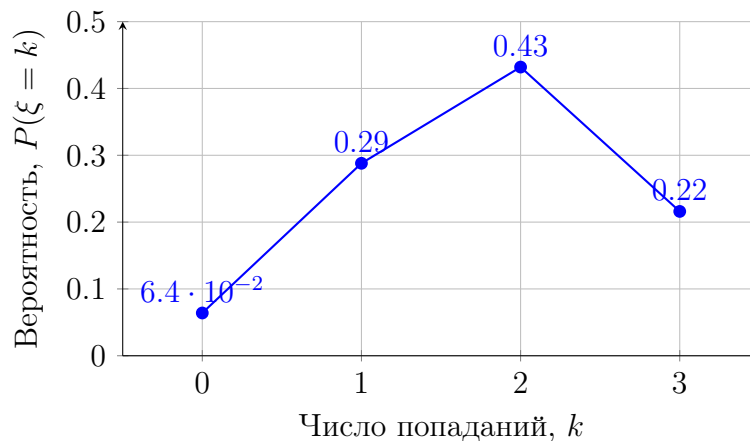


Рисунок 18. График многоугольника биномиального распределения

□

6.3. Распределение Пуассона

Пусть в испытаниях Бернулли $n \rightarrow \infty$, $p \rightarrow 0$. Стандартная формула Бернулли при таких условиях трудновычислима и дает большие погрешности. Поэтому введем

$\lambda = np$ — относительно небольшое число. Тогда вероятность возникновения ожидаемого события m раз в n испытаниях $P_n(m)$ приближенно определяется по уже изученной формуле Пуассона:

$$P\{\xi = m\} = \frac{\lambda^m}{m!} \cdot e^{-\lambda}$$

Определение 6.3. Пуассоновским называется распределение дискретной случайной величины, вероятности значений которой определяются по формуле Пуассона. Распределение Пуассона задается таблицей:

ξ	0	1	2	...	k	...
p	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2}{2!} e^{-\lambda}$...	$\frac{\lambda^k}{k!} e^{-\lambda}$...

Замечание. Распределение Пуассона характеризуется одним параметром: $\lambda = np$, по сути являющимся мат.ожиданием для испытаний Бернулли.

Замечание. Производящей функцией для распределения Пуассона является:

$$\varphi(z) = e^{\lambda(z-1)}$$

Теорема 6.2. Математическое ожидание для пуассоновского распределения можно вычислить по формуле:

$$M(\xi) = \lambda$$

Дисперсия же в силу $q \rightarrow 1$ также равна:

$$D(\xi) = \lambda$$

Пример 6.2. Высокоточный фрезерный ЧПУ-станок выпускает 1000 шестерней за смену. Вероятность того, что шестерня окажется бракованной, составляет 0.008. Случайная величина ξ — число бракованных шестерней среди партии, выпущенной за смену. Составить закон распределения д.с.в. ξ . Найти мат.ожидание и дисперсию. Нарисовать многоугольник распределения ξ .

Решение. В данной задаче число «испытаний» $n = 1000$, а вероятность $p = 0.008$. Величины слишком различаются по порядку для использования формулы Бернулли, а их произведение $\lambda = np = 8$ является числом небольшим. Значит, данная ситуация удовлетворяет условию применения формулы Пуассона и случайная величина ξ имеет пуассоновское распределение.

Вычислим вероятности нескольких значений случайной величины:

$$P\{\xi = 0\} = \frac{8^0}{0!} \cdot e^{-8} = 0.0003; \quad P\{\xi = 1\} = \frac{8^1}{1!} \cdot e^{-8} = 0.0027$$

$$P\{\xi = 2\} = \frac{8^2}{2!} \cdot e^{-8} = 0.0107; \quad P\{\xi = 7\} = \frac{8^7}{7!} \cdot e^{-8} = 0.1396$$

$$P\{\xi = 8\} = \frac{8^8}{8!} \cdot e^{-8} = 0.1396; \quad P\{\xi = 9\} = \frac{8^9}{9!} \cdot e^{-8} = 0.1241$$

Составим закон распределения в общей форме:

ξ	0	1	2	...	7	8	9	...	k	...
p	0.0003	0.0027	0.0107	...	0.1396	0.1396	0.1241	...	$\frac{8^k}{k!} \cdot e^{-8}$...

