

PT LEC 03

isagila

Собрано 29.10.2023 в 14:08



Содержание

1. Лекции	3
1.1. Лекция 23.09.05.	3
1.2. Лекция 23.09.12.	5
1.3. Лекция 23.09.19.	9
1.4. Лекция 23.09.26.	12
1.5. Лекция 23.10.03.	16
1.6. Лекция 23.10.10.	19
1.7. Лекция 23.10.17.	22
1.8. Лекция 23.10.24.	27
1.9. Лекция 23.10.31.	33

TODO: Хочу на cover поставить другой арт с Тосакой

1. Лекции

1.1. Лекция 23.09.05.

Пусть проводится n реальных экспериментов, в которых событие A появилось n_A раз. Отношение $P(A) = \frac{n_A}{n}$ называется частотой события A . Эксперименты показывают, что частота «стабилизируется» около некоторого числа, под которым и подразумевается статистическая вероятность.

Def 1.1.1. Пространством элементарных исходов Ω называется множество, содержащее все возможные результаты данного эксперимента, из которых при испытании происходит ровно один. Элементы этого множества называются элементарными исходами и обозначаются ω .

Def 1.1.2. Случайными событиями называются подмножества $A \subseteq \Omega$. Говорят, что в ходе эксперимента событие A наступило, если произошел один из элементарных исходов, принадлежащий A .

Пример 1.1.3. 1. Подбрасывание монеты. $\Omega = \{\Gamma, P\}$

2. Подбрасывание кубика. $\Omega = \{1, 2, 3, 4, 5, 6\}$, тогда, например, 2 — это элементарный исход, а $A = \{2, 4, 6\}$ это событие «выпало четное число».
3. Подбрасывание монеты дважды. Если учитывать порядок бросков, то $\Omega = \{\Gamma\Gamma, \Gamma P, P\Gamma, PP\}$, если не учитывать порядок бросков, то $\Omega = \{\Gamma\Gamma, \Gamma P, PP\}$.
4. Подбрасывание кубика дважды. $\Omega = \{(i, j) \mid 1 \leq i, j \leq 6\}$. Событию A «разница выпавших очков делится на 3» благоприятствуют исходы $(1, 4), (4, 1), (2, 5), (5, 2), (3, 6), (6, 3), (1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)$.
5. Монета подбрасывается до выпадения герба. $\Omega = \{\Gamma, P\Gamma, PPG, PPPG, \dots\}$ — пространство элементарных исходов бесконечно, но счетно.
6. Монета бросается на координатную плоскость. $\Omega = \{(x, y) \mid x, y \in \mathbb{R}\}$.

Операции над событиями

Выделим два специальных события. Ω — достоверное (универсальное) событие, которое наступает всегда, т.к. содержит все элементарные исходы. Пустое (невозможное) событие \emptyset , которое никогда не наступает.

Def 1.1.4. Суммой $A + B$ называется событие, состоящее в том, что произошло событие A или B .

Def 1.1.5. Произведением AB называется событие, состоящее в том, что произошло событие A и событие B (т.е. оба).

Замечание 1.1.6. Сумма $A_1 + \dots + A_n$ — произошло хотя бы одно из этих событий. Произведение $A_1 \dots A_n$ — произошли все события.

Def 1.1.7. Противоположным к A называется событие \bar{A} , состоящее в том, что событие A не произошло.

Def 1.1.8. Дополнением $A \setminus B$ называется событие, состоящее в том, что событие A произошло, а событие B — нет, т.е. $A \setminus B = A\bar{B}$.

Def 1.1.9. События A и B называются несовместными, если из произведения равно \emptyset , т.е. при одном эксперименте может произойти только одно из этих них.

Def 1.1.10. Событие A влечет событие B , если $A \subseteq B$.

Подходы к определению вероятности

Каждому случайному событию A хотим присвоить числовую характеристику, отражающую частоту события A : $0 \leq P(A) \leq 1$ — вероятность события A .

Подход I. Классическое определение вероятности

Пусть Ω содержит конечное число равновозможных исходов, тогда применимо классическое определение вероятности.

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

где m это число исходов, которые благоприятствуют событию A , а n — число всех возможных исходов. В частности, если A_i — элементарный исход, то $P(A_i) = \frac{1}{n}$. Выделим некоторые свойства.

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

$$2. \quad P(\Omega) = \frac{n}{n} = 1$$

$$3. \quad P(\emptyset) = \frac{0}{n} = 0$$

4. Если A и B — несовместные события, то $P(A + B) = P(A) + P(B)$.

Приведем несложное доказательство 4° свойства.

\square Если A и B — несовместные события, то $|A + B| = |A| + |B|$, значит

$$P(A + B) \frac{|A + B|}{|\Omega|} = \frac{|A| + |B|}{|\Omega|} = P(A) + P(B)$$

■

Пример 1.1.11. Найти вероятность того, что на кубике выпадет четное число при одном броске.

$$\left. \begin{array}{l} \Omega = \{1, 2, 3, 4, 5, 6\} \\ A = \{2, 4, 6\} \end{array} \right\} \Rightarrow P(A) = \frac{|A|}{|\Omega|} = \frac{3}{6} = 0.5$$

Подход II. Геометрическое определение вероятности

Замечание 1.1.12. Определение «меры» не вводим, ограничимся тем, что мерой отрезкой является его длина, плоскости — площадь, а пространства — объем.

Пусть Ω можно изобразить в виде замкнутой ограниченной области $\Omega \subseteq \mathbb{R}^n$ и мера $\mu(\Omega)$ конечна. В эту область наугад бросается точка. Здесь «наугад» означает, что вероятность A зависит лишь от меры A и не зависит от расположения A (грубо говоря, попадание в любую точку равновозможно). При выполнении этих условий применимо геометрическое определение вероятности.

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

Свойства геометрического определения вероятности полностью аналогичны свойствам классического определения вероятности.

Замечание 1.1.13. Т.к. мера точки равна нулю, то вероятность попадания в нее также равна нулю, хотя попасть в точку мы можем.

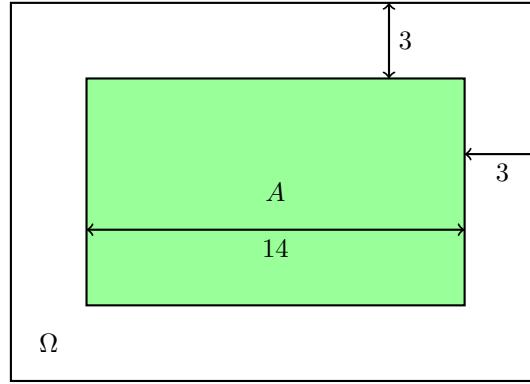
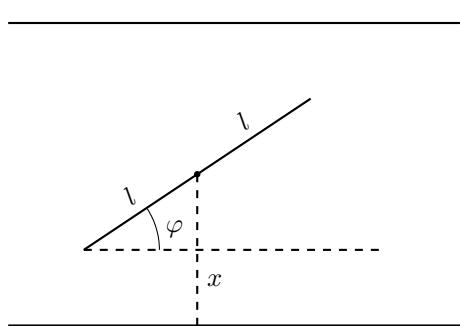


Рис. 1.1.14: Иллюстрация к 1.1.15

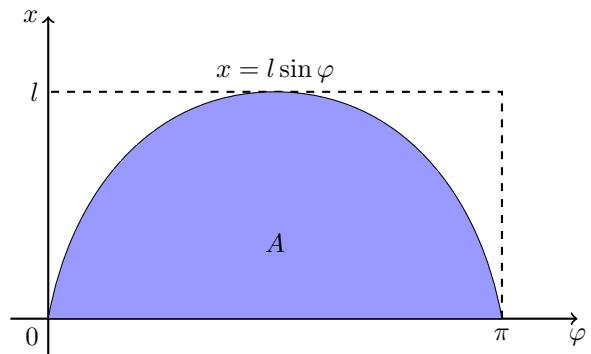
Пример 1.1.15. Монета диаметром 6 см наудачу бросается на пол, вымощенный квадратной плиткой со стороной 20 см. Какова вероятность того, что монета окажется целиком на одной плитке?

Решение: положение монеты определяется положением ее центра. Сделаем рисунок (рис. 1.1.14), из которого видно, что монета полностью окажется на одной из плиток, если ее центра окажется во внутреннем (зеленом) квадрате со стороной 14 см. Итого имеем

$$P(A) = \frac{14^2}{20^2} = \frac{196}{400} = 0.49$$



(a) Схематичное изображение иголки на ламинате



(b) Применение геометрического определения вероятности

Рис. 1.1.16: Иллюстрация к 1.1.17

Пример 1.1.17. Пол застелен ламинатом, на который бросается игла длиной равной длине одной доски ламината. Требуется найти вероятность того, что игла пересечет стык.

Решение: положение иглы определяется ее центром и углом поворота. Стоит отметить, что эти две величины независимы. Пусть игла имеет длину $2l$, сделаем рисунок (рис. 1.1.16a). Обозначим через $x \in [0; l]$ расстояние от середины иглы до ближайшего края доски, а через $\varphi \in [0; \pi]$ — угол между направлением иглы и доской. Таким образом $\Omega = [0; \pi] \times [0; l]$. Игла пересечет доску, если $x \leq l \sin \varphi$. Изобразим это (рис. 1.1.16b) и вычислим площадь области A .

$$S_A = \int_0^\pi l \sin \varphi = -l \cos \varphi \Big|_0^\varphi = 2l$$

Учитывая, что $S_\Omega = \pi l$, получаем, что искомая вероятность будет равна $P(A) = \frac{2l}{\pi l} = \frac{2}{\pi}$.

Статистика Максвелла–Больцмана

Поделим прямоугольную область, в которой «летает» r частиц, на n ячеек. Пусть частицы различимы, в одной ячейки может одновременно находиться несколько частиц и все размещения равновероятны. Итого вероятность попадания частицы в конкретную область будет равна $\frac{1}{n^r}$.

Статистика Бозе–Эйнштейна

Пусть теперь частицы неразличимы, но в одной ячейке все еще может одновременно находиться несколько частиц. Получаем следующую формулу для вероятности попадания частицы в конкретную область $\frac{1}{C_{n+r-1}^r}$.

Статистика Ферми–Дирака

Пусть частицы неразличимы и в одной ячейки может находиться равно одна частица. Значит формула для вероятности будет иметь вид $\frac{1}{C_n^r}$.

1.2. Лекция 23.09.12.

Будем строить аксиоматическое определение вероятности. Пусть у нас есть Ω — пространство элементарных исходов случайного эксперимента.

Def 1.2.1. Система \mathcal{F} подмножеств Ω называется σ -алгеброй событий, если выполнены следующие свойства

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \Rightarrow \bar{A} \in \mathcal{F}$ (замкнутость относительно дополнения)
3. $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ (замкнутость относительно счетного объединения)

Замечание 1.2.2. В 1.2.1 свойство 1° избыточно и может быть выведено из 2° и 3°.

$$A \in \mathcal{F} \xrightarrow{2^\circ} \bar{A} \in \mathcal{F} \xrightarrow{3^\circ} \underbrace{(A \cup \bar{A})}_{\Omega} \in \mathcal{F}$$

Свойства σ -алгебры событий

Lm 1.2.3.

$$\emptyset \in \mathcal{F}$$

□

$$\Omega \in \mathcal{F} \xrightarrow{2^\circ} \emptyset = \bar{\Omega} \in \mathcal{F}$$

Lm 1.2.4. σ -алгебра событий замкнута относительно счетного объединения.

$$A_1, A_2, \dots \in \mathcal{F} \implies \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$$

□ Пусть $A_1, A_2, \dots \in \mathcal{F}$, тогда

$$A_1, A_2, \dots \in \mathcal{F} \xrightarrow{2^\circ} \overline{A_1}, \overline{A_2}, \dots \in \mathcal{F} \xrightarrow{3^\circ} \overline{\bigcup_{i=1}^{\infty} A_i} \in \mathcal{F} \xrightarrow{2^\circ} \overline{\bigcup_{i=1}^{\infty} \overline{A_i}} \in \mathcal{F}$$

Применяя закон де Моргана (который работает не только в конечном, но и в счетном случае), получаем искомое.

$$\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$$

Lm 1.2.5.

$$A, B \in \mathcal{F} \implies A \setminus B \in \mathcal{F}$$

□

$$A, B \in \mathcal{F} \xrightarrow{2^\circ} \overline{B} \in \mathcal{F} \xrightarrow{3^\circ} A \setminus B = A\overline{B} \in \mathcal{F}$$

Пример 1.2.6. 1. $\mathcal{F} = \{\emptyset, \Omega\}$ — тривиальная.

$$2. \mathcal{F} = \{\emptyset, \Omega, A, \overline{A}\}$$

3. Пусть $\Omega = \mathbb{R}$, тогда \mathcal{F} это σ -алгебра Борелевских множеств на прямой.

Def 1.2.7. Пусть Ω — пространство элементарных исходов и \mathcal{F} его σ -алгебра. Вероятностью на $\langle \Omega, \mathcal{F} \rangle$ называется функция $P: \mathcal{F} \rightarrow \mathbb{R}$, обладающая следующими свойствами

1. Неотрицательность. $\forall A \in \mathcal{F} \mid P(A) \geq 0$

2. Счетная аддитивность. Если события A_1, A_2, \dots попарно несовместны, то $P\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$.

3. Нормированность. $P(\Omega) = 1$

Замечание 1.2.8. Таким образом, вероятность это нормированная мера, а свойства 1–3 называются аксиомами вероятности.

Def 1.2.9. Тройка $\langle \Omega, \mathcal{F}, P \rangle$ называется вероятностным пространством, где Ω это пространство элементарных исходов, \mathcal{F} его σ -алгебра, а P — нормированная мера.

Свойства вероятности

Lm 1.2.10.

$$P(\emptyset) = 0$$

□

$$1 = P(\Omega) = P(\emptyset + \Omega) = P(\emptyset) + P(\Omega) = P(\emptyset) + 1 \implies P(\emptyset) = 0$$

Lm 1.2.11 (Формула обратной вероятности).

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

□ Т.к. A и \overline{A} несовместны и $A + \overline{A} = \Omega$, то

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

Lm 1.2.12.

$$0 \leq P(A) \leq 1$$

□ $P(A) \geq 0$ по первой аксиоме вероятности. Далее по 1.2.11 имеем, что $P(A) = 1 - P(\bar{A})$. Т.к. $P(\bar{A}) \geq 0$, то $P(A) \leq 1$. ■

Аксиома непрерывности

Пусть имеется убывающая цепочка событий $A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$ и $\bigcap_{i=1}^{\infty} A_i = \emptyset$. Тогда $P(A_n) \rightarrow 0$ при $n \rightarrow \infty$.

Замечание 1.2.13. Смысл аксиомы непрерывности заключается в том, что при непрерывном изменении области $A \subset \Omega \subset \mathbb{R}^n$ соответствующая вероятность $P(A)$ также должна изменяться непрерывно.

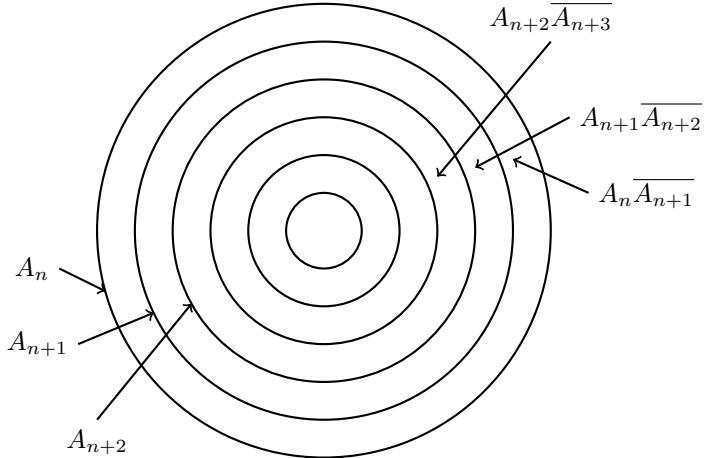


Рис. 1.2.14: Иллюстрация к 1.2.15

Теорема 1.2.15. Аксиома непрерывности следует из аксиомы счетной аддитивности.

