PT LEC 03

isagila ablearthy

Собрано 22.12.2023 в 17:43

https://github.com/isagila/tesc

Содержание

Предисловие

Данные материалы основаны на видеозаписях лекций (2022 и 2021) и практик (2022).

1. Лекции

1.1. Лекция 23.09.05.

Пусть проводится n реальных экспериментов, в которых событие A появилось n_A раз. Отношение $\mathsf{P}(A) = \dfrac{n_A}{n}$ 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. $\Omega = \left\{(i, j) \,\middle|\, 1 \leqslant i, j \leqslant 6\right\}$. 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. $\Omega = \{\Gamma, P\Gamma, PP\Gamma, PPP\Gamma, \ldots\}$ — пространство элементарных исходов бесконечно, но счетно.
6. Монета бросается на координатную плоскость. $\Omega = \left\{(x, y) \,\middle|\, x, y \in \mathbb{R}\right\}$.

Операции над событиями

Выделим два специальных события. Ω — достоверное (универсальное) событие, которое наступает всегда, т.к. содержит все элементарные исходы. Пустое (невозможное) событие $\varnothing$, 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 $A_1 + \ldots + A_n$ — произошло хотя бы одно из этих событий. Произведение $A_1 \ldots A_n$ — произошли все события.

Def 1.1.7. Противоположным к A называется событие $\overline{A}$, 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. $A \setminus B = A\overline{B}$.

Def 1.1.9. События A и B называются несовместными, если из произведение равно $\varnothing$, 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 приписать числовую характеристику, отражающую частоту события A: $0 \leqslant \mathsf{P}(A) \leqslant 1$ — вероятность события A.

Подход I. Классическое определение вероятности

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

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

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

1. $0 \leqslant \mathsf{P}(A) \leqslant 1$
2. $\mathsf{P}(\Omega) = \dfrac{n}{n} = 1$
3. $\mathsf{P}(\varnothing) = \dfrac{0}{n} = 0$
4. Если A и B — несовместные события, то $\mathsf{P}(A+B) = \mathsf{P}(A) + \mathsf{P}(B)$.

Приведем несложное доказательство 4° свойства.

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

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

■

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

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

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.

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

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.

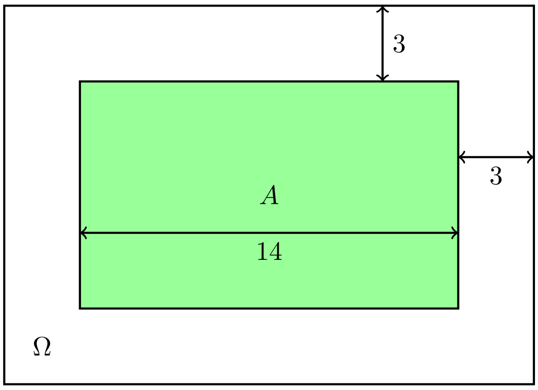
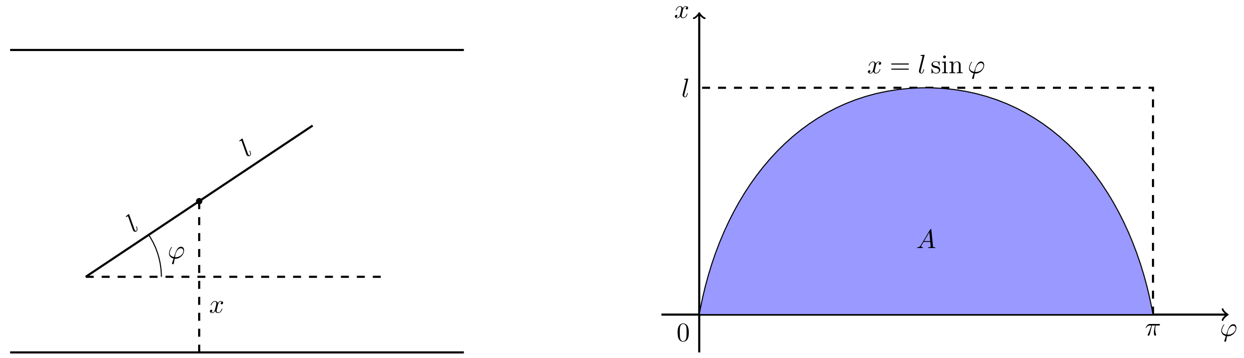


Рис. 1.1.14: Иллюстрация к 1.1.15

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

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



(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 $x \leqslant l \sin \varphi$. 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 $\mathsf{P}(A) = \dfrac{2l}{\pi l} = \dfrac{2}{\pi}$.

Статистика Максвелла–Больцмана

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

Статистика Бозе–Эйнштейна

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

Статистика Ферми–Дирака

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

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 $\mathfrak{F}$ подмножеств Ω называется σ-алгеброй событий, если выполнены следующие свойства

1. $\Omega \in \mathfrak{F}$
2. $A \in \mathfrak{F} \Longrightarrow \overline{A} \in \mathfrak{F}$ (замкнутость относительно дополнения)
3. $A_1, A_2, \ldots \in \mathfrak{F} \Longrightarrow \bigcup_{i=1}^\infty A_i \in \mathfrak{F}$ (замкнутость относительно счетного объединения)

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

$$A \in \mathfrak{F} \overset{2^\circ}{\Longrightarrow} \overline{A} \in \mathfrak{F} \overset{3^\circ}{\Longrightarrow} \underbrace{(A \cup \overline{A})}_{\Omega} \in \mathfrak{F}$$

Свойства σ-алгебры событий

Lm 1.2.3.

$$\varnothing \in \mathfrak{F}$$

□
$$\Omega \in \mathcal{F} \xRightarrow{2^\circ} \varnothing = \overline{\Omega} \in \mathcal{F}$$
■

<u>Lm</u> 1.2.4. σ-алгебра событий замкнута относительно счетного объединения.
$$A_1, A_2, \ldots \in \mathcal{F} \Longrightarrow \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$$

□ Пусть $A_1, A_2, \ldots \in \mathcal{F}$, тогда
$$A_1, A_2, \ldots \in \mathcal{F} \xRightarrow{2^\circ} \overline{A_1}, \overline{A_2}, \ldots \in \mathcal{F} \xRightarrow{3^\circ} \bigcup_{i=1}^{\infty} \overline{A_i} \in \mathcal{F} \xRightarrow{2^\circ} \overline{\bigcup_{i=1}^{\infty} \overline{A_i}} \in \mathcal{F}$$
Применяя закон де Моргана (который работает не только в конечном, но и в счетном случае), получаем искомое.
$$\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$$
■

<u>Lm</u> 1.2.5.
$$A, B \in \mathcal{F} \Longrightarrow A \setminus B \in \mathcal{F}$$

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

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

2. $\mathcal{F} = \left\{\varnothing, \Omega, A, \overline{A}\right\}$
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 $\mathsf{P}\colon \mathcal{F} \to \mathbb{R}$, обладающая следующими свойствами

1. Неотрицательность. $\forall A \in \mathcal{F} \,\Big|\, \mathsf{P}(A) \geqslant 0$
2. Счетная аддитивность. Если события $A_1, A_2, \ldots$ попарно несовместны, то $\mathsf{P}\left(\sum\limits_{i=1}^{\infty} A_i\right) = \sum\limits_{i=1}^{\infty} \mathsf{P}(A_i)$.
3. Нормированность. $\mathsf{P}(\Omega) = 1$

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

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

Свойства вероятности

<u>Lm</u> 1.2.10.
$$\mathsf{P}(\varnothing) = 0$$

□
$$1 = \mathsf{P}(\Omega) = \mathsf{P}(\varnothing + \Omega) = \mathsf{P}(\varnothing) + \mathsf{P}(\Omega) = \mathsf{P}(\varnothing) + 1 \Longrightarrow \mathsf{P}(\varnothing) = 0$$
■

<u>Lm</u> 1.2.11 (Формула обратной вероятности)**.**
$$\mathsf{P}(A) = 1 - \mathsf{P}\left(\overline{A}\right)$$

□ Т.к. A и $\overline{A}$ несовместны и $A + \overline{A} = \Omega$, то
$$1 = \mathsf{P}\left(A + \overline{A}\right) = \mathsf{P}(A) + \mathsf{P}\left(\overline{A}\right) \Longrightarrow \mathsf{P}(A) = 1 - \mathsf{P}\left(\overline{A}\right)$$
■

Lm 1.2.12.

$$0 \leqslant \mathsf{P}(A) \leqslant 1$$

□ $\mathsf{P}(A) \geqslant 0$ по первой аксиоме вероятности. Далее по 1.2.11 имеем, что $\mathsf{P}(A) = 1 - \mathsf{P}(\overline{A})$. Т.к. $\mathsf{P}(\overline{A}) \geqslant 0$, то $\mathsf{P}(A) \leqslant 1$. ■

Аксиома непрерывности

Пусть имеется убывающая цепочка событий $A_1 \supset A_2 \supset \ldots A_n \supset \ldots$ и $\bigcap\limits_{i=1}^{\infty} A_i = \varnothing$. Тогда $\mathsf{P}(A_n) \to 0$ при $n \to \infty$.