Согласно упрощенным формулам, мат.ожидание и дисперсия д.с.в. ξ :

$$M(\xi) = D(\xi) = \lambda = 8$$

Построим многоугольник распределения:

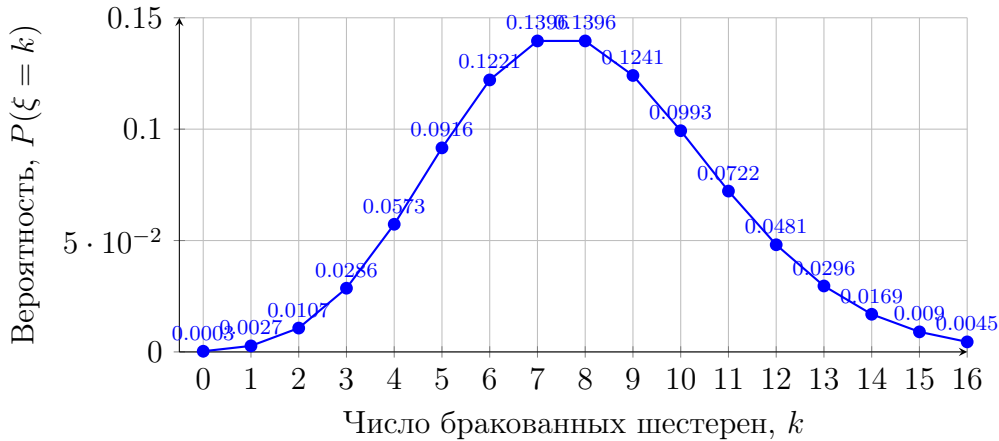


Рисунок 19. График многоугольника распределения Пуассона

□

6.4. Геометрическое распределение

Рассмотрим случай, когда независимые испытания производятся до тех пор, пока в них не возникнет некоторое событие A . При этом в каждой попытке A может возникнуть с вероятностью p . Тогда число попыток до появления события A (включая успешную) является дискретной случайной величиной со значениями от 1 до ∞ . Вероятности таких значений по теореме умножения вероятностей для независимых событий равны:

$$P\{\xi = m\} = pq^{m-1}$$

Определение 6.4. Геометрическим называется распределение дискретной случайной величины, вероятности значений которой определяются по формуле m -го члена геометрической прогрессии с базой p и знаменателем q . Геометрическое распределение задается таблицей:

ξ	1	2	3	...	k	...
p	p	pq	pq^2	...	pq^{k-1}	...

Замечание. Геометрическое распределение характеризуется параметром p — вероятностью наступления ожидаемого события, а также вытекающим из него $q = 1 - p$.

Замечание. Производящей функцией для геометрического распределения является:

$$\varphi(z) = \frac{pz}{1 - zq}$$

Теорема 6.3. Математическое ожидание для геометрически распределенной случайной величины можно вычислить по формуле:

$$M(\xi) = \frac{1}{p}$$

А дисперсию:

$$D(\xi) = \frac{q}{p^2}$$

Пример 6.3. Футболист бьет мячи в ворота до первого попадания. Вероятность попадания в ворота при ударе составляет 0.7. Случайная величина ξ — число совершенных ударов. Составить закон распределения д.с.в. ξ . Найти мат.ожидание и дисперсию. Нарисовать многоугольник распределения ξ .