□ Изобразим события A_1, A_2, \dots графически (рис. 1.2.14). Выразим A_n согласно полученному рисунку.

$$A_n = \underbrace{\sum_{i=n}^{\infty} A_i \overline{A_{i+1}}}_{\text{кольца}} + \underbrace{\bigcap_{i=n}^{\infty} A_i}_{\text{«сердцевина»}} \quad (1)$$

Далее заметим, что

$$\bigcap_{i=n}^{\infty} A_i = A_n \cap \bigcap_{i=n+1}^{\infty} A_i = \bigcap_{i=1}^n A_i \cap \bigcap_{i=n+1}^{\infty} A_i = \bigcap_{i=1}^{\infty} A_i \stackrel{\text{усл}}{=} \emptyset \quad (2)$$

Подставим (2) в (1) и получим, что

$$\begin{aligned} A_n &= \sum_{i=n}^{\infty} A_i \overline{A_{i+1}} \\ A_1 &= \sum_{i=1}^{\infty} A_i \overline{A_{i+1}} = \sum_{i=1}^n A_i \overline{A_{i+1}} + \underbrace{\sum_{i=n}^{\infty} A_i \overline{A_{i+1}}}_{A_n} \end{aligned} \quad (3)$$

Далее воспользуемся аксиомой счетной аддитивности.

$$P(A_1) = \sum_{i=1}^n P(A_i \overline{A_{i+1}}) + P(A_n) = \underbrace{\sum_{i=1}^{\infty} P(A_i \overline{A_{i+1}})}_S \quad (4)$$

Получаем числовой ряд S , который будет сходиться, т.е. его сумма равна $P(A_1)$, которая принадлежит отрезку $[0; 1]$. Тогда $P(A_n)$ это «хвост» ряда. Т.к. ряд сходится, то его хвост стремится к нулю при $n \rightarrow \infty$. Итого $P(A_n) \rightarrow 0$ при $n \rightarrow \infty$. ■

Замечание 1.2.16. Можно доказать, что счетная аддитивность следует из конечной аддитивности и аксиомы непрерывности.

Свойства операций сложения и умножения

Свойство дистрибутивности $A(B+C) = AB + AC$. Также если события A и B несовместны, то $P(A+B) = P(A) + P(B)$ (это частный случай аксиомы 2°).

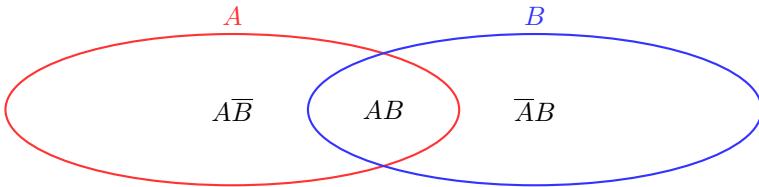


Рис. 1.2.17: Иллюстрация к 1.2.18

Lm 1.2.18. В общем случае (если про совместность событий A и B ничего не известно) имеем формулу

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

□ Т.к. $A + B = A\bar{B} + AB + \bar{A}B$ (рис. 1.2.17), то по второй аксиоме

$$\begin{aligned} P(A + B) &= P(A\bar{B}) + P(AB) + P(\bar{A}B) \\ &= \underbrace{(P(A\bar{B}) + P(AB))}_{P(A)} + \underbrace{(P(\bar{A}B) + P(AB))}_{P(B)} - P(AB) \\ &= P(A) + P(B) - P(AB) \end{aligned}$$

Пример 1.2.19. Из колоды в 36 карт достали одну карту. Какова вероятность того, что это будет дама или пика?

Решение: пусть событие D — «выпала дама», а событие S — «выпала пика». Обозначив A искомое событие, имеем

$$P(A) = P(D) + P(S) - P(DS) = \frac{4}{36} + \frac{9}{36} - \frac{1}{36} = \frac{12}{36} = \frac{1}{3}$$

При $n = 3$ формула усложняется.

$$P(A_1 + A_2 + A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1A_2) - P(A_1A_3) - P(A_2A_3) + P(A_1A_2A_3)$$

В общем же случае получаем следующую формулу

$$P(A_1 + A_2 + \dots + A_n) = \sum_{i=1}^n P(A_i) - \sum_{i < j} P(A_iA_j) + \sum_{i < j < k} P(A_iA_jA_k) - \dots + (-1)^{n-1}P(A_1 \dots A_n)$$

Пример 1.2.20. В n конвертов раскладываются n писем. Какова вероятность того, что хотя бы одно письмо попадет в свой конверт? Куда стремится эта вероятность при $n \rightarrow \infty$.

Решение: пусть A_i — i -тое письмо в своем конверте, а событие A — хотя бы одно письмо в своем конверте. Тогда $A = A_1 + \dots + A_n$. Применим общую формулу для вероятности суммы. Сначала вычислим вероятности пар, троек и т.д.

$$P(A_i) = \frac{1}{n} \quad P(A_iA_j) = \frac{1}{A_n^2} \quad P(A_iA_jA_k) = \frac{1}{A_n^3} \quad P(A_1 \dots A_n) = \frac{1}{A_n^n} = \frac{1}{n!}$$

Далее вычислим число пар, троек и т.д.

$$P(A_i) \rightarrow n = C_n^1 \quad P(A_iA_j) \rightarrow C_n^2 \quad P(A_iA_jA_k) \rightarrow C_n^3$$

Подставим это в формулу.

$$P(A) = n \cdot \frac{1}{n} - C_n^2 \cdot \frac{1}{A_n^2} + C_n^3 \cdot \frac{1}{A_n^3} - \dots + (-1)^{n-1} \frac{1}{n!} = 1 - \frac{1}{2!} + \frac{1}{3!} - \dots + (-1)^{n-1} \frac{1}{n!}$$

С помощью разложения в ряд Тейлора функции e^{-1} можно показать, что полученная сумма будет примерно равна $1 - e^{-1} \approx 0.63$.

Независимые события

Под независимыми событиями логично понимать события, не связанные причинно-следственной связью, т.е. факт наступления одного события не влияет на оценку вероятности второго события.

Рассмотрим это на примере классической вероятности. Пусть дано Ω — пространство элементарных исходов, и два события $|A| = m_1$ и $|B| = m_2$. Проведем пару независимых испытаний, тогда получим пространство элементарных исходов $\Omega \times \Omega$, $|\Omega \times \Omega| = n^2$ и $|AB| = m_1m_2$. Значит

$$P(AB) = \frac{m_1m_2}{n^2} = \frac{m_1}{n} \cdot \frac{m_2}{n} = P(A)P(B)$$

Def 1.2.21. События A и B называются независимыми, если $P(AB) = P(A)P(B)$.

Lm 1.2.22. Если A и B независимы, то A и \bar{B} , \bar{A} и B , \bar{A} и \bar{B} также независимы.

□ Т.к. $A = A(B + \bar{B}) = AB + A\bar{B}$, то

$$\begin{aligned} P(A) &= P(AB) + P(A\bar{B}) \implies \\ P(\bar{A}\bar{B}) &= P(A) - P(AB) \\ &= P(A) - P(A)P(B) \\ &= P(A)(1 - P(B)) \\ &= P(A)P(\bar{B}) \end{aligned}$$

Таким образом события A и \bar{B} независимы. Остальные свойства доказываются аналогично. ■

Def 1.2.23. События A_1, \dots, A_n называются независимыми в совокупности, если для любого набора номеров вероятность произведения событий с этими номерами равняется произведению отдельных вероятностей.

Замечание 1.2.24. Из независимости в совокупности при $k = 2$ следует попарная независимость. Обратное в общем случае неверно. Независимость в совокупности это более сильное свойство.

Пример 1.2.25. Три грани правильного тетраэдра раскрашены в красный, синий, зеленый цвета, а четвертая грань — во все эти три цвета. Пусть A — «выпала грань с красным цветом», B — «выпала грань с синим цветом» и C — «выпала грань с зеленым цветом».

$$\begin{aligned} P(A) &= P(B) = P(C) = \frac{2}{4} = \frac{1}{2} \\ P(AB) &= P(AC) = P(BC) = \frac{1}{4} \end{aligned}$$

Т.к. $P(AB) = P(A)P(B)$, то A и B независимы. Аналогично попарно независимы A и C , B и C , т.е. все события попарно независимы.

$$P(ABC) = \frac{1}{4} \neq P(A)P(B)P(C) = \frac{1}{8}$$

Т.е. события A , B и C не являются независимыми в совокупности.

Замечание 1.2.26. В дальнейшем под «независимыми событиями» будем подразумевать независимые в совокупности события.

Пример 1.2.27. Какова вероятность того, что при четырех бросках кости выпадет хотя бы одна шестерка.

Решение: пусть событие A_i — на i -том броске ($1 \leq i \leq 4$) выпала шестерка. Обозначим B событие «выпала хотя бы одна шестерка», тогда $B = A_1 + A_2 + A_3 + A_4$. Получаем, что \bar{B} это событие «не выпала ни одна шестерка», тогда $\bar{B} = \bar{A}_1 \cdot \bar{A}_2 \cdot \bar{A}_3 \cdot \bar{A}_4$. Т.к. броски кубиков независимые, то

$$P(\bar{B}) = P(\bar{A}_1)P(\bar{A}_2)P(\bar{A}_3)P(\bar{A}_4) = \left(\frac{5}{6}\right)^4 \approx 0.48$$

Значит, $P(B) = 1 - P(\bar{B}) \approx 0.52$.

1.3. Лекция 23.09.19.

Если два события не являются независимыми и известно, что одно из них произошло, то на основе этой информации мы можем уточнить вероятность наступления другого. Вероятность данного в таком случае называется условной. Условная вероятность $P(A|B)$ это вероятность события A вычисленная в предположении, что событие B уже произошло.

Замечание 1.3.1. Существует другой обозначение $P_B(A)$, но мы не будем его использовать.

Пример 1.3.2. Подбросили кубик. Известно, что выпало больше трех очков. Какова вероятность того, что выпало четное число очков?

Решение: пусть A — «выпало четное число очков», а B — «выпало более трех очков». Тогда

$$P(A) = \frac{3}{6} = \frac{1}{2} \quad P(B) = \frac{3}{6} = \frac{1}{2}$$

Тогда получаем, что

$$P(A|B) = \frac{2}{3} = \frac{\frac{2}{6}}{\frac{3}{6}} = \frac{P(AB)}{P(B)}$$

Такое же соотношение можно получить с помощью геометрической вероятности (рис. 1.3.3).

$$\mathbb{P}(A|B) = \frac{S_{AB}}{S_B} = \frac{\frac{S_{AB}}{S_\Omega}}{\frac{S_B}{S_\Omega}} = \frac{\mathbb{P}(AB)}{\mathbb{P}(B)}$$

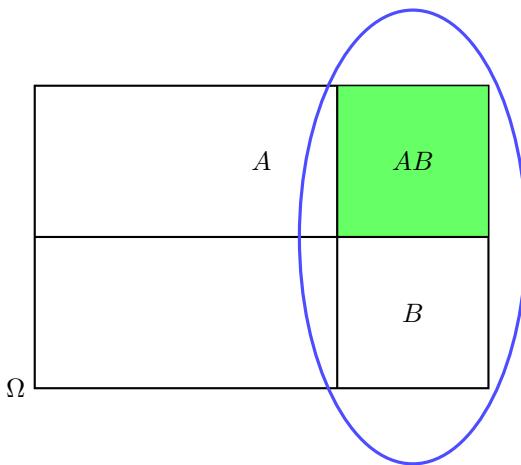


Рис. 1.3.3: Геометрический подход к условной вероятности

Замечание 1.3.4. Вывести эту формулу из аксиом вероятности нельзя, поэтому она берется в качестве определения.

Def 1.3.5. Условной вероятностью события A при условии, что имело место событие B называется величина

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(AB)}{\mathbb{P}(B)}$$

Формула произведения вероятностей

Из определения условной вероятности (1.3.5) следует, что

$$\mathbb{P}(AB) = \mathbb{P}(B)\mathbb{P}(A|B) = \mathbb{P}(A)\mathbb{P}(B|A)$$

Или, в общем случае, имеем

Теорема 1.3.6. (Формула произведения вероятностей)

$$\mathbb{P}(A_1 \dots A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1A_2) \dots \mathbb{P}(A_n|A_1 \dots A_{n-1})$$

□ Мат. индукция.

База: $n = 2$.

Этот случай уже рассмотрен ранее, он следует из определения условной вероятности.

Переход: $n - 1 \rightarrow n$

$$\begin{aligned} \mathbb{P}\left(\underbrace{A_1 \dots A_{n-1}}_{\mathbb{P}(A_1 \dots A_{n-1})} A_n\right) &= \mathbb{P}(A_1 \dots A_{n-1})\mathbb{P}(A_n|A_1 \dots A_{n-1}) \\ &= \underbrace{\mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1A_2) \dots \mathbb{P}(A_{n-1}|A_1 \dots A_{n-2})}_{\mathbb{P}(A_1 \dots A_{n-1})} \mathbb{P}(A_n|A_1 \dots A_{n-1}) \end{aligned}$$

Замечание 1.3.7. Условная вероятность определена при $\mathbb{P}(B) \neq 0$, поэтому формула умножения (1.3.6) верна при $\mathbb{P}(A_1 \dots A_{n-1}) \neq 0$.

Lm 1.3.8. Покажем, что равносильны два определения независимости: $\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B)$ и $\mathbb{P}(A|B) = \mathbb{P}(A)$

□

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(AB)}{\mathbb{P}(B)} = \mathbb{P}(A) \iff \mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B)$$

Пример 1.3.9. В коробке 3 красных карандаша и 2 синих. Вынули три карандаша. Найти вероятность того, что первые два красные, а третий — синий.

Решение: пусть A_1 — «первый красный», A_2 — «второй красный» и A_3 — «третий синий».

$$\mathbb{P}(A_1A_2A_3) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1A_2) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} = \frac{1}{5}$$

Полная группа событий

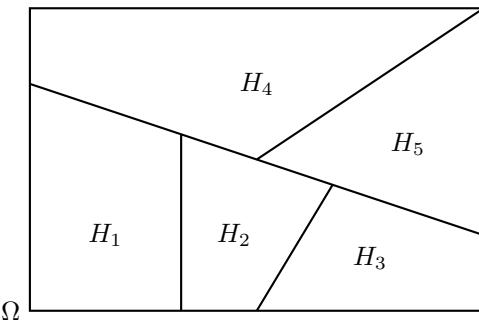


Рис. 1.3.10: Иллюстрация полной группы событий

Def 1.3.11. События H_1, \dots, H_n образуют полную группу событий, если они попарно несовместны и содержат все элементарные исходы.

$$\begin{cases} \forall i \neq j \mid H_i H_j = \emptyset \\ H_1 + \dots + H_n = \Omega \end{cases}$$

Следствие 1.3.12.

$$P(H_1) + \dots + P(H_n) = 1$$

□ Т.к. события H_1, \dots, H_n несовместны, то вероятность их суммы равна сумме их вероятностей, и при этом она равна единице, т.к. эти события содержат все элементарные исходы. ■

Замечание 1.3.13. События H_i еще называются гипотезами.

Формула полной вероятности

Теорема 1.3.14. Пусть H_1, \dots, H_n, \dots — полная группа событий. Тогда

$$P(A) = \sum_{i=1}^{\infty} P(H_i) P(A \mid H_i)$$

□

$$P(A) = P(\Omega A) = P((H_1 + \dots + H_n + \dots) A) = P(H_1 A + \dots + H_n A + \dots)$$

Заметим, что т.к. события H_i из полной группы, то они несовместны, значит и полученные произведения $H_i A$ также несовместны. Применяя формулу произведения вероятностей имеем

$$P(A) = \sum_{i=1}^{\infty} P(H_i A) = \sum_{i=1}^{\infty} P(H_i) P(A \mid H_i)$$

Формула Байеса (формула проверки гипотез)

Теорема 1.3.15. Пусть H_1, \dots, H_n — полная группа событий. Известно, что событие A произошло. Тогда

$$P(H_k \mid A) = \frac{P(H_k) P(A \mid H_k)}{\sum_{i=1}^{\infty} P(H_i) P(A \mid H_i)}$$

□ Раскрываем условную вероятность по определению, после чего в числителе применяем формулу произведения вероятностей, а в знаменателе — формулу полной вероятности.

$$P(H_k \mid A) = \frac{P(H_k A)}{P(A)} = \frac{P(H_k) P(A \mid H_k)}{\sum_{i=1}^{\infty} P(H_i) P(A \mid H_i)}$$

Пример 1.3.16. В первой коробке 4 белых и 2 черных шара, а во второй — 1 белый и 2 черных. Из первой коробки во вторую переложили два шара. Затем из второй коробки достали шар. Какова вероятность того, что он оказался белым?