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

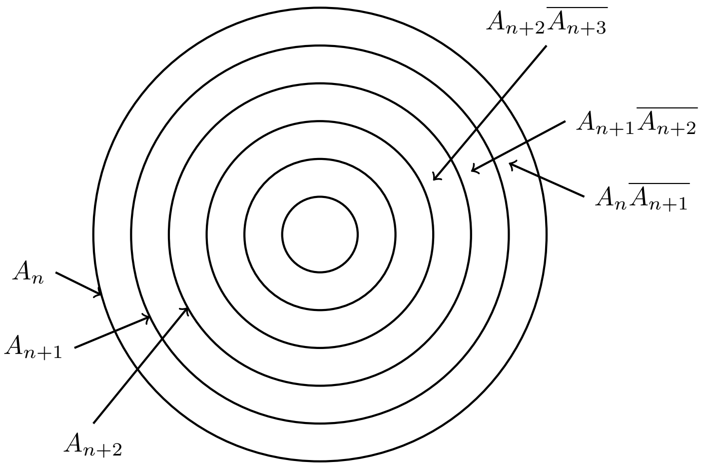


Рис. 1.2.14: Иллюстрация к 1.2.15

Теорема 1.2.15. Аксиома непрерывности следует из аксиомы счетной аддитивности.

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

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

Далее заметим, что

$$\bigcap_{i=n}^{\infty} A_i = A_n \cap \bigcap_{i=n+1}^{\infty} A_i = \bigcap_{i=1}^{n} A_i \cap \bigcap_{i=n+1}^{\infty} A_i = \bigcap_{i=1}^{\infty} A_i \overset{\text{усл}}{=\!=} \varnothing \tag{2}$$

Подставим (2) в (1) и получим, что

$$\begin{gathered} A_n = \sum_{i=n}^{\infty} A_i\overline{A_{i+1}} \\ A_1 = \sum_{i=1}^{\infty} A_i\overline{A_{i+1}} = \sum_{i=1}^{n} A_i\overline{A_{i+1}} + \underbrace{\textstyle\sum_{i=n}^{\infty} A_i\overline{A_{i+1}}}_{A_n} \end{gathered} \tag{3}$$

Далее воспользуемся аксиомой счетной аддитивности.

$$\mathsf{P}(A_1) = \sum_{i=1}^{n} \mathsf{P}\left(A_i\overline{A_{i+1}}\right) + \mathsf{P}(A_n) = \underbrace{\textstyle\sum_{i=1}^{\infty} \mathsf{P}\left(A_i\overline{A_{i+1}}\right)}_{S} \tag{4}$$

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

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 $\mathsf{P}(A+B) = \mathsf{P}(A) + \mathsf{P}(B)$ (это частный случай аксиомы $2°$).

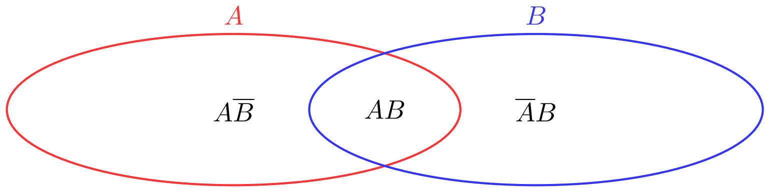


Рис. 1.2.17: Иллюстрация к 1.2.18

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

$$\mathsf{P}(A+B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(AB)$$

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

$$\begin{aligned}
\mathsf{P}(A+B) &= \mathsf{P}\left(A\overline{B}\right) + \mathsf{P}(AB) + \mathsf{P}\left(\overline{A}B\right) \\
&= \underbrace{\left(\mathsf{P}\left(A\overline{B}\right) + \mathsf{P}(AB)\right)}_{\mathsf{P}(A)} + \underbrace{\left(\mathsf{P}\left(\overline{A}B\right) + \mathsf{P}(AB)\right)}_{\mathsf{P}(B)} - \mathsf{P}(AB) \\
&= \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(AB)
\end{aligned}$$

■

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

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

При $n = 3$ формула усложняется.

$$\mathsf{P}(A_1 + A_2 + A_3) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \mathsf{P}(A_3) - \mathsf{P}(A_1A_2) - \mathsf{P}(A_1A_3) - \mathsf{P}(A_2A_3) + \mathsf{P}(A_1A_2A_3)$$

В общем же случае получаем следующую формулу

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

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

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

Далее вычислим число пар, троек и т.д.

$$\mathsf{P}(A_i) \to n = C_n^1 \qquad \mathsf{P}(A_iA_j) \to C_n^2 \qquad \mathsf{P}(A_iA_jA_k) \to C_n^3$$

Подставим это в формулу.

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

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

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

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

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

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

$$\mathsf{P}(A) = \mathsf{P}(AB) + \mathsf{P}\left(A\overline{B}\right) \Longrightarrow$$

$$\begin{aligned}\mathsf{P}\left(A\overline{B}\right) &= \mathsf{P}(A) - \mathsf{P}(AB)\\ &= \mathsf{P}(A) - \mathsf{P}(A)\,\mathsf{P}(B)\\ &= \mathsf{P}(A)\,(1 - \mathsf{P}(B))\\ &= \mathsf{P}(A)\,\mathsf{P}\left(\overline{B}\right)\end{aligned}$$

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

Def 1.2.23. События $A_1, \ldots, A_n$ 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».

$$\mathsf{P}(A) = \mathsf{P}(B) = \mathsf{P}(C) = \frac{2}{4} = \frac{1}{2}$$
$$\mathsf{P}(AB) = \mathsf{P}(AC) = \mathsf{P}(BC) = \frac{1}{4}$$

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

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

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 $(1 \leqslant i \leqslant 4)$ выпала шестерка. Обозначим B событие «выпала хотя бы одна шестерка», тогда $B = A_1 + A_2 + A_3 + A_4$. Получаем, что $\overline{B}$ это событие «не выпала ни одна шестерка», тогда $\overline{B} = \overline{A_1} \cdot \overline{A_2} \cdot \overline{A_3} \cdot \overline{A_4}$. Т.к. броски кубиков независимые, то

$$\mathsf{P}\left(\overline{B}\right) = \mathsf{P}\left(\overline{A_1}\right)\mathsf{P}\left(\overline{A_2}\right)\mathsf{P}\left(\overline{A_3}\right)\mathsf{P}\left(\overline{A_4}\right) = \left(\frac{5}{6}\right)^4 \approx 0.48$$

Значит, $\mathsf{P}(B) = 1 - \mathsf{P}\left(\overline{B}\right) \approx 0.52$.

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 $\mathsf{P}\left(A \,\middle|\, B\right)$ это вероятность события A вычисленная в предположении, что событие B уже произошло.

Замечание 1.3.1. Существует другой обозначение $\mathsf{P}_B(A)$, 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

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

Тогда получаем, что

$$\mathsf{P}\left(A \,\middle|\, B\right) = \frac{2}{3} = \frac{\frac{2}{6}}{\frac{3}{6}} = \frac{\mathsf{P}(AB)}{\mathsf{P}(B)}$$

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

$$\mathsf{P}\left(A \,\middle|\, B\right) = \frac{S_{AB}}{S_B} = \frac{\frac{S_{AB}}{S_\Omega}}{\frac{S_B}{S_\Omega}} = \frac{\mathsf{P}\left(AB\right)}{\mathsf{P}\left(B\right)}$$

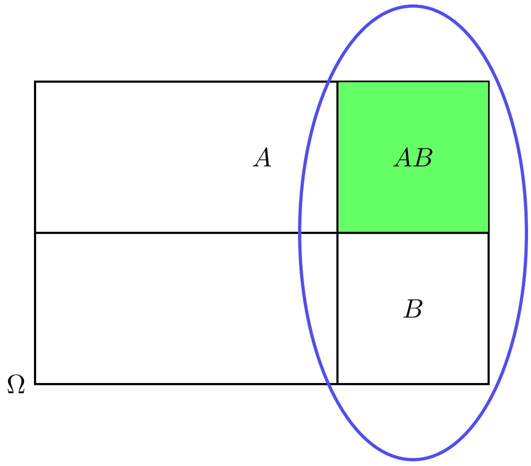


Рис. 1.3.3: Геометрический подход к условной вероятности

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

$$\mathsf{P}\left(A \,\middle|\, B\right) = \frac{\mathsf{P}\left(AB\right)}{\mathsf{P}\left(B\right)}$$

Формула произведения вероятностей

Из определения условной вероятности (1.3.5) следует, что

$$\mathsf{P}\left(AB\right) = \mathsf{P}\left(B\right)\mathsf{P}\left(A \,\middle|\, B\right) = \mathsf{P}\left(A\right)\mathsf{P}\left(B \,\middle|\, A\right)$$