Решение. Вероятность попадания при каждом ударе: $p = 0.7$. Удары совершаются до первого попадания, значит, д.с.в. имеет геометрическое распределение. Вычислим вероятности нескольких первых значений д.с.в.:

$$P\{\xi = 1\} = 0.7 \cdot 0.3^{1-1} = 0.7; \quad P\{\xi = 2\} = 0.7 \cdot 0.3^{2-1} = 0.21$$

$$P\{\xi = 3\} = 0.7 \cdot 0.3^{3-1} = 0.063; \quad P\{\xi = 4\} = 0.7 \cdot 0.3^{4-1} = 0.0189$$

Составим закон распределения:

ξ	1	2	3	4	...	k	...
p	0.7	0.21	0.063	0.0189	...	$0.7 \cdot 0.3^{k-1}$...

Посчитаем мат.ожидание и дисперсию по упрощенным формулам:

$$M(\xi) = \frac{1}{p} = \frac{1}{0.7} = 1.4286$$

$$D(\xi) = \frac{q}{p^2} = \frac{0.3}{0.49} = 0.6122$$

Построим многоугольник распределения:

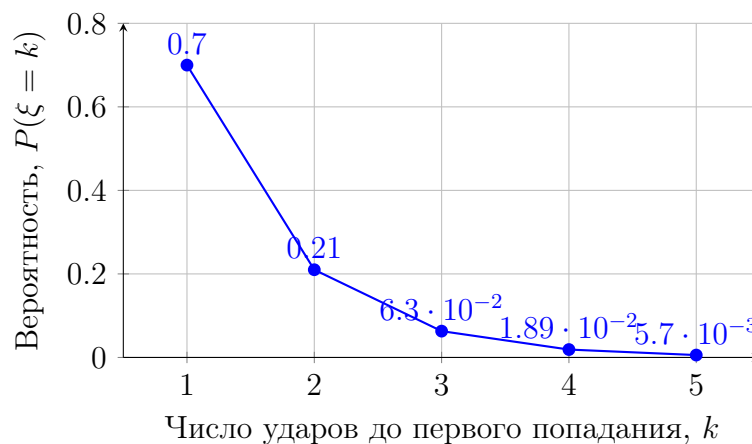


Рисунок 20. График многоугольника геометрического распределения

□

6.5. Гипергеометрическое распределение

Вспомним решение задачи о выборке (раздел 2.3 данного пособия): Имеется N однородных предметов. K из них «особые». Произвольным образом из N выбираются n предметов. Тогда вероятность того, что среди выбранных n предметов ровно k будут «особыми»:

$$P_n(k) = \frac{C_K^k \cdot C_{N-K}^{n-k}}{C_N^n}$$

Если же число k заранее неизвестно и является случайным, то можно ввести случайную величину ξ , равную кол-ву «особых» предметов в выборке.

Определение 6.5. Гипергеометрическим называется распределение дискретной случайной величины, вероятности значений которой определяются по формуле решения задачи о выборке. Гипергеометрическое распределение задается таблицей:

ξ	0	1	2	...	m
p	$\frac{C_K^0 \cdot C_{N-K}^n}{C_N^n}$	$\frac{C_K^1 \cdot C_{N-K}^{n-1}}{C_N^n}$	$\frac{C_K^2 \cdot C_{N-K}^{n-2}}{C_N^n}$...	$\frac{C_K^m \cdot C_{N-K}^{n-m}}{C_N^n}$

, где $m = \min(n; K)$

Замечание. Гипергеометрическое распределение характеризуется параметрами N — общим числом предметов, K — числом «особых» предметов, n — числом выбранных предметов.

Теорема 6.4. Математическое ожидание для гипергеометрического распределения случайной величины можно вычислить по формуле:

$$M(\xi) = n \cdot \frac{K}{N}$$

А дисперсию:

$$D(\xi) = n \cdot \frac{K}{N} \cdot \left(1 - \frac{K}{N}\right) \cdot \frac{N-n}{N-1}$$

Пример 6.4. В ящике находится 20 деталей, 4 из которых бракованные. Случайным образом извлекают 6 деталей без возвращения. Случайная величина ξ — количество бракованных деталей в выборке. Составить закон распределения ξ , найти математическое ожидание и дисперсию. Построить многоугольник распределения.