Решение: пусть H_1 — во вторую коробку переложили 2 белых шара, H_2 — переложили 2 черных шара, H_3 — переложили 2 шара разного цвета. A — из второй коробки достали белый шар, тогда

$$P(H_1) = \frac{4}{6} \cdot \frac{3}{5} = \frac{6}{15} \quad P(H_2) = \frac{2}{6} \cdot \frac{1}{5} = \frac{1}{15} \quad P(H_3) = \frac{4}{6} \cdot \frac{2}{5} + \frac{2}{6} \cdot \frac{4}{5} = \frac{8}{15}$$

$$P(A|H_1) = \frac{3}{5} \quad P(A|H_2) = \frac{1}{5} \quad P(A|H_3) = \frac{2}{5}$$

По формуле полной вероятности получаем

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + P(H_3)P(A|H_3) = \frac{6}{15} \cdot \frac{3}{5} + \frac{1}{15} \cdot \frac{1}{5} + \frac{8}{15} \cdot \frac{2}{5} = \frac{18+1+16}{75} = \frac{7}{15}$$

Пример 1.3.17. По статистике раком болен 1% населения. Тест дает правильный результат в 99% случаев. Анализ оказался положительным. Какова вероятность того, что человек болен?

Решение: пусть H_1 — человек болен, а H_2 — здоров. Событие A — анализ положительный. Тогда

$$\begin{aligned} P(H_1) &= 0.01 & P(H_2) &= 0.99 \\ P(A|H_1) &= 0.99 & P(A|H_2) &= 0.01 \end{aligned}$$

Применяем формулу Байеса.

$$P(H_1|A) = \frac{P(H_1)P(A|H_1)}{P(H_1)P(A|H_1) + P(H_2)P(A|H_2)} = \frac{0.01 \cdot 0.99}{0.01 \cdot 0.99 + 0.99 \cdot 0.01} = \frac{1}{2}$$

Пример 1.3.18. Пусть после ситуации, описанной в [1.3.17](#), провели второй **независимый** анализ, и он оказался положительным. Какова вероятность того, что человек болен?

Решение: используя обозначения из предыдущего примера получаем

$$\begin{aligned} P(H_1) &= 0.01 & P(H_2) &= 0.99 \\ P(AA|H_1) &= 0.99 \cdot 0.99 & P(AA|H_2) &= 0.01 \cdot 0.01 \end{aligned}$$

Применяем формулу Байеса.

$$P(H_1|AA) = \frac{P(H_1)P(AA|H_1)}{P(H_1)P(AA|H_1) + P(H_2)P(AA|H_2)} = \frac{0.01 \cdot 0.99^2}{0.01 \cdot 0.99^2 + 0.99 \cdot 0.01^2} = 0.99$$

Статистические рассуждения

Пусть проверили 10 000 человек. По статистике из них 100 больных и 9 900 здоровых. У 99 человек из 100 больных анализ окажется положительным, а у одного — отрицательным. У 99 из 9900 здоровых людей теста будет положительным, а у остальных — отрицательным. Таким образом мы видим, что даже у человека положительный тест, то вероятность того, что он болен лишь 50%.

Пример 1.3.19. В студии три двери. За одной из них приз. Ведущий предлагает игроку угадывать дверь с призом, после чего он открывает одну из оставшихся дверей и показывает, что за ней приза нет. Затем он предлагает игроку поменять свой выбор. Следует ли игроку соглашаться?

Решение: пользоваться формулой Байеса здесь нельзя, т.к. выбор ведущего (событие A в формуле) это не случайное событие. Можно воспользоваться примерными статистическими рассуждениями. Пусть игрок сыграл в эту игру 300 раз. Тогда он 100 раз изначально выбрал верную дверь и 200 раз — неверную. Если он после вопроса ведущего поменяет свой выбор, то окажется прав в 200 случаях из 300. Таким образом, если игрок поменяет свой выбор, то вероятность его победы будет равна $\frac{2}{3}$.

1.4. Лекция 23.09.26.

Последовательности независимых испытаний

Def 1.4.1. Схемой Бернулли называется серия одинаковых независимых испытаний, каждое из которых имеет лишь два исхода: произошло интересующее нас событие (успех) или не произошло (неудача). Обозначения:

1. p — вероятность успеха при одном испытании.
2. $q = 1 - p$ — вероятность неудачи при одном испытании.
3. n — число независимых испытаний.
4. Y_n — число успехов при n испытаниях.
5. N_n — число неудач при n испытаниях.

6. Для краткости обозначим $p(Y_n = k) = p_n(k)$.

Формула Бернулли

Теорема 1.4.2. Вероятность того, что при n испытаниях произойдет ровно k успехов будет равна

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

□ Пусть $A = \{Y_n = k\}$. Рассмотрим один из элементарных исходов, благоприятствующих событию A .

$$A_1 = \underbrace{Y \dots Y}_k \underbrace{N \dots N}_{n-k} \quad P(Y) = p, P(N) = q$$

Т.к. испытания независимы, то

$$P(A_1) = p^k q^{n-k}$$

Остальные элементарные исходы (благоприятствующие A) отличаются от данного лишь расстановкой k успехов по n испытаниям, а вероятности их те же самые. Всего таких способов расставить будет C_n^k . В результате получаем искомую формулу. ■

Пример 1.4.3. Вероятность попадания стрелка при одном выстреле равна 0.8. Найти вероятность того, что из пяти выстрелов три будут точными.

Решение: Обозначим $n = 5$, $k = 3$, $p = 0.8$, $q = 0.2$. Подставим это в формулу и получим

$$P_5(3) = C_5^3 \cdot 0.8^3 \cdot 0.2^2 = 0.2048$$

Наиболее вероятное число успехов

Выясним, при каком значении k , вероятность предшествующего числа успехов $k - 1$ будет не более, чем вероятность k успехов. Формально получаем

$$\begin{aligned} P_n(k-1) &\leq P_n(k) \\ C_n^{k-1} \cdot p^{k-1} \cdot q^{n-k+1} &\leq C_n^k \cdot p^k \cdot q^n \\ \frac{n!}{(k-1)!(n-k+1)!} \cdot q &\leq \frac{n!}{k!(n-k)!} \cdot p \\ k \cdot (1-p) &\leq (n-k+1) \cdot p \\ k - kp &\leq np - kp + p \\ k &\leq np + p \end{aligned}$$

Т.к. k целое, то получаем, что $np + p - 1 \leq k \leq np + p$. Рассмотрим 3 ситуации.

Случай I np целое

Тогда $np + p$ — не целое и $k = np$ это наиболее вероятное число успехов.

Случай II $np + p$ не целое

Тогда $k = [np + p]$. Это обобщение случая I.

Случай III $np + p$ целое

В этом случае получаем два наиболее вероятных числа успехов $k = np + p$ и $k = np + p - 1$.

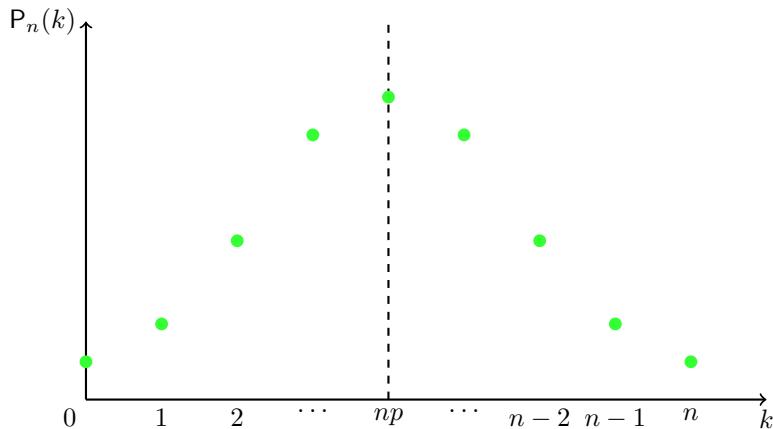


Рис. 1.4.4: Геометрическая иллюстрация наиболее вероятного числа успехов

Если рассмотреть рисунок рис. 1.4.4 и устремить $k \rightarrow \infty$, то через полученные точки можно будет провести плавную кривую. Полученный график (рис. 1.4.5) называется нормальным распределением. Такое распределение имеет место быть, если вероятности успеха и неудачи примерно равны. Если же какая-либо из этих вероятностей много больше другой, то «горб» будет сильно смещен к одной из границ, при этом одна из его сторон будет крутой, а другая — пологой. Такое распределение называется показательным (об этом мы поговорим позже).

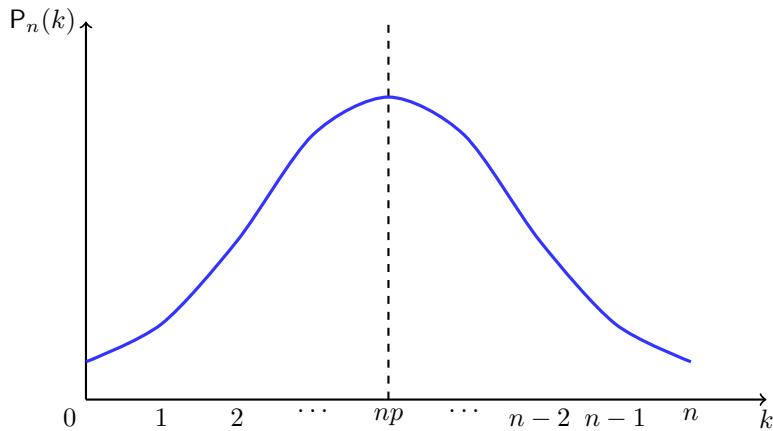


Рис. 1.4.5: Нормальное распределение

Пределевые теоремы в схеме Бернулли

Если требуется найти вероятность точного числа успехов, то применяем следующую теорему:

Теорема 1.4.6. (Локальная теорема Муавра-Лапласа)

$$P_n(k) \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{npq}} \varphi(x) \quad \begin{cases} \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \\ x = \frac{k-np}{\sqrt{npq}} \end{cases}$$

Замечание 1.4.7. Свойства $\varphi(x)$ (функция Гаусса)

1. $\varphi(-x) = \varphi(x)$ — функция четная.
2. $x > 5 \implies \varphi(x) \approx 0$ — функция быстро убывает.

Если требуется найти вероятность того, что число успехов находится в данном диапазоне, то применяем другую теорему:

Теорема 1.4.8. (Интегральная теорема Муавра-Лапласа)

$$P_n(k_1 \leq k \leq k_2) \xrightarrow{n \rightarrow \infty} \Phi(x_2) - \Phi(x_1) \quad \begin{cases} \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz \\ x_1 = \frac{k_1 - np}{\sqrt{npq}} \\ x_2 = \frac{k_2 - np}{\sqrt{npq}} \end{cases}$$

Замечание 1.4.9. Свойства $\Phi(x)$ (функция Лапласа)

1. $\Phi(-x) = -\Phi(x)$ — функция нечетная.
2. $x > 5 \implies \Phi(x) \approx 0.5$

Замечание 1.4.10. В различных источниках под функцией Лапласа подразумевают несколько иные функции, чаще всего функцию

$$F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-z^2/2} dz$$

Это функция стандартного нормального распределения.

Замечание 1.4.11. Интеграл Пуассона

$$\int_{-\infty}^{\infty} e^{-z^2/2} dz = \sqrt{2\pi}$$

Используя его можно получить, что

$$F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-z^2/2} dz + \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz$$

Замечание 1.4.12. Эти формулы (1.4.6 и 1.4.8) обычно применяются при $n \geq 100$ и когда p и q отличаются друг от друга не слишком сильно. Если же p и q отличаются друг от друга очень сильно, то сходимость все равно будет, но при очень больших n , которые не всегда можно получить на практике. Для таких случаев применяется формула Пуассона (формула редких событий), которая будет рассмотрена на следующей лекции.

Пример 1.4.13. Вероятность попадания стрелка в цель 0.8. Сделано 400 выстрелов, найти вероятность того, что произошло ровно 330 попаданий.

Решение: Обозначим $n = 400$, $p = 0.8$, $q = 0.2$, $k = 330$. По локальной формуле Лапласа

$$x = \frac{k - np}{\sqrt{npq}} = \frac{330 - 400 \cdot 0.8}{\sqrt{400 \cdot 0.8 \cdot 0.2}} = \frac{330 - 320}{\sqrt{64}} = 1.25$$

$$P_{400}(330) \approx \frac{1}{\sqrt{npq}} \varphi(x) = \frac{1}{8} \varphi(1.25) = 0.125 \cdot 0.1826 \approx 0.0228$$

Пример 1.4.14. Вероятность попадания стрелка в цель 0.8. Сделано 400 выстрелов, найти вероятность того, что произошло от 312 до 336 попаданий.

Решение: Обозначим $n = 400$, $p = 0.8$, $q = 0.2$, $k_1 = 312$, $k_2 = 336$. По интегральной формуле Лапласа

$$x_1 = \frac{k_1 - np}{\sqrt{npq}} = \frac{312 - 400 \cdot 0.8}{\sqrt{400 \cdot 0.8 \cdot 0.2}} = \frac{312 - 320}{\sqrt{64}} = -1 \quad x_2 = \frac{k_2 - np}{\sqrt{npq}} = \frac{336 - 400 \cdot 0.8}{\sqrt{400 \cdot 0.8 \cdot 0.2}} = \frac{336 - 320}{\sqrt{64}} = 2$$

$$P_{400}(312 \leq k \leq 336) \approx \Phi(2) - \Phi(-1) = \Phi(2) + \Phi(1) = 0.4772 + 0.3413 \approx 0.8185$$

Вернемся к статистическому определению вероятности. Пусть проводится n независимых испытаний, событие A появилось n_A раз. Относительная частота равна $\frac{n_A}{n}$, тогда статистическое определение вероятности утверждает, что $\frac{n_A}{n} \approx P(A) = p$.

Вероятность отклонения относительной частоты от вероятности события

Пусть p — вероятность события A , а $\frac{n_A}{n}$ — относительная частота. По интегральной формуле Лапласа

$$P\left(\left|\frac{n_A}{n} - p\right| < \varepsilon\right) = P(-n\varepsilon < n_A - np < n\varepsilon) = P(np - n\varepsilon < n_A < np + n\varepsilon)$$

$$x_1 = \frac{np - n\varepsilon - np}{\sqrt{npq}} = -\frac{n\varepsilon}{\sqrt{npq}} \quad x_2 = \frac{np + n\varepsilon - np}{\sqrt{npq}}$$

$$P\left(\left|\frac{n_A}{n} - p\right| < \varepsilon\right) \approx \Phi\left(\frac{n\varepsilon}{\sqrt{npq}}\right) - \Phi\left(-\frac{n\varepsilon}{\sqrt{npq}}\right) = \Phi\left(\frac{n\varepsilon}{\sqrt{npq}}\right) + \Phi\left(\frac{n\varepsilon}{\sqrt{npq}}\right) = 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right)$$

Итак, получили следующую формулу

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

Закон больших чисел Бернулли

Рассмотрим выведенную формулу. При $n \rightarrow \infty$ аргумент функции Лапласа будет стремиться к ∞ , а значит по 2° свойству значение функции будет стремиться к $\frac{1}{2}$. Значит

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{n_A}{n} - p\right| < \varepsilon\right) = 1$$

Это равенство и называется законом больших чисел Бернулли. Такая сходимость называется сходимостью по вероятности.

Пример 1.4.15. Теща Кисы Воробьянинова с вероятностью 0.9 зашила бриллианты в одном из 12 стульев, а с вероятностью 0.1 подшутила над ним. В первых одиннадцати стульях бриллиантов не оказалось. Какова вероятность того, что они найдутся в двенадцатом стуле?