Или, в общем случае, имеем

Теорема 1.3.6. (Формула произведения вероятностей)

$$\mathsf{P}\left(A_1 \dots A_n\right) = \mathsf{P}\left(A_1\right)\mathsf{P}\left(A_2 \,\middle|\, A_1\right)\mathsf{P}\left(A_3 \,\middle|\, A_1A_2\right)\dots\mathsf{P}\left(A_n \,\middle|\, A_1 \dots A_{n-1}\right)$$

□ Мат. индукция.
База: $n = 2$.
Этот случай уже рассмотрен ранее, он следует из определения условной вероятности.
Переход: $n - 1 \to n$

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

■

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

<u>Lm</u> 1.3.8. Покажем, что равносильны два определения независимости: $\mathsf{P}\left(AB\right) = \mathsf{P}\left(A\right)\mathsf{P}\left(B\right)$ и $\mathsf{P}\left(A \,\middle|\, B\right) = \mathsf{P}\left(A\right)$

□

$$\mathsf{P}\left(A \,\middle|\, B\right) = \frac{\mathsf{P}\left(AB\right)}{\mathsf{P}\left(B\right)} = \mathsf{P}\left(A\right) \Longleftrightarrow \mathsf{P}\left(AB\right) = \mathsf{P}\left(A\right)\mathsf{P}\left(B\right)$$

■

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».

$$\mathsf{P}\left(A_1A_2A_3\right) = \mathsf{P}\left(A_1\right)\mathsf{P}\left(A_2 \,\middle|\, A_1\right)\mathsf{P}\left(A_3 \,\middle|\, A_1A_2\right) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} = \frac{1}{5}$$

Полная группа событий

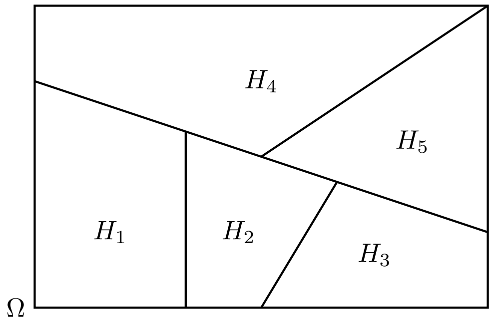


Рис. 1.3.10: Иллюстрация полной группы событий

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

$$\begin{cases} \forall i \neq j \,\Big|\, H_i H_j = \varnothing \\ H_1 + \ldots + H_n = \Omega \end{cases}$$

Следствие 1.3.12.

$$\mathsf{P}\left(H_1\right) + \ldots + \mathsf{P}\left(H_n\right) = 1$$

□ Т.к. события $H_1, \ldots, H_n$ 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 $H_1, \ldots, H_n, \ldots$ — полная группа событий. Тогда

$$\mathsf{P}\left(A\right) = \sum_{i=1}^{\infty} \mathsf{P}\left(H_i\right) \mathsf{P}\left(A \,\middle|\, H_i\right)$$

□

$$\mathsf{P}\left(A\right) = \mathsf{P}\left(\Omega A\right) = \mathsf{P}\left(\left(H_1 + \ldots + H_n + \ldots\right) A\right) = \mathsf{P}\left(H_1 A + \ldots + H_n A + \ldots\right)$$

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

$$\mathsf{P}\left(A\right) = \sum_{i=1}^{\infty} \mathsf{P}\left(H_i A\right) = \sum_{i=1}^{\infty} \mathsf{P}\left(H_i\right) \mathsf{P}\left(A \,\middle|\, H_i\right)$$

■

Формула Байеса (формула проверки гипотез)

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

$$\mathsf{P}\left(H_k \,\middle|\, A\right) = \frac{\mathsf{P}\left(H_k\right) \mathsf{P}\left(A \,\middle|\, H_k\right)}{\sum_{i=1}^{\infty} \mathsf{P}\left(H_i\right) \mathsf{P}\left(A \,\middle|\, H_i\right)}$$

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

$$\mathsf{P}\left(H_k \,\middle|\, A\right) = \frac{\mathsf{P}\left(H_k A\right)}{\mathsf{P}\left(A\right)} = \frac{\mathsf{P}\left(H_k\right) \mathsf{P}\left(A \,\middle|\, H_k\right)}{\sum_{i=1}^{\infty} \mathsf{P}\left(H_i\right) \mathsf{P}\left(A \,\middle|\, H_i\right)}$$

■

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

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

$$\mathsf{P}\left(A \,\middle|\, H_1\right) = \tfrac{3}{5} \quad \mathsf{P}\left(A \,\middle|\, H_2\right) = \tfrac{1}{5} \quad \mathsf{P}\left(A \,\middle|\, H_3\right) = \tfrac{2}{5}$$

По формуле полной вероятности получаем

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

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

$$\begin{array}{ll} \mathsf{P}(H_1) = 0.01 & \mathsf{P}(H_2) = 0.99 \\ \mathsf{P}\left(A \,\middle|\, H_1\right) = 0.99 & \mathsf{P}\left(A \,\middle|\, H_2\right) = 0.01 \end{array}$$

Применяем формулу Байеса.

$$\mathsf{P}\left(H_1 \,\middle|\, A\right) = \frac{\mathsf{P}(H_1)\,\mathsf{P}\left(A \,\middle|\, H_1\right)}{\mathsf{P}(H_1)\,\mathsf{P}\left(A \,\middle|\, H_1\right) + \mathsf{P}(H_2)\,\mathsf{P}\left(A \,\middle|\, H_2\right)} = \frac{0.01\cdot 0.99}{0.01\cdot 0.99 + 0.99\cdot 0.01} = \frac{1}{2}$$

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

$$\begin{array}{ll} \mathsf{P}(H_1) = 0.01 & \mathsf{P}(H_2) = 0.99 \\ \mathsf{P}\left(AA \,\middle|\, H_1\right) = 0.99\cdot 0.99 & \mathsf{P}\left(AA \,\middle|\, H_2\right) = 0.01\cdot 0.01 \end{array}$$

Применяем формулу Байеса.

$$\mathsf{P}\left(H_1 \,\middle|\, AA\right) = \frac{\mathsf{P}(H_1)\,\mathsf{P}\left(AA \,\middle|\, H_1\right)}{\mathsf{P}(H_1)\,\mathsf{P}\left(AA \,\middle|\, H_1\right) + \mathsf{P}(H_2)\,\mathsf{P}\left(AA \,\middle|\, H_2\right)} = \frac{0.01\cdot 0.99^2}{0.01\cdot 0.99^2 + 0.99\cdot 0.01^2} = 0.99$$

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

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

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

$$A_1 = \underbrace{Y \dots Y}_{k}\underbrace{N \dots N}_{n-k} \qquad \mathsf{P}\left(Y\right) = p, \mathsf{P}\left(N\right) = q$$

Т.к. испытания независимы, то

$$\mathsf{P}\left(A_1\right) = p^k q^{n-k}$$

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. $\blacksquare$

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

$$\mathsf{P}_5(3) = C_5^3 \cdot 0.8^3 \cdot 0.2^2 = 0.2048$$

Наиболее вероятное число успехов

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

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

Т.к. k целое, то получаем, что $np + p - 1 \leqslant k \leqslant np + p$. 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 $k = np + p - 1$.

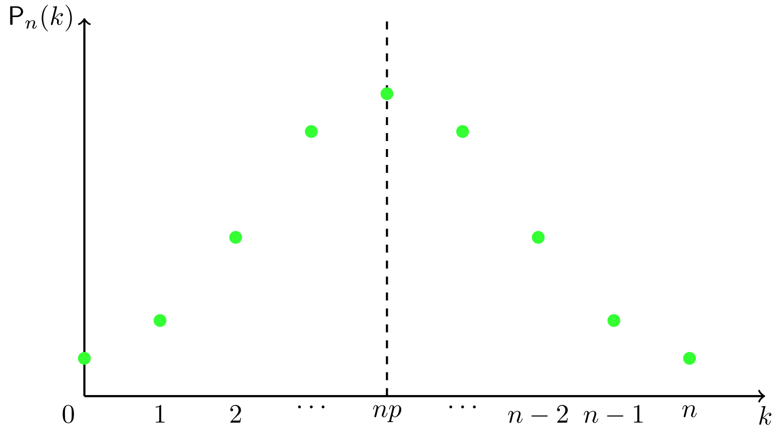


Рис. 1.4.4: Геометрическая иллюстрация наиболее вероятного числа успехов

Если рассмотреть рисунок рис. 1.4.4 и устремить $k \to \infty$, 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).