Решение. Налицо задача о выборке. Поэтому случайная величина имеет гипергеометрическое распределение с параметрами $N = 20$ — общее количество деталей, $K = 4$ — количество бракованных деталей, $n = 6$ — размер выборки.

Вычислим вероятности для всех возможных значений:

$$P\{\xi = 0\} = \frac{C_4^0 \cdot C_{16}^6}{C_{20}^6} \approx 0.2067$$

$$P\{\xi = 1\} = \frac{C_4^1 \cdot C_{16}^5}{C_{20}^6} \approx 0.4508$$

$$P\{\xi = 2\} = \frac{C_4^2 \cdot C_{16}^4}{C_{20}^6} \approx 0.2817$$

$$P\{\xi = 3\} = \frac{C_4^3 \cdot C_{16}^3}{C_{20}^6} \approx 0.0578$$

$$P\{\xi = 4\} = \frac{C_4^4 \cdot C_{16}^2}{C_{20}^6} \approx 0.0030$$

Закон распределения в табличной форме:

ξ	0	1	2	3	4
p	0.2067	0.4508	0.2817	0.0578	0.0030

Математическое ожидание и дисперсия:

$$M(\xi) = n \cdot \frac{K}{N} = 6 \cdot \frac{4}{20} = 1.2$$

$$D(\xi) = n \cdot \frac{K}{N} \cdot \left(1 - \frac{K}{N}\right) \cdot \frac{N - n}{N - 1} = 6 \cdot 0.2 \cdot 0.8 \cdot \frac{14}{19} \approx 0.7074$$

Многоугольник распределения:

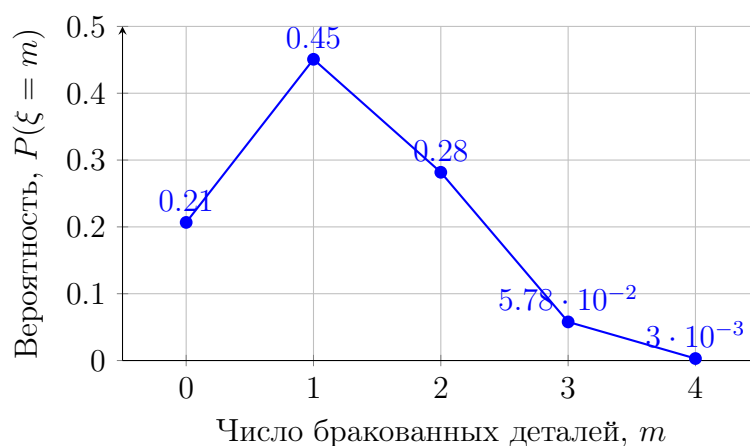


Рисунок 21. График многоугольника гипергеометрического распределения

□

6.6. Равномерное распределение

Перейдем к рассмотрению законов распределения для непрерывных случайных величин.

Определение 6.6. Равномерным называется распределение непрерывной случайной величины, плотность которой постоянна на всей области значений.

Если область значений н.с.в. — это отрезок $[a; b]$, то плотность определяется как:

$$f(x) = \begin{cases} C, & x \in [a; b] \\ 0, & x \notin [a; b] \end{cases}$$

Причем из условия нормировки вытекает значение константы C , так что:

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

А функция распределения определяется как:

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

Замечание. Равномерное распределение характеризуется лишь областью значений н.с.в, т.е. отрезком распределения.

Теорема 6.5. Математическое ожидание для равномерно распределенной случайной величины можно вычислить по формуле:

$$M(\xi) = \frac{a+b}{2}$$

А дисперсию:

$$D(\xi) = \frac{(b-a)^2}{12}$$

Пример 6.5. Автобус приходит на остановку строго по расписанию с интервалом в 20 минут. Пассажир приходит на остановку в случайный момент времени. Случайная величина ξ — время ожидания автобуса. Найти функцию плотности распределения $f(x)$, функцию распределения $F(x)$, математическое ожидание и дисперсию. Построить графики функции распределения и плотности, а также определить вероятность того, что пассажиру придётся ждать от 5 до 15 минут.