Решение: A_i — бриллианты в i -том стуле, $P(A_i) = \frac{3}{40}$. Получаем

$$P(A_{12} | \overline{A_1} \dots \overline{A_{11}}) = \frac{P(A_{12} \overline{A_1} \dots \overline{A_{11}})}{P(\overline{A_1} \dots \overline{A_{11}})}$$

$$P(\overline{A_1} \dots \overline{A_{11}}) = 1 - P(A_1 + \dots + A_{11}) = 1 - (P(A_1) + \dots + P(A_{11})) = \frac{7}{40}$$

$$P(A_{12} \overline{A_1} \dots \overline{A_{11}}) = P(A_{12}) = \frac{3}{40}$$

$$P(A_{12} | \overline{A_1} \dots \overline{A_{11}}) = \frac{3}{7}$$

Пример 1.4.16. Для оценки доли p курящих людей берется выборка объема n . Далее делается оценка $p^* = \frac{n_k}{n}$. Какой должен быть объем n , чтобы с вероятностью $\gamma = 0.95$ данная оценка отличалась от истинного значения не более чем на $\varepsilon = 0.01$.

Решение: используем формулу отклонения относительной частоты от вероятности

$$\begin{aligned} \mathbb{P}(|p^* - p| \leq \varepsilon) &\approx 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right) \geq 0.95 \\ \Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right) &\geq 0.475 \\ \frac{\sqrt{n}\varepsilon}{\sqrt{pq}} &\geq 1.96 \\ \sqrt{n} &\geq \frac{1.96 \cdot \sqrt{pq}}{\varepsilon} \\ n &\geq \frac{1.96^2 \cdot pq}{0.01^2} \\ n &\geq 196^2 \cdot p(1-p) \end{aligned}$$

Функция $p(1-p)$ достигает наибольшего значения при $p = 0.5$, значит получаем $n \geq 9604$.

1.5. Лекция 23.10.03.

Схемы испытаний и соответствующие распределения

Обозначим n — число испытаний, p — вероятность успеха при одном испытании, $q = 1 - p$ — вероятность неудачи при одном испытании.

Схема I. Схема Бернулли

Пусть Y_n — число успехов при n испытаниях. По формуле Бернулли

$$\mathbb{P}_n(Y_n = k) = C_n^k p^k q^{n-k} \quad 0 \leq k \leq n$$

Def 1.5.1. Соответствие $k \rightarrow C_n^k p^k q^{n-k}$, $0 \leq k \leq n$ называется биномиальным распределением вероятностей и обозначается $B_{n,p}$ или $B(n,p)$.

Схема II. Схема до первого успешного испытания

Пусть проводится бесконечная серия независимых испытаний, которая заканчивается после первого успешного испытания под номером τ .

Lm 1.5.2.

$$\mathbb{P}(\tau = k) = q^{k-1}p \quad k = 0, 1, \dots$$

□

$$\mathbb{P}(\tau = k) = \mathbb{P}\left(\underbrace{N \dots N}_{k-1} Y\right) = q^{k-1}p$$

■

Def 1.5.3. Соответствие $k \rightarrow q^{k-1}p$, $k \in \mathbb{N}$ называется геометрическим распределением и обозначается G_p .

Геометрическое распределение обладает свойством «не старения» или «отсутствия последействия».

Теорема 1.5.4.

$$\mathbb{P}(\tau = k) = q^{k-1}p, k \in \mathbb{N} \implies \forall n, k \geq 0 \mid \mathbb{P}(\tau > n+k \mid \tau > n) = \mathbb{P}(\tau > k)$$

□ Заметим, что $\mathbb{P}(\tau > m) = q^m$, т.к. нас интересуют ситуации, в которых каким минимум первые m испытаний были неудачными. Применим определение и получим

$$\mathbb{P}(\tau > n+k \mid \tau > n) = \frac{\mathbb{P}(\tau > n+k, \tau > n)}{\mathbb{P}(\tau > n)} = \frac{q^{n+k}}{q^n} = q^k = \mathbb{P}(\tau > k)$$

■

Схема III. Испытания с несколькими исходами (полиномиальная схема)

Пусть при n независимых испытаниях могут произойти m несовместных исходов, вероятности которых равны p_i при одном отдельном испытании.

Теорема 1.5.5. Вероятность того, что при n испытаниях первый исход произойдет n_1 раз, второй — n_2 раз, …, m -ый — n_m раз равна

$$P_n(n_1, \dots, n_m) = \frac{n!}{n_1! \dots n_m!} \cdot p_1^{n_1} \cdot \dots \cdot p_m^{n_m}$$

□ Рассмотрим один элементарный исход благоприятный данному событию.

$$A_1 = \underbrace{1 \dots 1}_{n_1} \underbrace{2 \dots 2}_{n_2} \dots \underbrace{m \dots m}_{n_m} = p_1^{n_1} \cdot \dots \cdot p_m^{n_m}$$

Остальные благоприятные исходы имеют те же вероятности и отличаются лишь расстановкой i -ых исходов по n местам. Число таких исходов равно

$$C_n^{n_1} \cdot C_{n-n_1}^{n_2} \dots \cdot C_{n_m}^{n_m} = \underbrace{1 \dots 1}_{n_1} \underbrace{2 \dots 2}_{n_2} \dots \underbrace{m \dots m}_{n_m}$$

Это формула перестановок с повторениями. В итоге получаем требуемую формулу. ■

Замечание 1.5.6. При $n = 2$ доказанная формула превращается в уже рассмотренную формулу Бернулли.

Пример 1.5.7. Два одинаковых по силе шахматиста играют матч из шести партий. Вероятность ничьи при одной партии 0.5. Какова вероятность того, что второй игрок выиграет две партии, а еще три сведет к ничьей.

Решение: Обозначим $n = 6$, $m = 3$ (победа, поражение, ничья), $p_3 = 0.5$, $p_1 = p_2 = 0.25$. По формуле получаем

$$P_6(1, 2, 3) = \frac{6!}{1! \cdot 2! \cdot 3!} \cdot \left(\frac{1}{4}\right)^1 \cdot \left(\frac{1}{4}\right)^2 \cdot \left(\frac{1}{2}\right)^3 = \frac{15}{128} \approx 0.12$$

Схема IV. «Урновая» схема

В урне N шаров, из которых K — белые, $N - K$ — черные. Выбираем n шаров без учета порядка, обозначим k — число белых шаров в выборке.

Схема IV. (a) С возвратом

Пусть после вытаскивания шара мы сразу же возвращаем его в урну. Вероятность выбрать белый шар не меняется и равна $p = \frac{K}{N}$. Таким образом получается обычная схема Бернулли.

$$P_n(k) = C_n^k \cdot \left(\frac{K}{N}\right)^k \cdot \left(1 - \frac{K}{N}\right)^{n-k}$$

Схема IV. (b) Без возврата

$$P_{N,K}(n, k) = \frac{C_K^k \cdot C_{N-K}^{n-k}}{C_N^n} \quad 0 \leq k \leq K$$

Def 1.5.8. Соответствие $k \rightarrow \frac{C_K^k \cdot C_{N-K}^{n-k}}{C_N^n}$, $0 \leq k \leq K$ называется гипергеометрическим распределением.

Интуитивно ясно, что если число шаров N велико, а n — нет, то извлечение без возврата белых шаров не сильно влияет на пропорцию шаров в урне и поэтому результат по схеме (b) будут стремиться к схеме (a), а гипергеометрическое распределение будет приближаться к биномиальному.

Lm 1.5.9.

$$C_K^k \sim \frac{K^k}{k!} \quad K \rightarrow \infty, k = const$$

□

$$\begin{aligned} C_K^k &= \frac{K!}{k!(K-k)!} \\ &= \frac{K \cdot (K-1) \cdot \dots \cdot (K-k+1)}{k!} \\ &= \frac{K \cdot (K-1) \cdot \dots \cdot (K-k+1)}{K^k} \cdot \frac{K^k}{k!} \\ &= \underbrace{\left(1 - \frac{1}{K}\right) \left(1 - \frac{2}{K}\right) \dots \left(1 - \frac{k-1}{K}\right)}_{\rightarrow 1} \cdot \frac{K^k}{k!} \sim \frac{K^k}{k!} \end{aligned}$$

Теорема 1.5.10. Если $N \rightarrow \infty$ и $K \rightarrow \infty$, таким образом что $\frac{K}{N} \rightarrow p \in (0; 1)$, n и $0 \leq k \leq n$ фиксированы, то

$$\mathbb{P}_{N,K}(n, k) = \frac{C_K^k \cdot C_{N-K}^{n-k}}{C_N^n} \xrightarrow{n \rightarrow \infty} C_n^k \cdot p^k \cdot (1-p)^{n-k}$$

□ К каждому сочетанию применим лемму 1.5.9.

$$\frac{C_K^k \cdot C_{N-K}^{n-k}}{C_N^n} \sim \frac{K^k}{k!} \cdot \frac{(N-K)^{n-k}}{(n-k)!} \cdot \frac{n!}{N^n} = C_k^n \cdot \frac{K^k \cdot (N-K)^{n-k}}{N^n} = C_k^n \cdot \frac{K^k}{N^k} \cdot \frac{(N-K)^{n-k}}{N^{n-k}} = C_k^n \left(\frac{K}{N}\right)^k \cdot \left(1 - \frac{K}{N}\right)^{n-k}$$

Используя обозначение $p = \frac{K}{N}$ получаем исходную формулу. ■

Схема V. Схема Пуассона

Вероятность успеха p_n зависит от числа испытаний n таким образом, что их произведение $np_n \approx \lambda = const$. Под λ можно понимать интенсивность появления редких событий в заданную единицу времени.

Теорема 1.5.11. (Формула Пуассона) Пусть $n \rightarrow \infty$, тогда $p_n \rightarrow 0$ так, что $n \cdot p_n \rightarrow \lambda > 0$. Тогда вероятность k успехов при n испытаниях будет равна

$$\mathbb{P}(Y_n = k) = C_k^n \cdot p_n^k \cdot (1 - p_n)^{n-k} \xrightarrow{n \rightarrow \infty} \frac{\lambda^k}{k!} e^{-\lambda}$$

□ Положим $\lambda_n = n \cdot p_n$, тогда $\lambda_n \rightarrow \lambda$, $p_n = \frac{\lambda_n}{n}$. Имеем

$$\mathbb{P}(Y_n = k) = C_k^n \cdot p_n^k \cdot (1 - p_n)^{n-k} = C_k^n \cdot \left(\frac{\lambda_n}{n}\right)^k \cdot \left(1 - \frac{\lambda_n}{n}\right)^{n-k}$$

Далее применим лемму 1.5.9.

$$\mathbb{P}(Y_n = k) \xrightarrow{n \rightarrow \infty} \frac{n^k}{k!} \cdot \frac{\lambda_n^k}{n^k} \cdot \left(1 - \frac{\lambda_n}{n}\right)^{n-k} = \frac{\lambda_n^k}{k!} \cdot \left(1 - \frac{\lambda_n}{n}\right)^n \cdot \underbrace{\left(1 - \frac{\lambda_n}{n}\right)^{-k}}_{\rightarrow 1}$$

Т.к. $\frac{\lambda_n}{n} = p_n \rightarrow 0$ при $n \rightarrow \infty$, а $k = const$, то вторая скобка стремится к единице. Первую же скобку в пределе можно раскрыть с помощью второго замечательного предела. Итого

$$\begin{aligned} \left(1 - \frac{\lambda_n}{n}\right)^n &= \left(1 - \frac{\lambda_n}{n}\right)^{\frac{-n}{\lambda_n} \cdot (-\lambda_n)} \xrightarrow{n \rightarrow \infty} e^{-\lambda_n} \\ \mathbb{P}(Y_n = k) &\xrightarrow{n \rightarrow \infty} \frac{\lambda_n^k}{k!} \cdot e^{-\lambda_n} \xrightarrow{n \rightarrow \infty} \frac{\lambda^k}{k!} \cdot e^{-\lambda} \end{aligned}$$

Теорема 1.5.12. (Оценка погрешности в формуле Пуассона) Пусть Y_n это число успехов при n испытаниях в схеме Бернулли с вероятностью успеха p . Обозначим $\lambda = np$ и $A \subset \{0, 1, \dots, n\}$ произвольное подмножество целых чисел. Тогда

$$\left| \mathbb{P}(Y_n = k) - \sum_{k \in A} \frac{\lambda^k}{k!} \cdot e^{-\lambda} \right| \leq \min(p, \lambda p)$$

Def 1.5.13. Соответствие $k \rightarrow \frac{\lambda^k}{k!} \cdot e^{-\lambda}$, $k \in \mathbb{N}_0$ называется распределением Пуассона с параметром $\lambda > 0$ и обозначается Π_λ .

Пример 1.5.14. Прибор состоит из 1000 элементов. Вероятность отказа каждого элемента 0.001. Какова вероятность отказа больше двух элементов?

Решение: Обозначим $n = 1000$, $p = 0.001$, $\lambda = np = 1$, $k > 2$. Применим формулу Пуассона.

$$\mathbb{P}(k > 2) = 1 - \mathbb{P}(k \leq 2) = 1 - \mathbb{P}(0) - \mathbb{P}(1) - \mathbb{P}(2) = 1 - \frac{1^0}{0!} \cdot e^{-1} - \frac{1^1}{1!} \cdot e^{-1} - \frac{1^2}{2!} \cdot e^{-1} = 1 - e^{-1} \cdot \left(1 + 1 + \frac{1}{2}\right) \approx 0.0803$$

Применим 1.5.12, чтобы оценить погрешность.

$$\varepsilon \leq \min(p, \lambda p) = 0.001$$

1.6. Лекция 23.10.10.

Случайные величины

Def 1.6.1. Пусть имеется вероятностное пространство $\langle \Omega, \mathcal{F}, P \rangle$, функция $\xi: \Omega \rightarrow \mathbb{R}$ называется \mathcal{F} -измеримой, если

$$\forall x \in \mathbb{R} \left| \left\{ \omega \in \Omega \mid \xi(\omega) < x \right\} \in \mathcal{F} \right.$$

Таким образом $\xi^{-1}((-\infty; x)) \in \mathcal{F}$.

Def 1.6.2. Случайной величиной, заданной на вероятностном пространстве $\langle \Omega, \mathcal{F}, P \rangle$ называется \mathcal{F} -измеримая функция $\xi: \Omega \rightarrow \mathbb{R}$, сопоставляющая каждому элементарному исходу некоторое вещественное число.

Замечание 1.6.3. Не все функции являются измеримыми. Например, рассмотрим бросок кости, $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{F} = \{\emptyset, \Omega, A = \{2, 4, 6\}, \bar{A} = \{1, 3, 5\}\}$. Если $\xi(x) = i$, то такая функция не будет \mathcal{F} -измеримой, потому что при $x = 4$ получается

$$\left\{ \omega \in \Omega \mid \xi(\omega) < x \right\} = \{1, 2, 3\} \notin \mathcal{F}$$

Если же определить $\xi(x)$ как $\xi(2) = \xi(4) = \xi(6) = 1$, $\xi(1) = \xi(3) = \xi(5) = 0$, то $\xi(x)$ будет \mathcal{F} -измерима.

Смысл измеримости

Если задана случайная величина ξ , то мы можем задать вероятность попадания случайной величины в интервал $(-\infty; x)$, получаем $P(\xi \in (-\infty; x)) = P\left(\omega \in \Omega \mid \xi(\omega) < x\right)$. С помощью операций объединения, пересечения и дополнения из этих интервалов мы можем получить любые другие интервалы (включая точки) и также приписать им вероятности. Далее согласно теореме Каратеодори мы можем данную вероятностную меру однозначно продолжить на всю борелевскую σ -алгебру вещественной прямой и таким образом $\forall B \in \mathcal{B}(\mathbb{R})$ будет определена вероятность $P(B) = P\left(\left\{ \omega \in \Omega \mid \xi(\omega) \in B \right\}\right)$. Итак, пусть ξ — случайная величина, заданная на вероятностном пространстве $\langle \Omega, \mathcal{F}, P \rangle$. Тогда получаем новое вероятностное пространство:

$$\langle \Omega, \mathcal{F}, P \rangle \xrightarrow{\xi} \langle \mathbb{R}, \mathcal{B}, P_\xi \rangle$$

Полученное вероятностное пространство называется индуцированным с помощью случайной величины ξ .

Def 1.6.4. Распределением называется функция $P(B)$, сопоставляющая каждому борелевскому множеству на прямой вероятность. Это называется распределением вероятностей случайной величины $\xi(\omega)$.

Распределения бывают дискретными и абсолютно непрерывными.