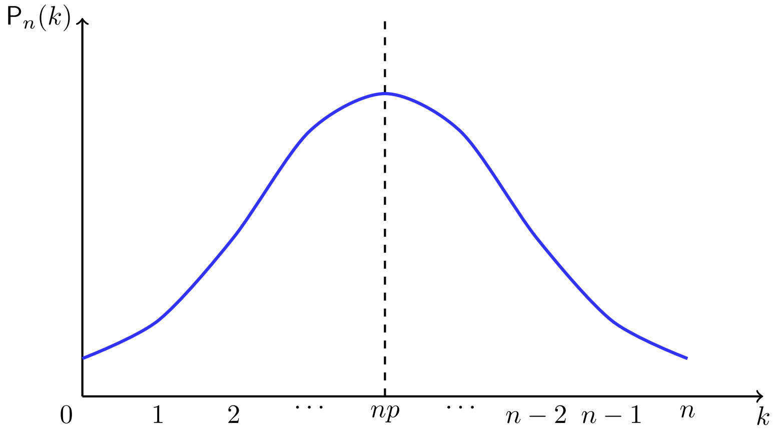


Рис. 1.4.5: Нормальное распределение

Предельные теоремы в схеме Бернулли

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

Теорема 1.4.6. (Локальная теорема Муавра—Лапласа)

$$\mathsf{P}_n(k) \xrightarrow{n\to\infty} \frac{1}{\sqrt{npq}}\varphi(x) \qquad \begin{cases} \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \\ x = \frac{k-np}{\sqrt{npq}} \end{cases}$$

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

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

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

Теорема 1.4.8. (Интегральная теорема Муавра—Лапласа)

$$\mathsf{P}_n(k_1 \leqslant k \leqslant k_2) \xrightarrow{n\to\infty} \Phi(x_2) - \Phi(x_1) \qquad \begin{cases} \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} \mathrm{d}z \\ x_1 = \frac{k_1 - np}{\sqrt{npq}} \\ x_2 = \frac{k_2 - np}{\sqrt{npq}} \end{cases}$$

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

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

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

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

Это функция стандартного нормального распределения.

Замечание 1.4.11. Интеграл Пуассона

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

Используя его можно получить, что

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

Замечание 1.4.12. Эти формулы (1.4.6 и 1.4.8) обычно применяются при $n \geqslant 100$ 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

$$x = \frac{k - np}{\sqrt{npq}} = \frac{330 - 400 \cdot 0.8}{\sqrt{400 \cdot 0.8 \cdot 0.2}} = \frac{330 - 320}{\sqrt{64}} = 1.25$$
$$\mathsf{P}_{400}(330) \approx \frac{1}{\sqrt{npq}}\varphi(x) = \frac{1}{8}\varphi(1.25) = 0.125 \cdot 0.1826 \approx 0.0228$$

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

$$x_1 = \frac{k_1 - np}{\sqrt{npq}} = \frac{312 - 400 \cdot 0.8}{\sqrt{400 \cdot 0.8 \cdot 0.2}} = \frac{312 - 320}{\sqrt{64}} = -1 \qquad x_2 = \frac{k_2 - np}{\sqrt{npq}} = \frac{336 - 400 \cdot 0.8}{\sqrt{400 \cdot 0.8 \cdot 0.2}} = \frac{336 - 320}{\sqrt{64}} = 2$$
$$\mathsf{P}_{400}(312 \leqslant k \leqslant 336) \approx \Phi(2) - \Phi(-1) = \Phi(2) + \Phi(1) = 0.4772 + 0.3413 \approx 0.8185$$

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 $\frac{n_A}{n} \approx \mathsf{P}(A) = p$.

Вероятность отклонения относительной частоты от вероятности события

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

$$\mathsf{P}\left(\left|\frac{n_A}{n} - p\right| < \varepsilon\right) = \mathsf{P}\left(-n\varepsilon < n_A - np < n\varepsilon\right) = \mathsf{P}\left(np - n\varepsilon < n_A < np + n\varepsilon\right)$$
$$x_1 = \frac{np - n\varepsilon - np}{\sqrt{npq}} = -\frac{n\varepsilon}{\sqrt{npq}} \qquad x_2 = \frac{np + n\varepsilon - np}{\sqrt{npq}}$$
$$\mathsf{P}\left(\left|\frac{n_A}{n} - p\right| < \varepsilon\right) \approx \Phi\left(\frac{n\varepsilon}{\sqrt{npq}}\right) - \Phi\left(-\frac{n\varepsilon}{\sqrt{npq}}\right) = \Phi\left(\frac{n\varepsilon}{\sqrt{npq}}\right) + \Phi\left(\frac{n\varepsilon}{\sqrt{npq}}\right) = 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right)$$

Итак, получили следующую формулу

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

Закон больших чисел Бернулли

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

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

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, $\mathsf{P}(A_i) = \frac{3}{40}$. Получаем

$$\mathsf{P}\left(A_{12} \,\middle|\, \overline{A_1} \dots \overline{A_{11}}\right) = \frac{\mathsf{P}\left(A_{12}\overline{A_1} \dots \overline{A_{11}}\right)}{\mathsf{P}\left(\overline{A_1} \dots \overline{A_{11}}\right)}$$
$$\mathsf{P}\left(\overline{A_1} \dots \overline{A_{11}}\right) = 1 - \mathsf{P}\left(A_1 + \dots + A_{11}\right) = 1 - \left(\mathsf{P}(A_1) + \dots + \mathsf{P}(A_{11})\right) = \frac{7}{40}$$
$$\mathsf{P}\left(A_{12}\overline{A_1} \dots \overline{A_{11}}\right) = \mathsf{P}(A_{12}) = \frac{3}{40}$$
$$\mathsf{P}\left(A_{12} \,\middle|\, \overline{A_1} \dots \overline{A_{11}}\right) = \frac{3}{7}$$

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

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

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

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

$$\mathsf{P}_n(Y_n = k) = C_n^k p^k q^{n-k} \qquad 0 \leqslant k \leqslant n$$

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

Схема II. Схема до первого успешного испытания

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

<u>Lm</u> 1.5.2.

$$\mathsf{P}\left(\tau = k\right) = q^{k-1}p \qquad k = 0, 1, \ldots$$

□

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

■

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

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

Теорема 1.5.4.

$$\mathsf{P}\left(\tau = k\right) = q^{k-1}p, k \in \mathbb{N} \Longrightarrow \forall n, k \geqslant 0 \,\Big|\, \mathsf{P}\left(\tau > n + k \,\middle|\, \tau > n\right) = \mathsf{P}\left(\tau > k\right)$$

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

$$\mathsf{P}\left(\tau > n + k \,\middle|\, \tau > n\right) = \frac{\mathsf{P}\left(\tau > n + k, \tau > n\right)}{\mathsf{P}\left(\tau > n\right)} = \frac{q^{n+k}}{q^n} = q^k = \mathsf{P}\left(\tau > k\right)$$

■

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

$$\mathsf{P}_n(n_1, \ldots, n_m) = \frac{n!}{n_1! \cdot \ldots n_m!} \cdot p_1^{n_1} \cdot \ldots \cdot p_m^{n_m}$$

□ Рассмотрим один элементарный исход благоприятный данному событию.

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

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

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

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

$$\mathsf{P}_6(1, 2, 3) = \frac{6!}{1! \cdot 2! \cdot 3!} \cdot \left(\frac{1}{4}\right)^1 \cdot \left(\frac{1}{4}\right)^2 \cdot \left(\frac{1}{2}\right)^3 = \frac{15}{128} \approx 0.12$$

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.

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

Схема IV. (b) Без возврата

$$\mathsf{P}_{N,K}(n, k) = \frac{C_K^k \cdot C_{N-K}^{n-k}}{C_N^n} \qquad 0 \leqslant k \leqslant K$$

Def 1.5.8. Соответствие $k \to \dfrac{C_K^k \cdot C_{N-K}^{n-k}}{C_N^n}$, $0 \leqslant k \leqslant K$ 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 результате по схеме (b) будут стремиться к схеме (a), а гипергеометрическое распределение будет приближаться к биномиальному.

Lm 1.5.9.

$$C_K^k \sim \frac{K^k}{k!} \qquad K \to \infty, k = const$$

□

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

■

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

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

$\square$ 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 искомую формулу. $\blacksquare$

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 $n \to \infty$, тогда $p_n \to 0$ так, что $n \cdot p_n \to \lambda > 0$. Тогда вероятность k успехов при n испытаниях будет равна

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

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

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

Далее применим лемму 1.5.9.

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

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

$$\left(1 - \frac{\lambda_n}{n}\right)^n = \left(1 - \frac{\lambda_n}{n}\right)^{\frac{-n}{\lambda_n}\cdot(-\lambda_n)} \xrightarrow{n\to\infty} e^{-\lambda_n}$$

$$\mathsf{P}(Y_n = k) \xrightarrow{n\to\infty} \frac{\lambda_n^k}{k!} \cdot e^{-\lambda_n} \xrightarrow{n\to\infty} \frac{\lambda^k}{k!} \cdot e^{-\lambda}$$

$\blacksquare$

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

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

Def 1.5.13. Соответствие $k \to \dfrac{\lambda^k}{k!} \cdot e^{-\lambda}$, $k \in \mathbb{N}_0$ 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.

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

Применим 1.5.12, чтобы оценить погрешность.

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

1.6. Лекция 23.10.10.