Решение. Так как время прихода пассажира на остановку случайно, то все значения величины ξ равновероятны, а значит, она имеет равномерное распределение на отрезке $[0; 20]$.

Согласно определению равномерного распределения, функция плотности имеет вид:

$$f(x) = \begin{cases} \frac{1}{20}, & x \in [0; 20] \\ 0, & x \notin [0; 20] \end{cases}$$

А функция распределения:

$$F(x) = \begin{cases} 0, & \text{если } x \leq 0 \\ \frac{x}{20}, & \text{если } 0 < x \leq 20 \\ 1, & \text{если } x > 20 \end{cases}$$

По упрощенным формулам найдем мат.ожидание н.с.в. ξ :

$$M(\xi) = \frac{a+b}{2} = \frac{0+20}{2} = 10$$

И дисперсию:

$$D(\xi) = \frac{(b-a)^2}{12} = \frac{(20-0)^2}{12} \approx 33.3$$

Вероятность ожидания от 5 до 15 минут представляет собой вероятность попадания ξ на отрезок $[5; 15]$:

$$P\{5 \leq \xi \leq 15\} = F(15) - F(5) = \frac{15}{20} - \frac{5}{20} = 0.5$$

Построим графики функции распределения и плотности н.с.в. ξ :

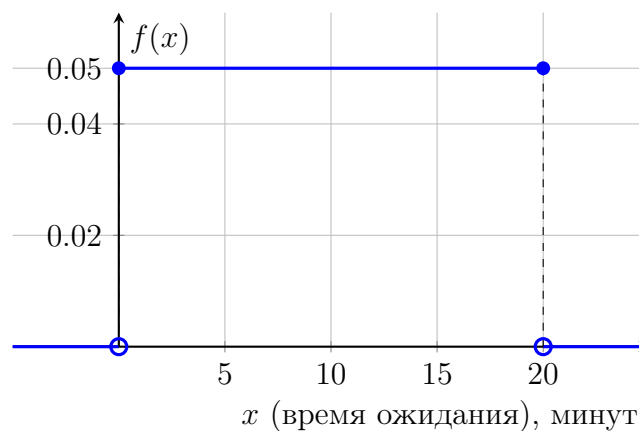


Рисунок 22. График плотности равномерного распределения

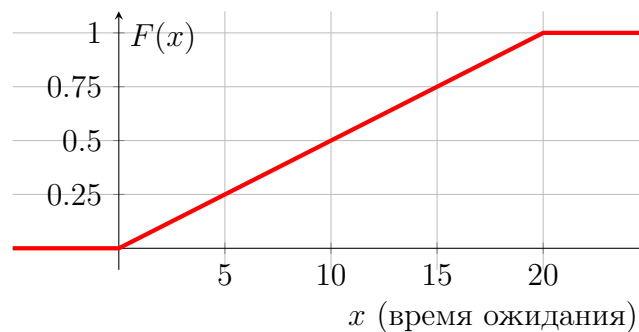


Рисунок 23. График функции равномерного распределения

□

6.7. Показательное распределение

Существует ряд случаев, в которых вероятности значений случайной величины уменьшаются по экспоненте. Тогда говорят про показательное распределение н.с.в.

Определение 6.7. Показательным (экспоненциальным) называется распределение непрерывной случайной величины, плотность которого убывает экспоненциально с ростом значения.

Если н.с.в. ξ описывает время между событиями с интенсивностью $\lambda > 0$, то её плотность имеет вид:

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

Функция распределения выражается формулой:

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

Замечание. Показательное распределение характеризуется одним параметром λ — интенсивностью затухания.