Дискретные случайные величины

Def 1.6.5. Случайная величина ξ имеет дискретное распределение, если она принимает не более, чем счетное число значений.

$$\exists \{x_1, \dots, x_n, \dots\} \quad \begin{cases} p_i = P(\xi = x_i) > 0 \\ \sum p_i = 1 \end{cases}$$

Таким образом дискретная случайная величина задается законом распределения ($\sum p_i = 1$):

ξ	x_1	x_2	...	x_n	...
p	p_1	p_2	...	p_n	...

Пример 1.6.6. Бросаем кость, получаем следующий закон распределения:

ξ	1	2	3	4	5	6	
p	1/6	1/6	1/6	1/6	1/6	1/6	

Все распределения из предыдущей лекции (биномиальное, геометрическое, гипергеометрическое, Пуассона) являются дискретными.

Пример 1.6.7. Индикатор события A .

$$\xi(\omega) = \begin{cases} 0, & \omega \notin A \text{ — событие } A \text{ произошло} \\ 1, & \omega \in A \text{ — событие } A \text{ не произошло} \end{cases}$$

Числовые характеристики дискретной случайной величины

I. Математическое ожидание (среднее значение)

Def 1.6.8. Математическим ожиданием дискретной случайной величины ξ называется величина

$$\mathbb{E}(\xi) = \sum_{i=1}^{\infty} x_i p_i$$

при условии, что данный ряд сходится абсолютно.

Замечание 1.6.9. Если $\sum_{i=1}^{\infty} |x_i| p_i = \infty$, то говорят, что математическое ожидание не существует.

Смысл: математическое ожидание — значение, вокруг которого группируются все остальные.

Статистический смысл: математическое ожидание — среднее арифметическое наблюдаемых значений случайной величины при большом числе экспериментов.

II. Дисперсия

Def 1.6.10. Дисперсией случайной величины ξ называется среднее квадратов ее отклонения от математического ожидания.

$$\mathbb{D}(\xi) = \mathbb{E}(\xi - \mathbb{E}(\xi))^2 = \sum_{i=1}^{\infty} (x_i - \mathbb{E}(\xi))^2 p_i$$

При условии, что данный ряд сходится.

Замечание 1.6.11. Дисперсию удобнее вычислять по формуле

$$\mathbb{D}(\xi) = \mathbb{E}(\xi^2) - (\mathbb{E}(\xi))^2 = \sum_{i=1}^{\infty} x_i^2 p_i - (\mathbb{E}(\xi))^2$$

Смысл: квадрат среднего разброса рассеивания случайной величины относительно ее математического ожидания.

III. Среднее квадратическое отклонение

Def 1.6.12. Средним квадратическим отклонением (СКО) называется величина

$$\sigma_{\xi} = \sqrt{\mathbb{D}(\xi)}$$

Пример 1.6.13. Бросание кости. Закон распределения описан в 1.6.6.

$$\begin{aligned}\mathbb{E}(\xi) &= \sum_{i=1}^6 x_i p_i = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5 \\ \mathbb{D}(\xi) &= \sum_{i=1}^6 x_i^2 p_i - (\mathbb{E}(\xi))^2 = 1 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 9 \cdot \frac{1}{6} + 16 \cdot \frac{1}{6} + 25 \cdot \frac{1}{6} + 36 \cdot \frac{1}{6} - 3.5^2 \approx 2.92 \\ \sigma_{\xi} &\approx \sqrt{2.92} \approx 1.71\end{aligned}$$

Пример 1.6.14. Индикатор события A обычно обозначается I_A . Опишем закон распределения случайной величины:

$$\begin{array}{c|cc|c} I_A & 0 & 1 \\ \hline p & 1 - P(A) & P(A) \end{array}$$

Тогда

$$\begin{aligned}\mathbb{E}(\xi) &= 0 \cdot (1 - P(A)) + 1 \cdot P(A) = P(A) \\ \mathbb{D}(\xi) &= 0^2 \cdot (1 - P(A)) + 1^2 \cdot P(A) - P(A)^2 = P(A)(1 - P(A)) = pq\end{aligned}$$

Свойства математического ожидания и дисперсии

Def 1.6.15. Случайная величина ξ имеет вырожденное распределение, если

$$\forall \omega \in \Omega \mid \xi(\omega) = C \text{ или } P(\xi = C) = 1$$

Тогда

$$\mathbb{E}(\xi) = C \quad \mathbb{D}(\xi) = 0$$

Lm 1.6.16.

$$\mathbb{E}(\xi + C) = \mathbb{E}(\xi) + C \quad \mathbb{D}(\xi + C) = \mathbb{D}(\xi)$$

□

$$\begin{aligned}\mathbb{E}(\xi + C) &= \sum_{i=1}^{\infty} (x_i + C) \cdot p_i = \sum_{i=1}^{\infty} x_i p_i + C \cdot \underbrace{\sum_{i=1}^{\infty} p_i}_{=1} = \mathbb{E}(\xi) + C \\ \mathbb{D}(\xi + C) &= \mathbb{E}((\xi + C - \mathbb{E}(\xi + C))^2) = \mathbb{E}((\xi + C - \mathbb{E}(\xi) - C)^2) = \mathbb{E}((\xi - \mathbb{E}(\xi))^2) = \mathbb{D}(\xi)\end{aligned}$$

Lm 1.6.17.

$$\mathbb{E}(C\xi) = C\mathbb{E}(\xi) \quad \mathbb{D}(C\xi) = C^2\mathbb{D}(\xi)$$

□

$$\begin{aligned}\mathbb{E}(C\xi) &= \sum_{i=1}^{\infty} Cx_i p_i = C \sum_{i=1}^{\infty} x_i p_i = C\mathbb{E}(\xi) \\ \mathbb{D}(C\xi) &= \mathbb{E}((C\xi - \mathbb{E}(C\xi))^2) = \mathbb{E}((C(\xi - \mathbb{E}(\xi)))^2) = C^2 \cdot \mathbb{E}((\xi - \mathbb{E}(\xi))^2) = C^2\mathbb{D}(\xi)\end{aligned}$$

Lm 1.6.18.

$$\mathbb{E}(\xi + \eta) = \mathbb{E}(\xi) + \mathbb{E}(\eta)$$

□

$$\mathbb{E}(\xi + \eta) = \sum_{i,j=1}^{\infty} (x_i + y_j) \mathbb{P}(\xi = x_i, \eta = y_j) = \sum_{i=1}^{\infty} x_i \left(\sum_{j=1}^{\infty} \mathbb{P}(\xi = x_i, \eta = y_j) \right) + \sum_{j=1}^{\infty} y_j \left(\sum_{i=1}^{\infty} \mathbb{P}(\xi = x_i, \eta = y_j) \right)$$

Применим формулу полной вероятности к внутренним суммам.

$$\mathbb{E}(\xi + \eta) = \sum_{i=1}^{\infty} x_i \mathbb{P}(\xi = x_i) + \sum_{j=1}^{\infty} y_j \mathbb{P}(\eta = y_j) = \mathbb{E}(\xi) + \mathbb{E}(\eta)$$

Def 1.6.19. Случайные величины ξ и η называются независимыми, если

$$\forall i, j \mid \mathbb{P}(\xi = x_i, \eta = y_j) = \mathbb{P}(\xi = x_i) \cdot \mathbb{P}(\eta = y_j)$$

т.е. случайные величины принимают свои значения независимо друг от друга.

Lm 1.6.20. Если ξ и η независимы, то

$$\mathbb{E}(\xi\eta) = \mathbb{E}(\xi)\mathbb{E}(\eta)$$

□

$$\mathbb{E}(\xi\eta) = \sum_{i,j=1}^{\infty} (x_i y_j) \mathbb{P}(\xi = x_i, \eta = y_j) = \sum_{i=1}^{\infty} x_i \left(\sum_{j=1}^{\infty} y_j \mathbb{P}(\xi = x_i, \eta = y_j) \right)$$

Т.к. события независимы, то по определению независимости имеем

$$\mathbb{E}(\xi\eta) = \sum_{i=1}^{\infty} x_i \mathbb{P}(\xi = x_i) \cdot \sum_{j=1}^{\infty} y_j \mathbb{P}(\eta = y_j) = \mathbb{E}(\xi)\mathbb{E}(\eta)$$

Lm 1.6.21.

$$\mathbb{D}(\xi) = \mathbb{E}(\xi^2) - (\mathbb{E}(\xi))^2$$

□

$$\mathbb{D}(\xi) = \mathbb{E}((\xi - \mathbb{E}(\xi))^2) = \mathbb{E}(\xi^2 - 2\xi\mathbb{E}(\xi) + (\mathbb{E}(\xi))^2) = \mathbb{E}(\xi^2) - 2(\mathbb{E}(\xi))^2 + (\mathbb{E}(\xi))^2 = \mathbb{E}(\xi^2) - (\mathbb{E}(\xi))^2$$

Lm 1.6.22.

$$\mathbb{D}(\xi + \eta) = \mathbb{D}(\xi) + \mathbb{D}(\eta) + 2 \operatorname{cov}(\xi, \eta) \quad \operatorname{cov}(\xi, \eta) = \mathbb{E}(\xi\eta) - \mathbb{E}(\xi)\mathbb{E}(\eta)$$

□

$$\begin{aligned}
\mathbb{D}(\xi + \eta) &= \mathbb{E}((\xi + \eta)^2) - (\mathbb{E}(\xi + \eta))^2 \\
&= \mathbb{E}(\xi^2 + 2\xi\eta + \eta^2) - (\mathbb{E}(\xi) + \mathbb{E}(\eta))^2 \\
&= \mathbb{E}(\xi^2) + 2\mathbb{E}(\xi\eta) + \mathbb{E}(\eta^2) - (\mathbb{E}(\xi))^2 - 2\mathbb{E}(\xi)\mathbb{E}(\eta) - (\mathbb{E}(\eta))^2 \\
&= \left(\mathbb{E}(\xi^2) - (\mathbb{E}(\xi))^2\right) + \left(\mathbb{E}(\eta^2) - (\mathbb{E}(\eta))^2\right) + 2\left(\mathbb{E}(\xi\eta) - \mathbb{E}(\xi)\mathbb{E}(\eta)\right) \\
&= \mathbb{D}(\xi) + \mathbb{D}(\eta) + 2\text{cov}(\xi, \eta)
\end{aligned}$$

Lm 1.6.23. Если случайные величины ξ и η независимы, то

$$\mathbb{D}(\xi + \eta) = \mathbb{D}(\xi) + \mathbb{D}(\eta)$$

□ Т.к. ξ и η , то $\text{cov}(\xi, \eta) = 0$ по 1.6.20. Подставим это в 1.6.22 и получим искомое равенство. ■

Другие числовые характеристики

IV Моменты старших порядков

Def 1.6.24. Момент k -ого порядка это

$$m_k = \mathbb{E}(\xi^k)$$

Def 1.6.25. Центральный момент k -ого порядка это

$$\mu_k = \mathbb{E}((\xi - \mathbb{E}(\xi))^k)$$

Замечание 1.6.26. $\mathbb{E}(\xi) = m_1$ — момент первого порядка. $\mathbb{D}(\xi) = \mu_2$ — центральный момент второго порядка.

Замечание 1.6.27. Центральные моменты можно выразить через относительные.

$$\begin{aligned}
\mu_2 &= \mathbb{D}(\xi) = \mathbb{E}(\xi^2) - (\mathbb{E}(\xi))^2 = m_2 - m_1^2 \\
\mu_3 &= m_3 - 3m_2m_1 + 2m_1^3 \\
\mu_4 &= m_4 - 4m_3m_1 + 6m_2m_1^2 - 3m_1^4
\end{aligned}$$

Замечание 1.6.28. Задачу 1.1.17 можно решить, используя математическое ожидание. Пусть $p = \mathbb{P}(A)$ — вероятность того, что иголка пересекла стык, а случайная величина ξ — число пересечений стыка. Тогда, ξ это индикатор события A , значит $\mathbb{E}(\xi) = p$.

Используя свойства математического ожидания получаем, что если растянуть иголку в два раза, то математическое ожидание также возрастет вдвое. Растянем иголку в π раз и сделаем из нее окружность. Тогда математическое ожидание всегда будет равно двум вне зависимости от того, как эта окружность упадет на доску, значит

$$\pi \cdot \mathbb{E}(\xi) = 2 \implies p = \frac{2}{\pi}$$

1.7. Лекция 23.10.17.

Стандартное дискретное распределение

I. Распределение Бернулли B_p (с параметром $0 < p < 1$)

Пусть случайная величина ξ — число успехов при одном испытании, где p — вероятность успеха. Закон распределения имеет вид:

ξ	0	1	
p	$q = 1 - p$	p	

Вычислим его числовые характеристики.

$$\begin{aligned}
\mathbb{E}(\xi) &= 0 \cdot (1 - p) + 1 \cdot p = p \\
\mathbb{D}(\xi) &= 0^2 \cdot (1 - p) + 1^2 \cdot p - p^2 = p - p^2 = p(1 - p) = pq
\end{aligned}$$

Индикатор случайной величины имеет распределение Бернулли.

II. Биномиальное распределение $B_{n,p}$

Пусть случайная величина ξ это число успехов в серии из n независимых испытаний, где p это вероятность успеха при одном испытании.

$$\xi \in B_{n,p} \iff \forall 0 \leq k \leq n, k \in \mathbb{Z} \mid P(\xi = k) = C_n^k p^k q^{n-k}$$

Закон распределения имеет вид:

$$\begin{array}{c|c|c|c|c|c|c} \xi & | 0 & | 1 & | \dots & | k & | \dots & | n \\ \hline p & | q^n & | C_n^1 p q^{n-1} & | \dots & | C_n^k p^k q^{n-k} & | \dots & | p^n \end{array}$$

Заметим, что ξ складывается из числа успехов в каждом испытании, поэтому $\xi = \xi_1 + \dots + \xi_n$, где $\xi_i \in B_p$ — число успехов при i -ом испытании. Получаем

$$\begin{aligned} E(\xi) &= E(\xi_1) + \dots + E(\xi_n) = np \\ D(\xi) &= D(\xi_1) + \dots + D(\xi_n) = npq \\ \sigma_\xi &= \sqrt{npq} \end{aligned}$$

III. Геометрическое распределение G_p ($0 < p < 1$)

Пусть случайная величина ξ это номер первого успешного испытания, где p — вероятность успеха при одном испытании.

$$\xi \in G_p \iff \forall k \in \mathbb{N} \mid P(\xi = k) = q^{k-1} p$$

Закон распределения будет иметь вид:

$$\begin{array}{c|c|c|c|c|c|c} \xi & | 1 & | 2 & | \dots & | k & | \dots & | \dots \\ \hline p & | p & | pq & | \dots & | q^{k-1} p & | \dots & | \dots \end{array}$$

Найдем числовые характеристики.

$$\begin{aligned} E(\xi) &= \sum_{k=1}^{\infty} k \cdot q^{k-1} p = p \sum_{k=1}^{\infty} (q^k)' = p \left(\sum_{k=1}^{\infty} q^k \right)' = p \left(\frac{1}{1-q} \right)' = p \cdot \frac{1}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p} \\ E(\xi^2) &= \sum_{k=1}^{\infty} k^2 \cdot q^{k-1} p = \sum_{k=1}^{\infty} (k(k-1) + k) \cdot q^{k-1} p = \underbrace{\sum_{k=1}^{\infty} k(k-1) \cdot q^{k-1} p}_{\text{ниже}} + \underbrace{\sum_{k=1}^{\infty} k \cdot q^{k-1} p}_{E(\xi)} \\ \sum_{k=1}^{\infty} k(k-1) \cdot q^{k-1} p &= pq \sum_{k=1}^{\infty} (q^k)'' = pq \left(\sum_{k=1}^{\infty} q^k \right)'' = pq \left(\frac{1}{1-q} \right)'' = pq \cdot \frac{2}{(1-q)^3} = \frac{2q}{p^2} \\ E(\xi^2) &= \frac{2q}{p^2} + \frac{1}{p} = \frac{2q+p}{p^2} = \frac{q+1}{p^2} \\ D(\xi) &= E(\xi^2) - (E(\xi))^2 = \frac{q+1}{p^2} - \left(\frac{1}{p} \right)^2 = \frac{q}{p^2} \end{aligned}$$