Случайные величины

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

$$\forall x \in \mathbb{R} \,\Big|\, \Big\{\omega \in \Omega \,\Big|\, \xi(\omega) < x\Big\} \in \mathcal{F}$$

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

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

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

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

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 $\mathsf{P}\left(\xi \in (-\infty; x)\right) = \mathsf{P}\Big(\omega \in \Omega \,\Big|\, \xi(\omega) < x\Big)$. 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 $\mathsf{P}(B) = \mathsf{P}\Big(\Big\{\omega \in \Omega \,\Big|\, \xi(\omega) \in B\Big\}\Big)$.
Итак, пусть ξ — случайная величина, заданная на вероятностном пространстве $\langle\Omega, \mathcal{F}, \mathsf{P}\rangle$. Тогда получаем новое вероятностное пространство:

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

Полученное вероятностное пространство называется индуцированным с помощью случайной величины ξ.

Def 1.6.4. Распределением называется функция $\mathsf{P}(B)$, 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 распределения, если она принимает не более, чем счетное число значений.

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

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

ξ	x_1	x_2	$\ldots$	x_n	$\ldots$
p	p_1	p_2	$\ldots$	p_n	$\ldots$

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

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

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.

$$\mathbb{E}(\xi) = \sum_{i=1}^{6} x_i p_i = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5$$

$$\mathbb{D}(\xi) = \sum_{i=1}^{6} x_i^2 p_i - (\mathbb{E}(\xi))^2 = 1 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 9 \cdot \frac{1}{6} + 16 \cdot \frac{1}{6} + 25 \cdot \frac{1}{6} + 36 \cdot \frac{1}{6} - 3.5^2 \approx 2.92$$

$$\sigma(\xi) \approx \sqrt{2.92} \approx 1.71$$

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

I_A	0	1
p	$1 - \mathsf{P}(A)$	$\mathsf{P}(A)$

Тогда

$$\mathbb{E}(\xi) = 0 \cdot (1 - \mathsf{P}(A)) + 1 \cdot \mathsf{P}(A) = \mathsf{P}(A)$$

$$\mathbb{D}(\xi) = 0^2 \cdot (1 - \mathsf{P}(A)) + 1^2 \cdot \mathsf{P}(A) - \mathsf{P}(A)^2 = \mathsf{P}(A)(1 - \mathsf{P}(A)) = pq$$

Свойства математического ожидания и дисперсии

Def 1.6.15. Случайная величина ξ имеет вырожденное распределение, если

$$\forall \omega \in \Omega \,\Big|\, \xi(\omega) = C \text{ или } \mathsf{P}(\xi = C) = 1$$

Тогда

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

Lm 1.6.16.

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

□

$$\mathbb{E}(\xi + C) = \sum_{i=1}^{\infty}(x_i + C) \cdot p_i = \sum_{i=1}^{\infty} x_i p_i + C \cdot \underbrace{\textstyle\sum_{i=1}^{\infty} p_i}_{=1} = \mathbb{E}(\xi) + C$$

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

■

Lm 1.6.17.

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

□

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

$$\mathbb{D}(C\xi) = \mathbb{E}\left((C\xi - \mathbb{E}(C\xi))^2\right) = \mathbb{E}\left(\left(C(\xi - \mathbb{E}(\xi))\right)^2\right) = C^2 \cdot \mathbb{E}\left((\xi - \mathbb{E}(\xi))^2\right) = C^2\mathbb{D}(\xi)$$

■

Lm 1.6.18.

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

□

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

Применим формулу полной вероятности к внутренним суммам.

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

■

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

$$\forall i, j \;\Big|\; \mathsf{P}(\xi = x_i, \eta = y_j) = \mathsf{P}(\xi = x_i) \cdot \mathsf{P}(\eta = y_j)$$

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

Lm 1.6.20. Если ξ и η независимы, то

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

□

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

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

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

■

Lm 1.6.21.

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

□

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

■

Lm 1.6.22.

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

□

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

■

<u>Lm</u> 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 $\operatorname{cov}(\xi,\eta) = 0$ 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

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

Def 1.6.25. Центральный момент k-ого порядка это

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

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.

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

Замечание 1.6.28. Задачу 1.1.17 можно решить, используя математическое ожидание. Пусть $p = \mathsf{P}(A)$ — 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

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

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.

$$\xi \in B_{n,p} \Longleftrightarrow \forall 0 \leqslant k \leqslant n, k \in \mathbb{Z} \Big| \mathsf{P}(\xi = k) = C_n^k p^k q^{n-k}$$

Закон распределения имеет вид:

ξ	0	1	$\dots$	k	$\dots$	n
p	q^n	$C_n^1 pq^{n-1}$	$\dots$	$C_n^k p^k q^{n-k}$	$\dots$	p^n

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

$$\mathbb{E}(\xi) = \mathbb{E}(\xi_1) + \ldots + \mathbb{E}(\xi_n) = np$$
$$\mathbb{D}(\xi) = \mathbb{D}(\xi_1) + \ldots + \mathbb{D}(\xi_n) = npq$$
$$\sigma(\xi) = \sqrt{npq}$$

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

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

$$\xi \in G_p \Longleftrightarrow \forall k \in \mathbb{N} \Big| \mathsf{P}(\xi = k) = q^{k-1}p$$

Закон распределения будет иметь вид:

ξ	1	2	$\dots$	k	$\dots$
p	p	pq	$\dots$	$q^{k-1}p$	$\dots$

Найдем числовые характеристики.

$$\mathbb{E}(\xi) = \sum_{k=1}^{\infty} k \cdot q^{k-1}p = p\sum_{k=1}^{\infty}\left(q^k\right)' = p\left(\sum_{k=1}^{\infty} q^k\right)' = p\left(\frac{1}{1-q}\right)' = p \cdot \frac{1}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}$$

$$\mathbb{E}\left(\xi^2\right) = \sum_{k=1}^{\infty} k^2 \cdot q^{k-1}p = \sum_{k=1}^{\infty}(k(k-1)+k)\cdot q^{k-1}p = \underbrace{\textstyle\sum_{k=1}^{\infty} k(k-1)\cdot q^{k-1}p}_{\text{ниже}} + \underbrace{\textstyle\sum_{k=1}^{\infty} k \cdot q^{k-1}p}_{\mathbb{E}(\xi)}$$

$$\sum_{k=1}^{\infty} k(k-1)\cdot q^{k-1}p = pq\sum_{k=1}^{\infty}\left(q^k\right)'' = pq\left(\sum_{k=1}^{\infty} q^k\right)'' = pq\left(\frac{1}{1-q}\right)'' = pq \cdot \frac{2}{(1-q)^3} = \frac{2q}{p^2}$$

$$\mathbb{E}\left(\xi^2\right) = \frac{2q}{p^2} + \frac{1}{p} = \frac{2q+p}{p^2} = \frac{q+1}{p^2}$$

$$\mathbb{D}(\xi) = \mathbb{E}\left(\xi^2\right) - (\mathbb{E}(\xi))^2 = \frac{q+1}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{q}{p^2}$$

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

$$\forall k \geqslant 0, k \in \mathbb{Z} \Big| \mathsf{P}(\xi = k) = \frac{\lambda^k}{k!}e^{-\lambda}$$

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

Вычислим числовые характеристики.

$$\mathbb{E}(\xi) = \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}\sum_{n=0}^{\infty}\frac{\lambda^n}{n!} = \lambda e^{-\lambda}\cdot e^{\lambda} = \lambda$$

$$\mathbb{E}\left(\xi^2\right) = \sum_{k=0}^{\infty} k^2\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=0}^{\infty}\frac{k(k-1)+k}{k!}\cdot\lambda^k e^{-\lambda} = \lambda^2 e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!} + \underbrace{\textstyle\sum_{k=0}^{\infty}\frac{k}{k!}\cdot\lambda^k e^{-\lambda}}_{\mathbb{E}(\xi)} = \lambda^2 e^{-\lambda}e^{\lambda} + \lambda = \lambda^2 + \lambda$$

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

Функция распределения

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

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

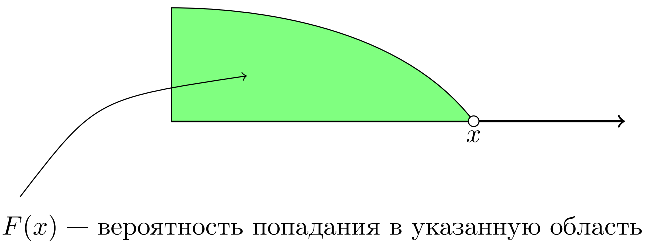


Рис. 1.7.3: Функция распределения

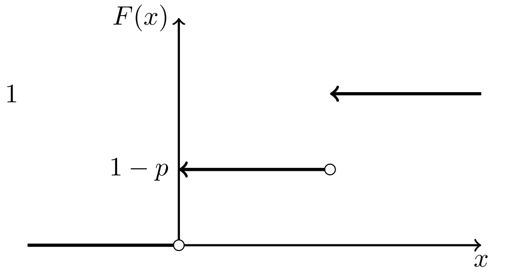


Рис. 1.7.4: Функция распределения Бернулли

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

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

Свойства функции распределения

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

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

Lm 1.7.7. $F(x)$ — неубывающая функция. $x_1 < x_2 \Longrightarrow F(x_1) \leqslant F(x_2)$

□

$$x_1 < x_2 \Longrightarrow \{\xi \in X_1\} \subset \{\xi \in X_2\} \Longrightarrow F(x_1) = \mathsf{P}\left(\xi < x_1\right) \leqslant \mathsf{P}\left(\xi < x_2\right) = F(x_2)$$

■

Lm 1.7.8.

$$\mathsf{P}\left(\alpha \leqslant \xi < \beta\right) = F(\beta) - F(\alpha)$$

□ Воспользуемся свойством аддитивности.

$$\mathsf{P}\left(\xi < \beta\right) = \mathsf{P}\left(\xi < \alpha\right) + \mathsf{P}\left(\alpha \leqslant \xi < \beta\right)$$
$$F(\beta) = F(\alpha) + \mathsf{P}\left(\alpha \leqslant \xi < \beta\right) \Longrightarrow \mathsf{P}\left(\alpha \leqslant \xi < \beta\right) = F(\beta) - F(\alpha)$$

■

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.