Теорема 6.6. Для показательного распределения математическое ожидание определяется по формуле:

$$M(\xi) = \frac{1}{\lambda}$$

А дисперсия:

$$D(\xi) = \frac{1}{\lambda^2}$$

Пример 6.6. В среднем сервер интернет-магазина обрабатывает 500 запросов в час. Случайная величина ξ — время между двумя последовательными запросами. Определить функцию плотности распределения. Найти вероятность того, что время между запросами: превысит 10 секунд; окажется между 5 и 15 секундами. Построить графики плотности и функции распределения.

Решение.

Нам дана средняя часовая нагрузка на сервер. Значит, можем определить среднее время между запросами:

$$M(\xi) = \frac{3600 \text{ [сек]}}{500} = 7.2$$

Параметр интенсивности:

$$\lambda = \frac{1}{M(\xi)} = \frac{1}{7.2} \approx 0.1389$$

Функция плотности имеет вид:

$$f(x) = \begin{cases} 0.1389 e^{-0.1389x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

А функция распределения:

$$F(x) = \begin{cases} 1 - e^{-0.1389x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Вероятность того, что время между запросами превысит 10 секунд:

$$P\{\xi > 10\} = 1 - P\{\xi \leq 10\} = 1 - F(10) = 1 - (1 - e^{-0.1389 \cdot 10}) = e^{-1.389} \approx 0.249$$

Вероятность того, что время между запросами будет в промежутке (5; 15):

$$P\{5 < \xi < 15\} = F(15) - F(5) = e^{-0.1389 \cdot 5} - e^{-0.1389 \cdot 15} \approx 0.497 - 0.123 = 0.374$$

Построим графики функций распределения и плотности:

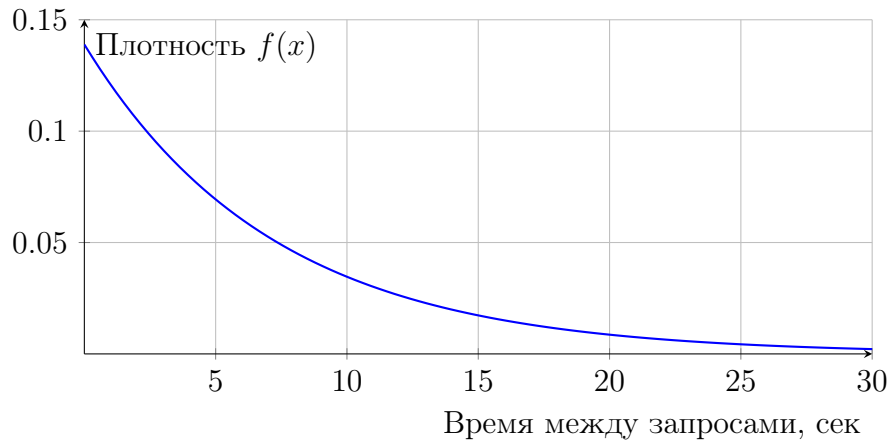


Рисунок 24. График плотности показательного распределения

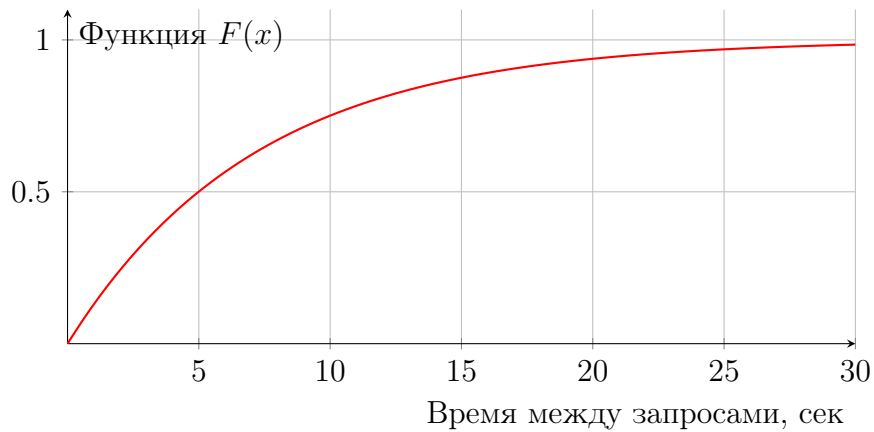


Рисунок 25. График функции показательного распределения

□

6.8. Нормальное распределение

7. Двумерные случайные величины

7.1. Дискретные двумерные с.в.