IV. Распределение Пуассона Π_λ (с параметром $\lambda > 0$)

Def 1.7.1. Случайная величина ξ имеет распределение Пуассона Π_λ с параметром $\lambda > 0$, если

$$\forall k \geq 0, k \in \mathbb{Z} \mid P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$\begin{array}{c|c|c|c|c|c|c} \xi & | 0 & | 1 & | \dots & | k & | \dots & | \dots \\ \hline p & | e^{-\lambda} & | \lambda e^{-\lambda} & | \dots & | \frac{\lambda^k}{k!} e^{-\lambda} & | \dots & | \dots \end{array}$$

Вычислим числовые характеристики.

$$\begin{aligned} E(\xi) &= \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = \lambda e^{-\lambda} \cdot e^\lambda = \lambda \\ E(\xi^2) &= \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=0}^{\infty} \frac{k(k-1)+k}{k!} \cdot \lambda^k e^{-\lambda} = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \underbrace{\sum_{k=0}^{\infty} \frac{k}{k!} \cdot \lambda^k e^{-\lambda}}_{E(\xi)} = \lambda^2 e^{-\lambda} e^\lambda + \lambda = \lambda^2 + \lambda \\ D(\xi) &= E(\xi^2) - (E(\xi))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda \end{aligned}$$

Функция распределения

Def 1.7.2. Функцией распределения (рис. 1.7.3) $F_\xi(x)$ случайной величины ξ называется функция

$$F_\xi(x) = \mathbb{P}(\xi < x), x \in \mathbb{R}$$

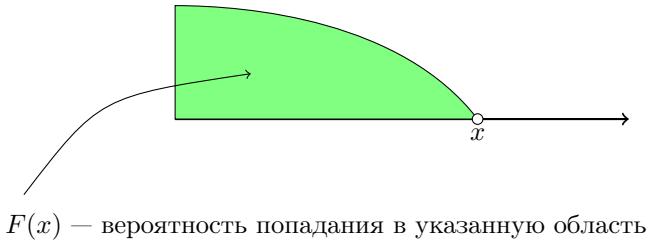


Рис. 1.7.3: Функция распределения

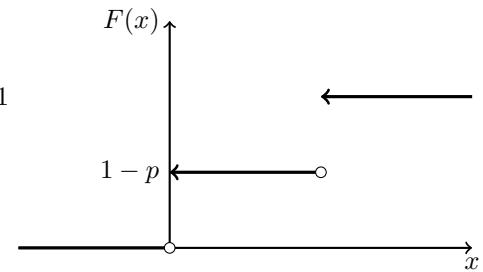


Рис. 1.7.4: Функция распределения Бернулли

Пример 1.7.5. Функция распределения Бернулли B_p (рис. 1.7.4).

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

Свойства функции распределения

Lm 1.7.6. $F(x)$ — ограниченная функция. $0 \leq F(x) \leq 1$

□ $F(x)$ это вероятность \Rightarrow по определению вероятности она лежит в отрезке $[0; 1]$. ■

Lm 1.7.7. $F(x)$ — неубывающая функция. $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$

□

$$x_1 < x_2 \Rightarrow \{\xi \in X_1\} \subset \{\xi \in X_2\} \Rightarrow F(x_1) = \mathbb{P}(\xi < x_1) \leq \mathbb{P}(\xi < x_2) = F(x_2)$$

Lm 1.7.8.

$$\mathbb{P}(\alpha \leq \xi < \beta) = F(\beta) - F(\alpha)$$

□ Воспользуемся свойством аддитивности.

$$\begin{aligned} \mathbb{P}(\xi < \beta) &= \mathbb{P}(\xi < \alpha) + \mathbb{P}(\alpha \leq \xi < \beta) \\ F(\beta) &= F(\alpha) + \mathbb{P}(\alpha \leq \xi < \beta) \Rightarrow \mathbb{P}(\alpha \leq \xi < \beta) = F(\beta) - F(\alpha) \end{aligned}$$

Замечание 1.7.9. Т.к. борелевская σ -алгебра порождается интервалами $(-\infty; x)$, то зная функцию распределения, можем найти вероятности попадания случайной величины в любое борелевское множество. Следовательно, функция распределения полностью задает распределение.

Lm 1.7.10.

$$\lim_{n \rightarrow -\infty} F(x) = 0 \quad \lim_{n \rightarrow +\infty} F(x) = 1$$

□ Т.к. функция распределения монотонна (1.7.7) и ограничена (1.7.6), то данные пределы существуют и конечны. Значит, достаточно найти эти пределы для последовательности $\{x_n\} \rightarrow \pm\infty$. Рассмотрим $A_n = \{n-1 \leq \xi_n < n\}, n \in \mathbb{Z}$ — несовместные события при разных n . По свойству счетной аддитивности получаем

$$1 = \mathbb{P}(\Omega) = \mathbb{P}\left(\sum_{n=-\infty}^{\infty} A_n\right) = \sum_{n=-\infty}^{\infty} \mathbb{P}(A_n) \stackrel{1.7.8}{=} \sum_{n=-\infty}^{\infty} (F(n) - F(n-1)) = \lim_{N \rightarrow \infty} \sum_{n=-N}^N (F(n) - F(n-1))$$

Заметим, что в сумме все слагаемые кроме первого и последнего сокращаются, поэтому получаем

$$1 = \lim_{N \rightarrow \infty} (F(N) - F(-N-1)) = \lim_{N \rightarrow \infty} F(N) - \lim_{N \rightarrow \infty} F(-N-1) \Rightarrow \lim_{N \rightarrow \infty} F(N) = 1 + \lim_{N \rightarrow \infty} F(-N-1)$$

Левый предел не превосходит единицы, а правый предел не меньше нуля, значит получаем, что левый предел в точности равен единице, а правый предел в точности равен нулю. ■

Lm 1.7.11. $F(x)$ непрерывна слева, т.е.

$$\forall x_0 \in \mathbb{R} \mid F(x_0 - 0) = F(x_0)$$

□ Т.к. $F(x)$ монотонна (1.7.7) и ограничена сверху $F(x_0)$, то данный предел существует. Пусть $B_n = \{x_0 - \frac{1}{n} \leq \xi < x_0\}$ — убывающая цепочка событий, т.е. $B_1 \supset B_2 \supset \dots \supset B_n \supset \dots$ и $\bigcup_{n=1}^{\infty} B_n = \emptyset$. Значит по аксиоме непрерывности $P(B_n) \xrightarrow{n \rightarrow \infty} 0$, получаем

$$0 = \lim_{n \rightarrow \infty} P(B_n) = \lim_{n \rightarrow \infty} P\left(x_0 - \frac{1}{n} \leq \xi < x_0\right) = \lim_{n \rightarrow \infty} \left(F(x_0) - F\left(x_0 - \frac{1}{n}\right)\right) = F(x_0) - \lim_{n \rightarrow \infty} F\left(x_0 - \frac{1}{n}\right) = 0$$

$$F(x_0) = \lim_{n \rightarrow \infty} F\left(x_0 - \frac{1}{n}\right)$$

Lm 1.7.12. Скачок в точке x_0 равен вероятности случайной величины попасть в данную точку.

$$F(x_0 + 0) = F(x_0) + P(\xi = x_0) = P(\xi \leq x_0)$$

□ Т.к. $F(x)$ монотонна (1.7.7) и ограничена снизу, то данный предел существует. Пусть $C_n = \{x_0 < \xi < x_0 + \frac{1}{n}\}$ — убывающая цепочка событий, т.е. $C_1 \supset C_2 \supset \dots \supset C_n \supset \dots$ и $\bigcup_{n=1}^{\infty} C_n = \emptyset$. Значит по аксиоме непрерывности $P(C_n) \xrightarrow{n \rightarrow \infty} 0$, получаем

$$P\left(x_0 < \xi < x_0 + \frac{1}{n}\right) + P(\xi = x_0) \xrightarrow{n \rightarrow \infty} 0 + P(\xi = x_0)$$

$$P\left(x_0 \leq \xi < x_0 + \frac{1}{n}\right) \xrightarrow{n \rightarrow \infty} P(\xi = x_0)$$

$$F\left(x_0 + \frac{1}{n}\right) - F(x_0) \xrightarrow{n \rightarrow \infty} P(\xi = x_0)$$

Lm 1.7.13. Если $F(x)$ — непрерывна в точке $x = x_0$, то $P(\xi = x_0) = 0$.

Lm 1.7.14. Если $F(x)$ — непрерывна, то

$$P(\alpha \leq \xi < \beta) = P(\alpha < \xi < \beta) = P(\alpha \leq \xi \leq \beta) F(\beta) - F(\alpha)$$

Теорема 1.7.15. Случайная величина ξ имеет дискретное распределение \iff ее функция распределения имеет ступенчатый вид.

Абсолютно непрерывные случайные величины

Def 1.7.16. Случайная величина ξ имеем абсолютно непрерывное распределение, если существует $f_\xi(x)$ такая, что

$$\forall B \in \mathcal{B}(\mathbb{R}) \mid P(\xi \in B) = \int_B f_\xi(x) dx$$

Функция $f_\xi(x)$ называется плотностью распределения.

Свойства плотности и функции распределения абсолютно непрерывного распределения

$$B = [\alpha; \beta] \implies P(\alpha \leq \xi \leq \beta) = \int_{\alpha}^{\beta} f_\xi(x) dx$$

Используя геометрический смысл определенного интеграла, получаем, что $P(\alpha \leq \xi \leq \beta)$ это площадь под графиком $f_\xi(x)$ на отрезке $[\alpha; \beta]$.

Lm 1.7.17. Условие нормировки: площадь всей фигуры под графиком $f_\xi(x)$ должна равняться единице, т.е.

$$\int_{-\infty}^{\infty} f_\xi(x) dx = 1$$

□ Искомое равенство получается из определения при $B = \mathbb{R}$

$$P(\xi \in \mathbb{R}) = \int_{-\infty}^{\infty} f_\xi(x) dx = 1$$

Lm 1.7.18.

$$F_\xi(x) = \int_{-\infty}^x f_\xi(x)dx$$

□ Искомое равенство получается из определения при $B = (-\infty; x)$.

$$F_\xi(x) = P(\xi < x) = \int_{-\infty}^x f_\xi(z)dz$$

Lm 1.7.19. $F_\xi(x)$ — абсолютно непрерывная функция.

□ Это свойство непрерывности интеграла с переменным верхним пределом.

Lm 1.7.20. $F_\xi(x)$ — дифференцируема почти всюду и $F'_\xi(x) = f_\xi(x)$ для почти всех x .

□ Это следует из теоремы Барроу.

Lm 1.7.21. $f_\xi(x) \geq 0$

□ Это справедливо, т.к. функция распределения неубывающая функция.

Lm 1.7.22.

$$\forall x_0 \in \mathbb{R} \mid P(\xi = x_0) = 0$$

□ Т.к. $F_\xi(x)$ — непрерывная функция, т.е. скачка в данной точке нет, то по 1.7.12 вероятность попадания в точку равна величине скачка, т.е. нулю.

Lm 1.7.23.

$$P(\alpha \leq \xi < \beta) = P(\alpha \leq \xi \leq \beta) = P(\alpha < \xi < \beta) = P(\alpha < \xi \leq \beta) = F(\beta) - F(\alpha)$$

□ Это следствие из 1.7.22.

Теорема 1.7.24.

$$\left. \begin{array}{l} \forall x \in \mathbb{R} \mid f(x) \geq 0 \\ \int_{-\infty}^{\infty} f(x)dx = 1 \end{array} \right\} \Rightarrow f(x) \text{ это плотность некоторого распределения}$$

Независимость случайных величин

Def 1.7.25. Случайные величины ξ и η называются независимыми, если

$$\forall B_1, B_2 \in \mathcal{B}(\mathbb{R}) \mid P(\xi \in B_1, \eta \in B_2) = P(\xi \in B_1)P(\eta \in B_2)$$

Def 1.7.26. Случайные величины ξ_1, \dots, ξ_n называются независимыми, если

$$\forall B_1, \dots, B_n \in \mathcal{B}(\mathbb{R}) \mid P(\xi_1 \in B_1, \dots, \xi_n \in B_n) = P(\xi_1 \in B_1) \dots P(\xi_n \in B_n)$$

Замечание 1.7.27. Из независимости в совокупности следует попарная независимость. Обратное в общем случае неверно.

Замечание 1.7.28. В дальнейшем под независимыми случайными величинами понимает независимые в совокупности случайные величины.

Числовые характеристики абсолютно непрерывного распределения

I. Математическое ожидание

Def 1.7.29. Математическим ожиданием абсолютно непрерывной случайной величины ξ называется число

$$\mathbb{E}(\xi) = \int_{-\infty}^{\infty} xf_\xi(x)dx$$

при условии, что данный интеграл сходится абсолютно.

Замечание 1.7.30.

$$\int_{-\infty}^{\infty} |x| f_\xi(x)dx \implies \text{математическое ожидание не существует}$$

II. Дисперсия

Def 1.7.31. Дисперсией абсолютно непрерывной случайной величины ξ называется число

$$\mathbb{D}(\xi) = \mathbb{E}((\xi - \mathbb{E}(\xi))^2) = \int_{-\infty}^{\infty} (x - \mathbb{E}(\xi))^2 f_{\xi}(x) dx$$

при условии, что данный интеграл сходится.

Замечание 1.7.32. Для вычисления дисперсии удобнее формула

$$\mathbb{D}(\xi) = \mathbb{E}(\xi^2) - (\mathbb{E}(\xi))^2 = \int_{-\infty}^{\infty} x^2 f_{\xi}(x) dx - (\mathbb{E}(\xi))^2$$

III. Среднее квадратичное отклонение

Def 1.7.33.

$$\sigma_{\xi} = \sqrt{\mathbb{D}(\xi)}$$

Замечание 1.7.34. Смысл и свойства этих числовых характеристик полностью идентичны случаю дискретной случайной величины.

Другие числовые характеристики

IV Моменты старших порядков

Def 1.7.35. Момент k -ого порядка это

$$m_k = \mathbb{E}(\xi^k) = \int_{-\infty}^{\infty} x^k f_{\xi}(x) dx$$

Def 1.7.36. Центральный момент k -ого порядка это

$$\mu_k = \mathbb{E}((\xi - \mathbb{E}(\xi))^k) = \int_{-\infty}^{\infty} (x - \mathbb{E}(\xi))^k f_{\xi}(x) dx$$

V Медиана

Def 1.7.37. Медианой Me называется значение случайной величины ξ такое, что

$$\mathbb{P}(\xi < Me) = \mathbb{P}(\xi > Me)$$

VI Мода

Def 1.7.38. Модой Mo абсолютно непрерывного распределения называется точка локального максимума плотности.

1.8. Лекция 23.10.24.

Стандартное абсолютно непрерывное распределение

I Равномерное распределение

Def 1.8.1. Случайная величина ξ равномерно распределена на отрезке $[a; b]$ (обозначается $\xi \in U(a; b)$), если ее плотность на данном отрезке постоянна.