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

□ 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 $\{x_n\} \to \pm\infty$. Рассмотрим $A_n = \{n-1 \leqslant \xi_n < n\}, n \in \mathbb{Z}$ — несовместные события при разных n. По свойству счетной аддитивности получаем

$$1 = \mathsf{P}\left(\Omega\right) = \mathsf{P}\left(\sum_{n=-\infty}^{\infty} A_n\right) = \sum_{n=-\infty}^{\infty} \mathsf{P}\left(A_n\right) \overset{1.7.8}{=\!=\!=} \sum_{n=-\infty}^{\infty} (F(n) - F(n-1)) = \lim_{N\to\infty} \sum_{n=-N}^{N} (F(n) - F(n-1))$$

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

$$1 = \lim_{N\to\infty} (F(N) - F(-N-1)) = \lim_{N\to\infty} F(N) - \lim_{N\to\infty} F(-N-1) \Longrightarrow \lim_{N\to\infty} F(N) = 1 + \lim_{N\to\infty} F(-N-1)$$

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. $F(x)$ непрерывная слева, т.е.

$$\forall x_0 \in \mathbb{R} \,\Big|\, F(x_0 - 0) = F(x_0)$$

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

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

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

■

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

$$F(x_0 + 0) = F(x_0) + \mathsf{P}\left(\xi = x_0\right) = \mathsf{P}\left(\xi \leqslant x_0\right)$$

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

$$\begin{aligned}
\mathsf{P}\left(x_0 < \xi < x_0 + \frac{1}{n}\right) + \mathsf{P}\left(\xi = x_0\right) &\xrightarrow{n\to\infty} 0 + \mathsf{P}\left(\xi = x_0\right) \\
\mathsf{P}\left(x_0 \leqslant \xi < x_0 + \frac{1}{n}\right) &\xrightarrow{n\to\infty} \mathsf{P}\left(\xi = x_0\right) \\
F\left(x_0 + \frac{1}{n}\right) - F(x_0) &\xrightarrow{n\to\infty} \mathsf{P}\left(\xi = x_0\right)
\end{aligned}$$

■

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

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

$$\mathsf{P}\left(\alpha \leqslant \xi < \beta\right) = \mathsf{P}\left(\alpha < \xi < \beta\right) = \mathsf{P}\left(\alpha \leqslant \xi \leqslant \beta\right) F(\beta) - F(\alpha)$$

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

$$\forall B \in \mathcal{B}(\mathbb{R}) \,\Big|\, \mathsf{P}\left(\xi \in B\right) = \int_B f_\xi(x)\mathrm{d}x$$

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

Свойства плотности и функции распределения абсолютно непрерывного распределения

$$B = [\alpha;\beta] \Longrightarrow \mathsf{P}\left(\alpha \leqslant \xi \leqslant \beta\right) = \int_\alpha^\beta f_\xi(x)\mathrm{d}x$$

Используя геометрический смысл определенного интеграла, получаем, что $\mathsf{P}\left(\alpha \leqslant \xi \leqslant \beta\right)$ 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.

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

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

$$\mathsf{P}\left(\xi \in \mathbb{R}\right) = \int_{-\infty}^{\infty} f_\xi(x)\mathrm{d}x = 1$$

■

<u>Lm</u> 1.7.18.

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

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

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

■

<u>Lm</u> 1.7.19. $F_\xi(x)$ — абсолютно непрерывная функция.

□ Это свойство непрерывности интеграла с переменным верхним пределом. ■

<u>Lm</u> 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. ■

<u>Lm</u> 1.7.21. $f_\xi(x) \geqslant 0$

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

<u>Lm</u> 1.7.22.

$$\forall x_0 \in \mathbb{R} \,\Big|\, \mathsf{P}(\xi = x_0) = 0$$

□ 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. ■

<u>Lm</u> 1.7.23.

$$\mathsf{P}(\alpha \leqslant \xi < \beta) = \mathsf{P}(\alpha \leqslant \xi \leqslant \beta) = \mathsf{P}(\alpha < \xi < \beta) = \mathsf{P}(\alpha < \xi \leqslant \beta) = F(\beta) - F(\alpha)$$

□ Это следствие из 1.7.22. ■

Теорема 1.7.24.

$$\left.\begin{aligned} &\forall x \in \mathbb{R} \,\Big|\, f(x) \geqslant 0 \\ &\textstyle\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1 \end{aligned}\right\} \Longrightarrow f(x) \text{ это плотность некоторого распределения}$$

Независимость случайных величин

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

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

Def 1.7.26. Случайные величины $\xi_1, \ldots \xi_n$ называются независимыми, если

$$\forall B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R}) \,\Big|\, \mathsf{P}(\xi_1 \in B_1, \ldots, \xi_n \in B_n) = \mathsf{P}(\xi_1 \in B_1) \ldots \mathsf{P}(\xi_n \in B_n)$$

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

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

при условии, что данный интеграл сходится абсолютно.

Замечание 1.7.30.

$$\int_{-\infty}^{\infty} |x|\, f_\xi(x)\mathrm{d}x = \infty \Longrightarrow \text{математическое ожидание не существует}$$

II. Дисперсия

Def 1.7.31. Дисперсией абсолютно непрерывной случайной величины ξ называется число

$$\mathbb{D}(\xi) = \mathbb{E}\left((\xi - \mathbb{E}(\xi))^2\right) = \int_{-\infty}^{\infty} (x - \mathbb{E}(\xi))^2 f_\xi(x)\mathrm{d}x$$

при условии, что данный интеграл сходится.

Замечание 1.7.32. Для вычисления дисперсии удобнее формула

$$\mathbb{D}(\xi) = \mathbb{E}\left(\xi^2\right) - (\mathbb{E}(\xi))^2 = \int_{-\infty}^{\infty} x^2 f_\xi(x)\mathrm{d}x - (\mathbb{E}(\xi))^2$$

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

$$m_k = \mathbb{E}\left(\xi^k\right) = \int_{-\infty}^{\infty} x^k f_\xi(x)\mathrm{d}x$$

Def 1.7.36. Центральный момент k-ого порядка это

$$\mu_k = \mathbb{E}\left((\xi - \mathbb{E}(\xi))^k\right) = \int_{-\infty}^{\infty} (x - \mathbb{E}(\xi))^k f_\xi(x)\mathrm{d}x$$

V Медиана

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

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

VI Мода

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

Элементы теории принятия решений

Пусть требуется принять одно из возможных решений $A_1, A_2, \ldots, A_m$. Результат принятия данного решения зависит от одной из возможных ситуаций $Q_1, Q_2, \ldots, Q_n$. Обозначим a_{ij} — результат при принятии решения A_i и случившейся ситуации Q_j. Матрица из этих чисел — платёжная матрица:

$$A = (a_{ij}) = \begin{array}{c} \begin{matrix} p_1 & p_2 & \ldots & p_n \end{matrix} \\ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 1 & 2 & 3 \\ 4 & 5 & 6 & 7 \end{pmatrix} \end{array}$$

Если вероятности P_j возможных ситуаций Q_j неизвестны, то говорят, что решение принимается в *условиях полной неопределённости.*

Если эти вероятности известны, то говорят, что решение принимается в *условиях рисках.*

Принципы принятия решений

I. Байесовский (максимизация прибыли)

В качестве оптимального выбирают то решение, где среднее ожидание прибыли $\overline{a}_i = \sum_{j=1}^{\infty} a_{ij}p_j$ является наибольшим.

	$\frac{1}{2}$	$\frac{1}{2}$	
A_1	-50	150	$\overline{a}_1 = 50$
A_2	40	50	$\overline{a}_2 = 45$

II. Минимизация риска

В качестве оптимального выбирают то решение, где меньше риск (дисперсия).

	баг есть 0.01	бага нет 0.99
верить	∞	0
не верить	0	10^6

$\sigma(a_1) = \infty$
$\sigma(a_2) = 10^{12} \cdot 0.99$

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.