7.2. Непрерывные двумерные с.в.

7.3. Корреляционная зависимость случайных величин

8. Основы математической статистики

8.1. Основные понятия статистики

Таблица значений функции Гаусса: $f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2}$

	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3823
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	7802	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
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

[illegible]

Приложение В. Таблица значений функции Лапласа

Таблица значений функции Лапласа: $\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^x e^{-z^2/2} dz$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000	0,32	0,1255	0,64	0,2389	0,96	0,3315
0,01	0,0040	0,33	0,1293	0,65	0,2422	0,97	0,3340
0,02	0,0080	0,34	0,1331	0,66	0,2454	0,98	0,3365
0,03	0,0120	0,35	0,1368	0,67	0,2486	0,99	0,3389
0,04	0,0160	0,36	0,1406	0,68	0,2517	1,00	0,3413
0,05	0,0199	0,37	0,1443	0,69	0,2549	1,01	0,3438
0,06	0,0239	0,38	0,1480	0,70	0,2580	1,02	0,3461
0,07	0,0279	0,39	0,1517	0,71	0,2611	1,03	0,3485
0,08	0,0319	0,40	0,1554	0,72	0,2642	1,04	0,3508
0,09	0,0359	0,41	0,1591	0,73	0,2673	1,05	0,3531
0,10	0,0398	0,42	0,1628	0,74	0,2703	1,06	0,3554
0,11	0,0438	0,43	0,1664	0,75	0,2734	1,07	0,3577
0,12	0,0478	0,44	0,1700	0,76	0,2764	1,08	0,3599
0,13	0,0517	0,45	0,1736	0,77	0,2794	1,09	0,3621
0,14	0,0557	0,46	0,1772	0,78	0,2823	1,10	0,3643
0,15	0,0596	0,47	0,1808	0,79	0,2852	1,11	0,3665
0,16	0,0636	0,48	0,1844	0,80	0,2881	1,12	0,3686
0,17	0,0675	0,49	0,1879	0,81	0,2910	1,13	0,3708
0,18	0,0714	0,50	0,1915	0,82	0,2939	1,14	0,3729
0,19	0,0753	0,51	0,1950	0,83	0,2967	1,15	0,3749
0,20	0,0793	0,52	0,1985	0,84	0,2995	1,16	0,3770
0,21	0,0832	0,53	0,2019	0,85	0,3023	1,17	0,3790
0,22	0,0871	0,54	0,2054	0,86	0,3051	1,18	0,3810
0,23	0,0910	0,55	0,2088	0,87	0,3078	1,19	0,3830
0,24	0,0948	0,56	0,2123	0,88	0,3106	1,20	0,3849
0,25	0,0987	0,57	0,2157	0,89	0,3133	1,21	0,3869
0,26	0,1026	0,58	0,2190	0,90	0,3159	1,22	0,3883
0,27	0,1064	0,59	0,2224	0,91	0,3186	1,23	0,3907
0,28	0,1103	0,60	0,2257	0,92	0,3212	1,24	0,3925
0,29	0,1141	0,61	0,2291	0,93	0,3238	1,25	0,3944
0,30	0,1179	0,62	0,2324	0,94	0,3264		
0,31	0,1217	0,63	0,2357	0,95	0,3289		

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,26	0,3962	1,59	0,4441	1,92	0,4726	2,50	0,4938
1,27	0,3980	1,60	0,4452	1,93	0,4732	2,52	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	0,499997