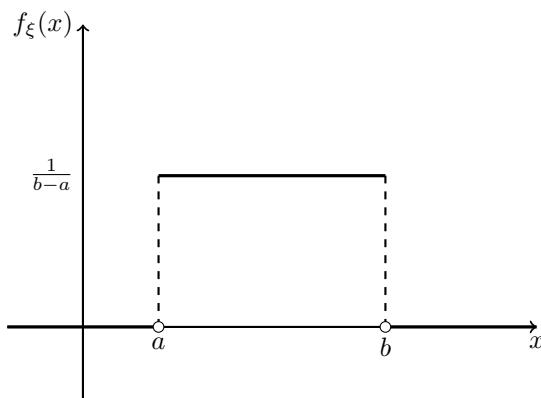


Рис. 1.8.2: Плотность равномерного распределения

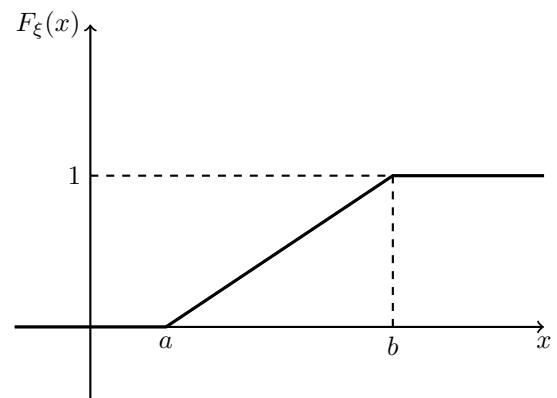


Рис. 1.8.3: Функция распределения

Учитывая условие нормировки (рис. 1.8.2) получаем, что плотность ξ будет иметь вид

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

Найдем функцию распределения (рис. 1.8.3).

$$\begin{aligned} F_{\xi}(x) &= \int_{-\infty}^{\infty} f_{\xi}(x) dx \\ x < a \implies F_{\xi}(x) &= 0 \\ a \leq x \leq b \implies F_{\xi}(x) &= \int_a^x \frac{dx}{b-a} = \frac{x-a}{b-a} \\ x > b \implies F_{\xi}(x) &= 1 \\ F_{\xi}(x) &= \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases} \end{aligned}$$

Вычислим числовые характеристики.

$$\begin{aligned} \mathbb{E}(\xi) &= \int_{-\infty}^{\infty} xf(x) dx = \int_a^b \frac{x dx}{b-a} = \frac{1}{b-a} \cdot \frac{x^2}{2} \Big|_a^b = \frac{a+b}{2} \\ \mathbb{E}(\xi^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx = \int_a^b \frac{x^2 dx}{b-a} = \frac{1}{b-a} \cdot \frac{x^3}{3} \Big|_a^b = \frac{a^2 + ab + b^2}{3} \\ \mathbb{D}(\xi) &= \mathbb{E}(\xi^2) - (\mathbb{E}(\xi))^2 = \frac{a^2 + ab + b^2}{3} - \frac{a^2 + 2ab + b^2}{4} = \frac{(b-a)^2}{12} \\ \sigma_{\xi} &= \sqrt{\mathbb{D}(\xi)} = \frac{b-a}{2\sqrt{3}} \end{aligned}$$

Вычислим вероятность попадания в интервал.

$$\mathbb{P}(\alpha < \xi < \beta) = \frac{\beta - \alpha}{b - a} \quad \alpha, \beta \in [a; b]$$

Равномерное распределение возникает в задачах, связанных со временем, а также в задачах, связанных с генерацией псевдослучайных чисел ($\xi \in U(0; 1)$).

II Показательное (экспоненциальное) распределение

Def 1.8.4. Случайная величина ξ имеет показательное распределение с параметром $\alpha > 0$ (обозначается $\xi \in E_{\alpha}$), если ее плотность имеет вид (рис. 1.8.5):

$$f_{\xi}(x) = \begin{cases} 0, & x < 0 \\ \alpha e^{-\alpha x}, & x \geq 0 \end{cases}$$

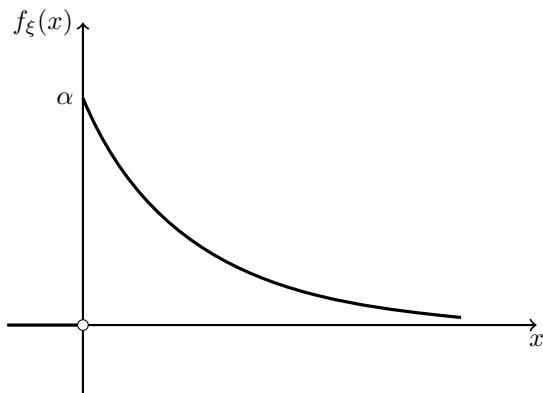


Рис. 1.8.5: Плотность показательного распределения

Найдем функцию распределения (рис. 1.8.6).

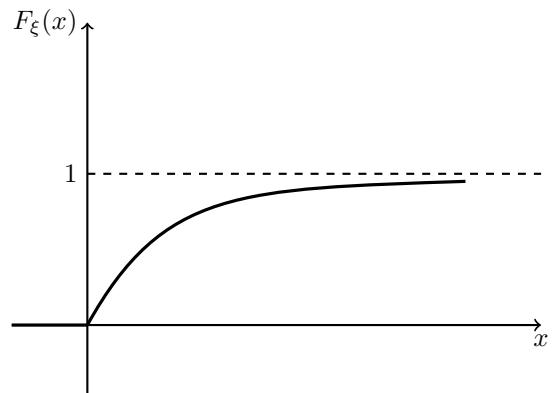


Рис. 1.8.6: Функция распределения

$$\begin{aligned}
F_\xi(x) &= \int_{-\infty}^{\infty} f_\xi(x) dx \\
x < 0 \implies F_\xi(x) &= 0 \\
x \geq 0 \implies F_\xi(x) &= \int_0^x \alpha e^{-\alpha x} = -e^{-\alpha x} \Big|_0^x = 1 - e^{-\alpha x} \\
iF_\xi(x) &= \begin{cases} 0, & x < 0 \\ 1 - e^{-\alpha x}, & x \geq 0 \end{cases}
\end{aligned}$$

Вычислим числовые характеристики.

$$\begin{aligned}
\mathbb{E}(\xi) &= \int_{-\infty}^{\infty} xf(x) dx = \int_0^{\infty} x \alpha e^{-\alpha x} dx = \left[\begin{array}{l} u = x \\ dv = \alpha e^{-\alpha x} dx \end{array} \right] \frac{du = dx}{v = -e^{-\alpha x}} \\
&= -xe^{-\alpha x} \Big|_0^{\infty} + \int_0^{\infty} e^{-\alpha x} dx = -\lim_{x \rightarrow \infty} \frac{x}{e^{\alpha x}} - \frac{1}{\alpha} e^{-\alpha x} \Big|_0^{\infty} = -\frac{1}{\alpha} \left(\lim_{x \rightarrow \infty} e^{-\alpha x} - e^0 \right) = \frac{1}{\alpha} \\
\mathbb{E}(\xi^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx = \int_0^{\infty} x^2 \alpha e^{-\alpha x} dx = \left[\begin{array}{l} u = x^2 \\ dv = \alpha e^{-\alpha x} dx \end{array} \right] \frac{du = 2x dx}{v = -e^{-\alpha x}} \\
&= -x^2 e^{-\alpha x} \Big|_0^{\infty} + 2 \int_0^{\infty} x e^{-\alpha x} dx = \frac{2}{\alpha} \int_0^{\infty} x \alpha e^{-\alpha x} dx = \frac{2}{\alpha} \mathbb{E}(\xi) = \frac{2}{\alpha^2} \\
\mathbb{D}(\xi) &= \mathbb{E}(\xi^2) - (\mathbb{E}(\xi))^2 = \frac{2}{\alpha^2} - \left(\frac{1}{\alpha} \right)^2 = \frac{1}{\alpha^2} \\
\sigma_\xi &= \sqrt{\mathbb{D}(\xi)} = \frac{1}{\alpha}
\end{aligned}$$

Используя функцию распределения, вычислим вероятность попадания в интервал.

$$\mathbb{P}(\alpha < \xi < \beta) = e^{-a\alpha} - e^{-b\alpha}$$

Из непрерывных распределений только показательное обладает свойством нестарения.

Теорема 1.8.7.

$$\mathbb{P}(\xi > x + y \mid \xi > x) = \mathbb{P}(\xi > y)$$

□

$$\begin{aligned}
\mathbb{P}(\xi > x + y \mid \xi > x) &= \frac{\mathbb{P}(\xi > x + y, \xi > x)}{\mathbb{P}(\xi > x)} = \frac{1 - \mathbb{P}(\xi < x + y)}{1 - \mathbb{P}(\xi < x)} = \frac{1 - F(x + y)}{1 - F(x)} = \frac{1 - 1 + e^{-\alpha(x+y)}}{1 - 1 + e^{-\alpha x}} \\
&= e^{-\alpha y} = 1 - (1 - e^{-\alpha y}) = 1 - F(y) = \mathbb{P}(\xi > y)
\end{aligned}$$

■

Показательное распределение возникает, (например) если нужно рассчитать время работы надежного прибора до поломки. Также показательное распределение можно трактовать как время между появлениемми двух соседних редких событий. Возникает в системах массового обслуживания, теории надежности, теории катастроф и ряде областей физики.

III Нормальное (Гауссовское) распределение

Def 1.8.8. Случайное величина ξ имеет нормальное распределение с параметрами a и $\sigma^2 > 0$ (обозначается $\xi \in N(a; \sigma^2)$), если ее плотность имеет вид

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} \quad x \in (-\infty; \infty)$$

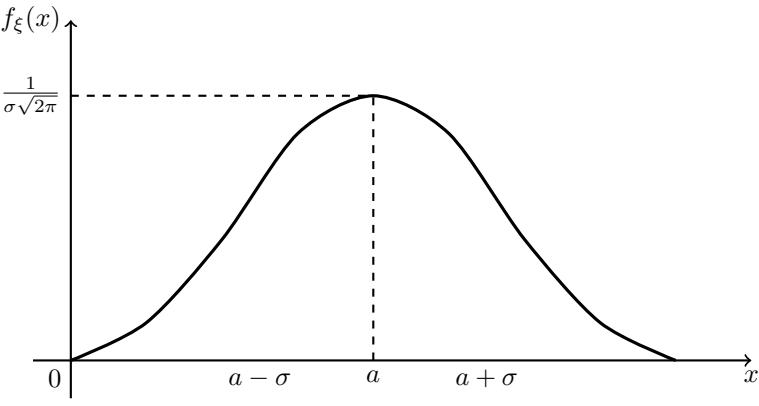


Рис. 1.8.9: Нормальное распределение

Смысл параметров распределения.

$$\mathbb{E}(\xi) = a \quad \mathbb{D}(\xi) = \sigma^2 \quad \sigma_\xi = \sigma$$

Функция распределения будет иметь вид

$$F_\xi(x) = \int_{-\infty}^x f_\xi(t) dt$$

Стандартное нормальное распределение

Def 1.8.11. Стандартным нормальным распределением называется нормальное распределение с параметрами $a = 0$, $\sigma^2 = 1$.

Получаем функцию плотности (функцию Гаусса, рис. 1.8.12):

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

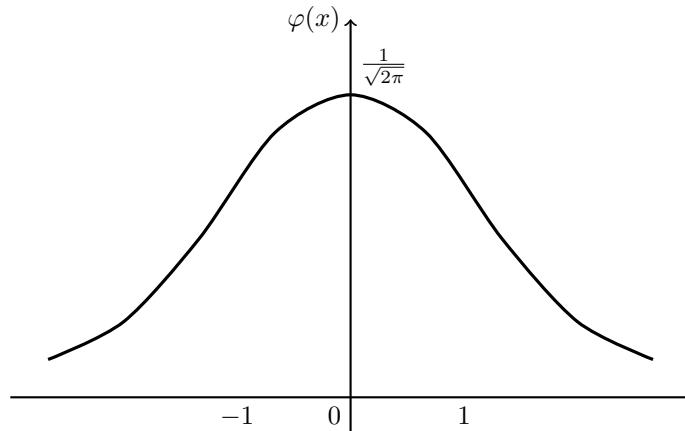


Рис. 1.8.12: Функция Гаусса

Функция распределения $F_0(x)$ будет иметь вид

$$F_0(x) = \frac{1}{2\pi} \int_{-\infty}^x e^{-z^2/2} dz$$

Замечание 1.8.13. По формуле Пуассона

$$\int_{-\infty}^{\infty} e^{-z^2/2} dz = \sqrt{2\pi}$$

Значит условие нормировки выполнено $\Rightarrow \varphi(x)$ это действительно плотность некоторого распределения.

Замечание 1.8.14.

$$F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-z^2/2} dz + \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz = \frac{1}{2} + \Phi(x)$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz$$

где $\Phi(x)$ это функция Лапласа.

Вычислим характеристики.

$$\mathbb{E}(\xi) = \int_{-\infty}^{\infty} x\varphi(x)dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} xe^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} d\left(\frac{x^2}{2}\right)$$

$$= -\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \Big|_{-\infty}^{\infty} = -\frac{1}{\sqrt{2\pi}} \left(\lim_{x \rightarrow \infty} e^{-x^2/2} - \lim_{x \rightarrow -\infty} e^{-x^2/2} \right) = 0$$

$$\mathbb{D}(\xi) = \int_{-\infty}^{\infty} x^2 \varphi(x)dx - (\mathbb{E}(\xi))^2 = \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx}_{0} = \left[\begin{array}{l} u = x \\ dv = xe^{-x^2/2} dx \\ du = dx \\ v = -e^{-x^2/2} \end{array} \right]$$

$$= \frac{1}{\sqrt{2\pi}} \left(\underbrace{-xe^{-x^2/2} \Big|_{-\infty}^{\infty}}_0 + \int_{-\infty}^{\infty} e^{-x^2/2} dx \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} dx = 1$$

Связь между нормальным и стандартным нормальным распределением

Lm 1.8.15.

$$\xi \in N(a; \sigma^2) \implies F_{\xi}(x) = F_0\left(\frac{x-a}{\sigma}\right)$$

□

$$F_{\xi}(x) = \frac{1}{\sigma\sqrt{2\pi}} = \int_{-\infty}^x e^{-\frac{(t-a)^2}{2\sigma^2}} dt = \left[\begin{array}{l} z = \frac{t-a}{\sigma} \\ z(-\infty) = -\infty \\ z(x) = \frac{x-a}{\sigma} \end{array} \right] t = \sigma z + a \quad dt = \sigma dz$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sigma}} e^{-z^2/2} \sigma dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sigma}} e^{-z^2/2} dz = F_0\left(\frac{x-a}{\sigma}\right)$$

■

Lm 1.8.16.

$$\xi \in N(a; \sigma^2) \implies \eta = \frac{\xi - a}{\sigma} \in N(0; 1)$$

□

$$F_{\eta}(x) = \mathbb{P}(\eta < x) = \mathbb{P}\left(\frac{\xi - a}{\sigma} < x\right) = \mathbb{P}(\xi < \sigma x + a) = F_{\xi}(\sigma x + a) = F_0\left(\frac{\sigma x + a - a}{\sigma}\right) = F_0(x) \implies \eta \in N(0; 1)$$

■

Lm 1.8.17.

$$\xi \in N(a; \sigma^2) \implies \begin{cases} \mathbb{E}(\xi) = a \\ \mathbb{D}(\xi) = \sigma^2 \end{cases}$$

□

$$\eta = \frac{\xi - a}{\sigma} \in N(0; 1) \implies \begin{cases} \mathbb{E}(\eta) = 0 \\ \mathbb{D}(\eta) = 1 \end{cases}$$

$$\xi = \sigma\eta + a \implies \begin{cases} \mathbb{E}(\xi) = \sigma\mathbb{E}(\eta) + a = a \\ \mathbb{D}(\xi) = \sigma^2 \mathbb{D}(\eta) = \sigma^2 \end{cases}$$

■

Lm 1.8.18.

$$\mathbb{P}(\alpha < \xi < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right)$$

Lm 1.8.19.

$$\mathbb{P}(\alpha < \xi < \beta) = F_\xi(\beta) - F_\xi(\alpha) = F_0\left(\frac{\beta - a}{\sigma}\right) - F_0\left(\frac{\alpha - a}{\sigma}\right) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right)$$

Lm 1.8.20. Вероятность отклонения нормальной случайной величины от ее среднего значения (или попадания в симметричный интервал относительно среднего) равна

$$\mathbb{P}(|\xi - a| < t) = 2\Phi\left(\frac{t}{\sigma}\right)$$

□

$$\mathbb{P}(|\xi - a| < t) = \mathbb{P}(a - t < \xi < a + t) = \Phi\left(\frac{a + t - a}{\sigma}\right) - \Phi\left(\frac{a - t - a}{\sigma}\right) = 2\Phi\left(\frac{t}{\sigma}\right)$$

■

Замечание 1.8.21. Если $\Phi(x)$ заменить на $F_0(x)$, то формула 1.8.20 приобретет вид

$$\mathbb{P}(|\xi - a| < t) = 2F_0\left(\frac{t}{\sigma}\right) - 1$$

Lm 1.8.22 (Правило «трех сигм»).

$$\mathbb{P}(|\xi - a| < 3\sigma) \approx 0.9973$$

□

$$\mathbb{P}(|\xi - a| < 3\sigma) = 2\Phi\left(\frac{3\sigma}{\sigma}\right) = 2\Phi(3) \approx 0.9973$$

■

Замечание 1.8.23. Как правило, при нормальном распределении случайной величины, мы почти гарантированно попадаем в интервал $[a - 3\sigma; a + 3\sigma]$.