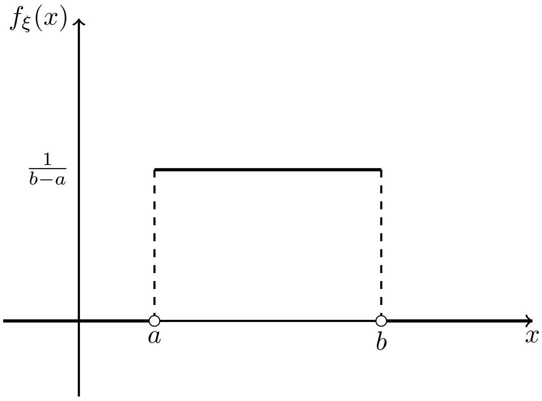


Рис. 1.8.2: Плотность равномерного распределения

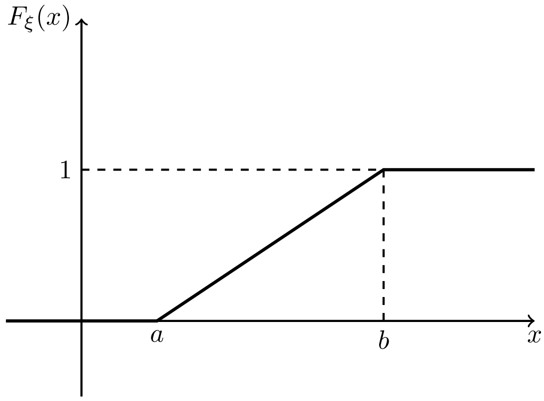


Рис. 1.8.3: Функция распределения

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

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

Найдем функцию распределения (рис. 1.8.3).

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

$$x < a \Longrightarrow F_\xi(x) = 0$$

$$a \leqslant x \leqslant b \Longrightarrow F_\xi(x) = \int_a^x \frac{\mathrm{d}x}{b-a} = \frac{x-a}{b-a}$$

$$x > b \Longrightarrow F_\xi(x) = 1$$

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

Вычислим числовые характеристики.

$$\mathbb{E}(\xi) = \int_{-\infty}^{\infty} xf(x)\mathrm{d}x = \int_a^b \frac{x\mathrm{d}x}{b-a} = \frac{1}{b-a} \cdot \frac{x^2}{2}\Big|_a^b = \frac{a+b}{2}$$

$$\mathbb{E}\left(\xi^2\right) = \int_{-\infty}^{\infty} x^2 f(x)\mathrm{d}x = \int_a^b \frac{x^2\mathrm{d}x}{b-a} = \frac{1}{b-a} \cdot \frac{x^3}{3}\Big|_a^b = \frac{a^2+ab+b^2}{3}$$

$$\mathbb{D}(\xi) = \mathbb{E}\left(\xi^2\right) - (\mathbb{E}(\xi))^2 = \frac{a^2+ab+b^2}{3} - \frac{a^2+2ab+b^2}{4} = \frac{(b-a)^2}{12}$$

$$\sigma(\xi) = \sqrt{\mathbb{D}(\xi)} = \frac{b-a}{2\sqrt{3}}$$

Вычислим вероятность попадания в интервал.

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

Равномерное распределение возникают 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):

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

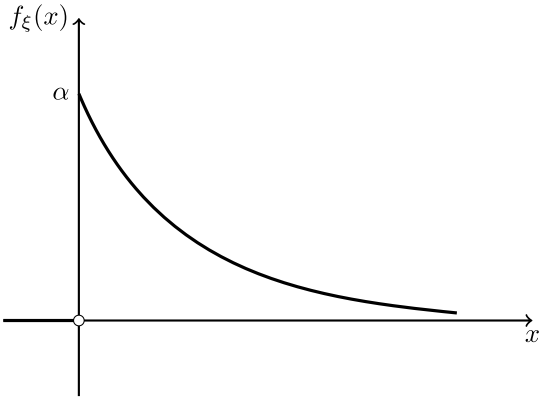


Рис. 1.8.5: Плотность показательного распределения

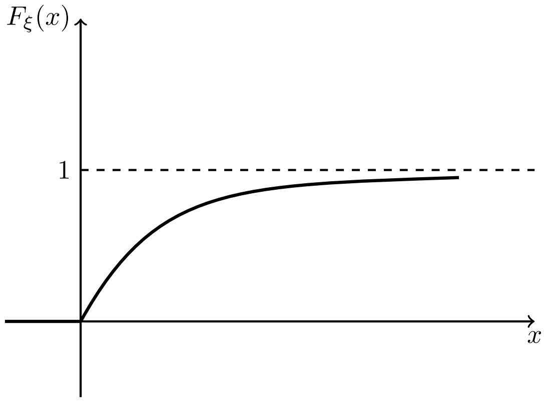


Рис. 1.8.6: Функция распределения

Найдем функцию распределения (рис. 1.8.6).

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

$$x < 0 \Longrightarrow F_\xi(x) = 0$$

$$x \geqslant 0 \Longrightarrow F_\xi(x) = \int_0^x \alpha e^{-\alpha x} = -e^{-\alpha x}\Big|_0^x = 1 - e^{-\alpha x}$$

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

Вычислим числовые характеристики.

$$\mathbb{E}(\xi) = \int_{-\infty}^{\infty} x f(x)\mathrm{d}x = \int_0^{\infty} x\alpha e^{-\alpha x}\mathrm{d}x = \begin{bmatrix} u = x & \mathrm{d}u = \mathrm{d}x \\ \mathrm{d}v = \alpha e^{-\alpha x}\mathrm{d}x & v = -e^{-\alpha x} \end{bmatrix}$$
$$= -xe^{-\alpha x}\Big|_0^{\infty} + \int_0^{\infty} e^{-\alpha x}\mathrm{d}x = -\lim_{x\to\infty}\frac{x}{e^{\alpha x}} - \frac{1}{\alpha}e^{-\alpha x}\Big|_0^{\infty} = -\frac{1}{\alpha}\left(\lim_{x\to\infty} e^{-\alpha x} - e^0\right) = \frac{1}{\alpha}$$

$$\mathbb{E}\left(\xi^2\right) = \int_{-\infty}^{\infty} x^2 f(x)\mathrm{d}x = \int_0^{\infty} x^2\alpha e^{-\alpha x}\mathrm{d}x = \begin{bmatrix} u = x^2 & \mathrm{d}u = 2x\mathrm{d}x \\ \mathrm{d}v = \alpha e^{-\alpha x}\mathrm{d}x & v = -e^{-\alpha x} \end{bmatrix}$$
$$= -x^2 e^{-\alpha x}\Big|_0^{\infty} + 2\int_0^{\infty} e^{-\alpha x}\mathrm{d}x = \frac{2}{\alpha}\int_0^{\infty} x\alpha e^{-\alpha x}\mathrm{d}x = \frac{2}{\alpha}\mathbb{E}(\xi) = \frac{2}{\alpha^2}$$

$$\mathbb{D}(\xi) = \mathbb{E}\left(\xi^2\right) - \left(\mathbb{E}(\xi)\right)^2 = \frac{2}{\alpha^2} - \left(\frac{1}{\alpha}\right)^2 = \frac{1}{\alpha^2}$$
$$\sigma(\xi) = \sqrt{\mathbb{D}(\xi)} = \frac{1}{\alpha}$$

Используя функцию распределения, вычислим вероятность попадания в интервал.

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

Из непрерывных распределений только показательное обладает свойством нестарения.

Теорема 1.8.7.

$$\mathsf{P}\left(\xi > x + y \,\middle|\, \xi > x\right) = \mathsf{P}\left(\xi > y\right)$$

□

$$\mathsf{P}\left(\xi > x + y \,\middle|\, \xi > x\right) = \frac{\mathsf{P}\left(\xi > x + y, \xi > x\right)}{\mathsf{P}\left(\xi > x\right)} = \frac{1 - \mathsf{P}\left(\xi < x + y\right)}{1 - \mathsf{P}\left(\xi < x\right)} = \frac{1 - F(x+y)}{1 - F(x)} = \frac{1 - 1 + e^{-\alpha(x+y)}}{1 - 1 + e^{-\alpha x}}$$
$$= e^{-\alpha y} = 1 - (1 - e^{-\alpha y}) = 1 - F(y) = \mathsf{P}\left(\xi > y\right)$$

■

Показательное распределение возникает, (например) если нужно рассчитать время работы надежного прибора до поломки. Также показательное распределение можно трактовать как время между появлениями 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