Lm 1.8.24 (Устойчивость относительно суммирования). Если ξ_1 и ξ_2 независимые случайные величины, то

$$\begin{cases} \xi_1 \in N(a_1; \sigma_1^2) \\ \xi_2 \in N(a_2; \sigma_2^2) \end{cases} \implies \xi_1 + \xi_2 \in N(a_1 + a_2; \sigma_1^2 + \sigma_2^2)$$

Коэффициенты асимметрии и эксцесса

Def 1.8.25. Асимметрией распределения называется число

$$A_S = \mathbb{E}\left(\left(\frac{\xi - a}{\sigma}\right)^3\right) = \frac{\mu_3}{\sigma^3}$$

Замечание 1.8.26. Если распределение симметрично относительно точки a , то $A_S = 0$. Если $A_S > 0$, то график плотности имеет более крутой спуск слева и наоборот.

Def 1.8.27. Эксцессом распределения называется число

$$E_k = \mathbb{E}\left(\left(\frac{\xi - a}{\sigma}\right)^4\right) - 3 = \frac{\mu_4}{\sigma^4} - 3$$

Замечание 1.8.28. Если случайная величина имеет нормальное распределение $E_k = 0$. При $E_k > 0$ имеет более «острую» вершину, чем у нормального распределения, а при $E_k < 0$ — более «тупую».

Стандартизация случайной величины

Def 1.8.29. Пусть имеется случайная величина ξ . Соответствующей ей стандартной случайной величиной называется случайная величина

$$\eta = \frac{\xi - \mathbb{E}(\xi)}{\sigma_\xi}$$

Lm 1.8.30.

$$\mathbb{E}(\eta) = 0 \quad \mathbb{D}(\eta) = 1$$

□

$$\mathbb{E}(\eta) = \mathbb{E}\left(\frac{\xi - \mathbb{E}(\xi)}{\sigma_\xi}\right) = \frac{1}{\sigma_\xi}(\mathbb{E}(\xi) - \mathbb{E}(\xi)) = 0$$

$$\mathbb{D}(\eta) = \mathbb{D}\left(\frac{\xi - \mathbb{E}(\xi)}{\sigma_\xi}\right) = \frac{1}{\sigma_\xi^2}\mathbb{D}(\xi - \mathbb{E}(\xi)) = \frac{1}{\sigma_\xi^2}\mathbb{D}(\xi) = \frac{1}{\sigma_\xi^2}\sigma_\xi^2 = 1$$

■

Замечание 1.8.31. Стандартизованная случайная величина не зависит от единиц измерения.

Замечание 1.8.32. При операции стандартизации в общем случае тип распределения может поменяться.

1.9. Лекция 23.10.31.

Сингулярное распределение

Def 1.9.1. Случайная величина ξ имеет сингулярное распределение, если существует борелевское множество $B \in \mathcal{B}$ с нулевой мерой Лебега такое, что $P(\xi \in B) = 1$, но $\forall x \in B \mid P(\xi = x) = 0$.

Заметим, что данное множество B — несчетное, т.к. в противном случае по свойству счетной аддитивности получили бы $P(\xi \in B) = 0$. Таким образом при сингулярном распределении случайная величина ξ распределена на несчетном множестве лебеговой меры ноль. Т.к. $P(\xi = x) = 0$, то функция сингулярного распределения является непрерывной.

Пример 1.9.2. Функция распределения случайной величины ξ — лестница Кантора.

$$F_\xi(x) = \begin{cases} 0, & x \leq 0 \\ \frac{1}{2}F_\xi(3x), & 0 < x \leq \frac{1}{3} \\ \frac{1}{2}, & \frac{1}{3} < x \leq \frac{2}{3} \\ \frac{1}{2} + \frac{1}{2}F_\xi(3x-2), & \frac{2}{3} < x < 1 \\ 1, & x \geq 1 \end{cases}$$

Теорема 1.9.3. (Лебега) Пусть $F_\xi(x)$ — функция распределения некоторой случайной величины ξ , то ее можно представить в виде линейной комбинации

$$F_\xi(x) = p_1 F_1(x) + p_2 F_2(x) + p_3 F_3(x)$$

где $p_1 + p_2 + p_3 = 1$ и $F_1(x)$, $F_2(x)$, $F_3(x)$ — это функции дискретного, абсолютно непрерывного и сингулярного распределения соответственно.

Таким образом существуют только дискретные, абсолютно непрерывные, сингулярные распределения и их смеси.

Преобразование случайных величин

Пусть ξ это случайная величина на вероятностном пространстве $\langle \Omega, \mathcal{F}, P \rangle$. Рассмотрим функцию $g(\xi): \Omega \rightarrow \mathbb{R}$, где $g(x): \mathbb{R} \rightarrow \mathbb{R}$. Вопрос: когда $g(\xi)$ также является случайной величиной?

Def 1.9.4. Функция $g(x)$ называется борелевской, если $\forall B \in \mathcal{B} \mid g^{-1}(B) \in \mathcal{B}$, т.е. прообраз любого борелевского множества является борелевским множеством.

Теорема 1.9.5. Если функция $g(x)$ — борелевская, ξ — случайная величина на вероятностном пространстве $\langle \Omega, \mathcal{F}, P \rangle$, то $g(\xi)$ также будет случайной величиной на том же вероятностном пространстве $\langle \Omega, \mathcal{F}, P \rangle$.

□ $\forall B \in \mathcal{B}$ прообраз $g^{-1}(B) = \xi^{-1}(B') \in \mathcal{F} \implies g(\xi)$ измеримая функция и является случайной величиной. ■

Замечание 1.9.6. На практике почти все функции борелевские.

Замечание 1.9.7. Возможная ситуация, когда ξ — абсолютно непрерывная случайная величина, $g(x)$ — непрерывная функция, однако $g(\xi)$ имеет дискретное распределение.

Теорема 1.9.8. Пусть $f_\xi(x)$ — плотность случайной величины ξ и функция $g(x)$ строго монотонная. Тогда случайная величина $\eta = g(\xi)$ имеет плотность

$$f_\eta(x) = |h'(x)| f_\xi(h(x)) \quad h(x) = g^{-1}(x)$$

□ Пусть $g(x)$ — возрастающая функция, тогда $h(x)$ также возрастающая и случайная величина $g(\xi)$ распределена на $(g(-\infty); g(+\infty))$. Ее функция распределения

$$F_\eta(x) = \mathbb{P}(\eta < x) = \mathbb{P}(g(\xi) < x) = \mathbb{P}(\xi < h(x)) = \int_{-\infty}^{h(x)} f_\xi(t) dt$$

$$\left[\begin{array}{l} t = h(y) \\ y(-\infty) = g(-\infty) \end{array} \right. \quad \left. \begin{array}{l} dt = h'(y) dy \\ y(h(x)) = g(h(x)) = x \end{array} \right] \quad y = g(t)$$

$$F_\eta(x) = \int_{g(-\infty)}^x f_\xi(h(y)) h'(y) dy \implies f_\eta(x) = |h'(x)| f_\xi(h(x))$$

Пусть $g(x)$ — убывающая функция, тогда $h(x)$ также убывающая и случайная величина $g(\xi)$ распределена на $(g(+\infty); g(-\infty))$. Ее функция распределения

$$F_\eta(x) = \mathbb{P}(\eta < x) = \mathbb{P}(g(\xi) < x) = \mathbb{P}(\xi > h(x)) = \int_{h(x)}^{+\infty} f_\xi(t) dt$$

$$\left[\begin{array}{l} t = h(y) \\ y(h(x)) = g(h(x)) = x \end{array} \right. \quad \left. \begin{array}{l} dt = h'(y) dy \\ y(+\infty) = g(+\infty) \end{array} \right] \quad y = g(t)$$

$$F_\eta(x) = \int_x^{g(+\infty)} f_\xi(h(y)) h'(y) dy = \int_{g(+\infty)}^x -h'(y) f_\xi(h(y)) dy \implies f_\eta(x) = |h'(x)| f_\xi(h(x))$$

■

Теорема 1.9.9. (О линейном преобразовании) Пусть случайная величина ξ имеет плотность $f_\xi(x)$. Тогда случайная величина $\eta = a\xi + b$ имеет плотность

$$f_\eta(x) = \frac{1}{|a|} f_\xi\left(\frac{x-b}{a}\right)$$

□ Линейное преобразование монотонно, имеем

$$g(x) = ax + b \quad h(x) = \frac{x-b}{a} \quad h'(x) = \frac{1}{a}$$

Подставляя все это в 1.9.8, получаем

$$f_\eta(x) = \frac{1}{|a|} f_\xi\left(\frac{x-b}{a}\right)$$

■

Lm 1.9.10.

$$\xi \in N(a; \sigma^2) \implies \eta = \gamma\xi + b \in N(\gamma a + b; \gamma^2 \sigma^2)$$

□

$$f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-a)^2}{2\sigma^2}} \quad x \in (-\infty; +\infty)$$

$$\eta = \gamma\xi + b \in (-\infty; +\infty)$$

$$f_\eta(x) = \frac{1}{|\gamma|} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\frac{x-b}{\gamma} - a)^2}{2\sigma^2}\right) = \frac{1}{\sigma|\gamma|\sqrt{2\pi}} \exp\left(-\frac{(x - b - \gamma a)^2}{2(\sigma\gamma)^2}\right)$$

Получили плотность нормального распределения с параметрами $\tilde{a} = b + \gamma a$ и $\tilde{\sigma}^2 = (\gamma\sigma)^2$

Lm 1.9.11. Если $\eta \in N(0; 1)$, то $\xi = \sigma\eta + a \in N(a; \sigma^2)$

□ Это следствие из предыдущей леммы.

Lm 1.9.12. Если $\eta \in U(0; 1)$, то $\xi = a\eta + b \in U(b; b + a)$.

□ **TODO:** на практике

Lm 1.9.13. Если $\xi \in E_\alpha$, то $\eta = \beta\xi \in E_{\frac{\alpha}{\beta}}$.

□ **TODO:** на практике

Квантильное преобразование

Теорема 1.9.14. Пусть функция распределения $F_\xi(x)$ случайной величины ξ непрерывна. Тогда случайная величина $\eta = F(\xi) \in U(0; 1)$.

□ Ясно, что новая случайная величина $\eta \in [0; 1]$, т.к. $0 \leq F(x) \leq 1$. Рассмотрим два случая.

Случай I. $F(x)$ — строго возрастающая функция

В этом случае $F(x)$ имеет обратную функцию $F^{-1}(x)$ и

$$F_\eta(x) = P(\eta < x) = P(F(\xi) < x) = P(\xi < F^{-1}(x)) = \begin{cases} 0, & x < 0 \\ F(F^{-1}(x)) = x, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

Получили функцию распределения $\in U(0; 1)$.

Случай II. $F(x)$ — не является строго возрастающей функцией, т.е. есть интервалы постоянства

В этом случае обозначим через $F^{-1}(x)$ самую левую точку такого интервала:

$$F^{-1}(x) = \min_t \{t \mid F(t) = x\}$$

Тогда при $0 < x < 1$

$$F_\eta(x) = P(F(\xi) < x) = P(\xi < F^{-1}(x)) = F(F^{-1}(x)) = x$$

Сформулируем обратную теорему. Пусть $F(x)$ — функция распределения случайной величины ξ , причем не обязательно непрерывная. Обозначим:

$$F^{-1}(x) = \inf_t \{t \mid F(t) \geq x\}$$

Теорема 1.9.15. Пусть случайная величина $\eta \in U(0; 1)$, а $F(x)$ — произвольная функция распределения. Тогда случайная величина $\xi = F^{-1}(\eta)$ имеет функцию распределения $F(x)$.

□

$$F_\xi(x) = P(\xi < x) = P(F^{-1}(\eta) < x) = P(\eta < F(x)) = \frac{F(x) - 0}{1 - 0} = F(x)$$

Замечание 1.9.16. Датчики случайных чисел имеют распределение $U(0; 1)$. С помощью квантильного преобразования мы можем смоделировать любое другое распределение.

Пример 1.9.17. Смоделируем показательное распределение E_α .

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

$$\eta = 1 - e^{-\alpha x} \implies x = -\frac{1}{\alpha} \ln(1 - \eta) = F^{-1}(x)$$

Таким образом, если $\eta \in U(0; 1)$, то $-\frac{1}{\alpha} \ln(1 - \eta) \in E_\alpha$.

Пример 1.9.18. Смоделируем стандартное нормальное распределение $N(0; 1)$.

$$F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-z^2/2} dz$$

Пусть $F_0^{-1}(x)$ — обратная функция. Таким образом, если $\eta \in U(0; 1)$, то $F_0^{-1}(\eta) \in N(0; 1)$.

Математическое ожидание преобразованной случайной величины

Теорема 1.9.19. Для произвольной борелевской функции $g(x)$ справедливо:

1. Если ξ это дискретная случайная величина, то

$$\mathbb{E}(g(\xi)) = \sum_{i=1}^{\infty} g(x_i) P(\xi = x_i)$$

при условии, что данный ряд сходится абсолютно.

2. Если ξ это абсолютно непрерывная случайная величина, то

$$\mathbb{E}(g(\xi)) = \int_{-\infty}^{\infty} g(x) f_{\xi}(x) dx$$

при условии, что данный интеграл сходится абсолютно.

Свойства моментов

Lm 1.9.20.

$$\xi \geq 0 \implies \mathbb{E}(\xi) \geq 0$$

Lm 1.9.21.

$$\xi \geq \eta \implies \mathbb{E}(\xi) \geq \mathbb{E}(\eta)$$

□

$$\xi \geq \eta \implies \xi - \eta \geq 0 \implies \mathbb{E}(\xi - \eta) = \mathbb{E}(\xi) - \mathbb{E}(\eta) \geq 0 \implies \mathbb{E}(\xi) \geq \mathbb{E}(\eta)$$

■

Lm 1.9.22.

$$|\xi| \geq |\eta| \implies \mathbb{E}(|\xi|^k) \geq \mathbb{E}(|\eta|^k)$$

Lm 1.9.23. Если существует момент m_t случайной величины ξ , то существуют ее моменты меньшего порядка m_s ($s < t$).

□ Заметим, что

$$\left. \begin{array}{ll} x \geq 1: & |x|^s \leq |x|^t \\ x < 1: & |x|^s \leq 1 \end{array} \right\} \implies |x|^s \leq \max(|x|^t, 1) \leq |x|^t + 1 \implies \mathbb{E}(|\xi|^s) \leq \mathbb{E}(|\eta|^t) + 1$$

Если $\mathbb{E}(|\eta|^t) < \infty$, то и $\mathbb{E}(|\eta|^s) < \infty$.

■

Теорема 1.9.24. (Неравенство Йенсена) Пусть функция $g(x)$ выпукла вниз. Тогда для любой случайной величины ξ верно неравенство

$$\mathbb{E}(\xi) \geq g(\mathbb{E}(\xi))$$

Если функция выпукла вверх, то знак неравенства поменяется на противоположный.

□ Если функция $g(x)$ выпукла вниз, то в любой точке ее графика можно провести прямую, лежащую ниже графика, т.е.

$$\forall x_0 \exists k(x_0) \mid g(x) \geq g(x_0) + k(x_0)(x - x_0)$$

Положим $x_0 = \mathbb{E}(\xi)$ и $x = \xi$, тогда

$$g(\xi) \geq g(\mathbb{E}(\xi)) + k(\mathbb{E}(\xi))(\xi - \mathbb{E}(\xi))$$

По 1.9.21 получаем, что

$$\begin{aligned} \mathbb{E}(g(\xi)) &\geq \mathbb{E}(g(\mathbb{E}(\xi))) + k(\mathbb{E}(\xi)) \underbrace{\mathbb{E}(\xi - \mathbb{E}(\xi))}_0 \\ \mathbb{E}(g(\xi)) &\geq g(\mathbb{E}(\xi)) \end{aligned}$$

■

Пример 1.9.25. С помощью 1.9.24 можно, например, получить следующие неравенства:

$$\begin{aligned} \mathbb{E}(e^{\xi}) &\geq e^{\mathbb{E}(\xi)} \\ \mathbb{E}(\xi^n) &\geq (\mathbb{E}(\xi))^n \quad (n > 1) \\ \mathbb{E}(\ln \xi) &\leq \ln(\mathbb{E}(\xi)) \end{aligned}$$