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

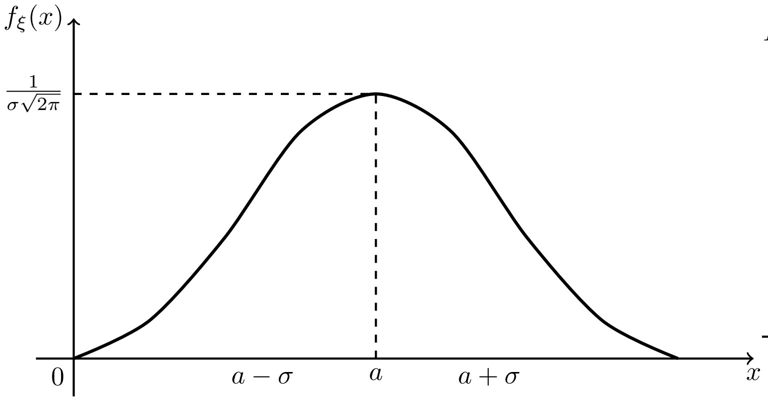


Рис. 1.8.9: Нормальное распределение

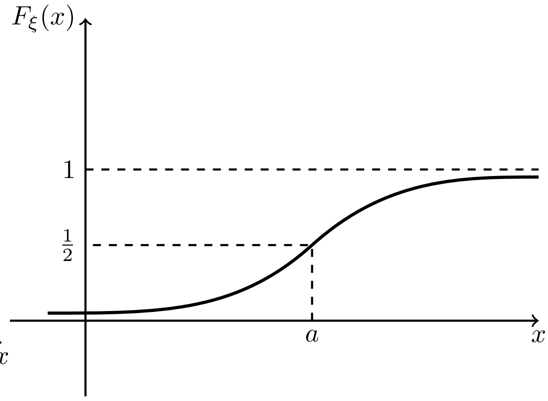


Рис. 1.8.10: Функция распределения

Смысл параметров распределения.

$$\mathbb{E}\left(\xi\right) = a \qquad \mathbb{D}\left(\xi\right) = \sigma^2 \qquad \sigma\left(\xi\right) = \sigma$$

Функция распределения будет иметь вид

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

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):

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

Функция распределения $F_0(x)$ будет иметь вид

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

Замечание 1.8.13. По формуле Пуассона

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

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

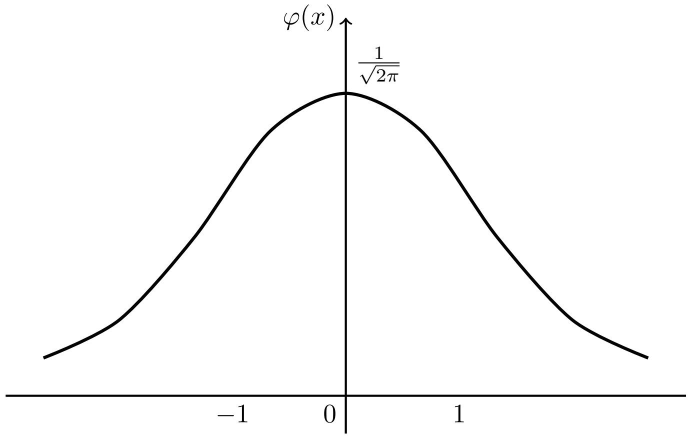


Рис. 1.8.12: Функция Гаусса

Замечание 1.8.14.

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

где $\Phi(x)$ это функция Лапласа.

Вычислим характеристики.

$$\begin{aligned}\mathbb{E}\,(\xi) &= \int_{-\infty}^{\infty} x\varphi(x)\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} xe^{-x^2/2}\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2}\mathrm{d}\left(\frac{x^2}{2}\right)\\ &= -\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\Big|_{-\infty}^{\infty} = -\frac{1}{\sqrt{2\pi}}\left(\lim_{x\to\infty} e^{-x^2/2} - \lim_{x\to-\infty} e^{-x^2/2}\right) = 0\end{aligned}$$

$$\begin{aligned}\mathbb{D}\,(\xi) &= \int_{-\infty}^{\infty} x^2\varphi(x)\mathrm{d}x - \underbrace{(\mathbb{E}\,(\xi))^2}_{0} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^2 e^{-x^2/2}\mathrm{d}x = \begin{bmatrix} u = x & \mathrm{d}u = \mathrm{d}x \\ \mathrm{d}v = xe^{-x^2/2}\mathrm{d}x & v = -e^{-x^2/2}\end{bmatrix}\\ &= \frac{1}{\sqrt{2\pi}}\left(\underbrace{-xe^{-x^2/2}\Big|_{-\infty}^{\infty}}_{0} + \int_{-\infty}^{\infty} e^{-x^2/2}\mathrm{d}x\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2}\mathrm{d}x = 1\end{aligned}$$

Связь между нормальным и стандартным нормальным распределением

<u>Lm</u> 1.8.15.

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

□

$$\begin{aligned}F_\xi(x) &= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}}\mathrm{d}t = \begin{bmatrix} z = \frac{t-a}{\sigma} & t = \sigma z + a & \mathrm{d}t = \sigma\mathrm{d}z \\ z(-\infty) = -\infty & z(x) = \frac{x-a}{\sigma} & \end{bmatrix}\\ &= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\frac{x-a}{\sigma}} e^{-z^2/2}\sigma\mathrm{d}z = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{x-a}{\sigma}} e^{-z^2/2}\mathrm{d}z = F_0\left(\frac{x-a}{\sigma}\right)\end{aligned}$$

■

<u>Lm</u> 1.8.16.

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

□

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

■

Lm 1.8.17.

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

□

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

$$\xi = \sigma\eta + a \Longrightarrow \begin{cases} \mathbb{E}\,(\xi) = \sigma\underbrace{\mathbb{E}\,(\eta)}_{0} + a = a \\ \mathbb{D}\,(\xi) = \sigma^2\underbrace{\mathbb{D}\,(\eta)}_{1} = \sigma^2 \end{cases}$$

■

Lm 1.8.18.

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

Lm 1.8.19.

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

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

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

□

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

■

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

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

Lm 1.8.22 (Правило «трех сигм»)**.**

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

□

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

■

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

$$\left.\begin{matrix} \xi_1 \in N(a_1;\sigma_1^2) \\ \xi_2 \in N(a_2;\sigma_2^2) \end{matrix}\right\} \Longrightarrow \xi_1 + \xi_2 \in N(a_1 + a_2;\sigma_1^2 + \sigma_2^2)$$

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.

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

□

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

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

■

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 $\mathsf{P}(\xi \in B) = 1$, но $\forall x \in B \,\Big|\, \mathsf{P}(\xi = x) = 0$.

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

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

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

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

$$F_\xi(x) = p_1F_1(x) + p_2F_2(x) + p_3F_3(x)$$

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 $\langle \Omega, \mathcal{F}, \mathsf{P}\rangle$. Рассмотрим функцию $g(\xi)\colon \Omega \to \mathbb{R}$, где $g(x)\colon \mathbb{R} \to \mathbb{R}$. Вопрос: когда $g(\xi)$ также является случайной величиной?

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

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

□ $\forall B \in \mathcal{B}$ прообраз $g^{-1}(B) = \xi^{-1}(B') \in \mathcal{F} \Longrightarrow g(\xi)$ 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

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

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

$$F_\eta(x) = \mathsf{P}\left(\eta < x\right) = \mathsf{P}\left(g(\xi) < x\right) = \mathsf{P}\left(\xi < h(x)\right) = \int_{-\infty}^{h(x)} f_\xi(t)\mathrm{d}t$$

$$\begin{bmatrix} t = h(y) & \mathrm{d}t = h'(y)\mathrm{d}y & y = g(t) \\ y(-\infty) = g(-\infty) & y(h(x)) = g(h(x)) = x & \end{bmatrix}$$

$$F_\eta(x) = \int_{g(-\infty)}^{x} f_\xi(h(y))h'(y)\mathrm{d}y \Longrightarrow f_\eta(x) = |h'(x)|\, f_\xi(h(x))$$

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

$$F_\eta(x) = \mathsf{P}\left(\eta < x\right) = \mathsf{P}\left(g(\xi) < x\right) = \mathsf{P}\left(\xi > h(x)\right) = \int_{h(x)}^{+\infty} f_\xi(t)\mathrm{d}t$$

$$\begin{bmatrix} t = h(y) & \mathrm{d}t = h'(y)\mathrm{d}y & y = g(t) \\ y(h(x)) = g(h(x)) = x & y(+\infty) = g(+\infty) & \end{bmatrix}$$

$$F_\eta(x) = \int_{x}^{g(+\infty)} f_\xi(h(y))h'(y)\mathrm{d}y = \int_{g(+\infty)}^{x} -h'(y) f_\xi(h(y))\mathrm{d}y \Longrightarrow f_\eta(x) = |h'(x)|\, f_\xi(h(x))$$

■

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

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

Подставляя все это в 1.9.8, получаем

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

■

<u>Lm</u> 1.9.10.

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

□

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

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

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

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

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

□ Это следствие из предыдущей леммы. ■

<u>Lm</u> 1.9.12. Если $\eta \in U\left(0; 1\right)$, то $\xi = a\eta + b \in U\left(b; b + a\right)$.

□ **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. $0 \leqslant F(x) \leqslant 1$. Рассмотрим два случая.

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

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

$$F_\eta(x) = \mathsf{P}\,(\eta < x) = \mathsf{P}\,(F(\xi) < x) = \mathsf{P}\left(\xi < F^{-1}(x)\right) = \begin{cases} 0, & x < 0 \\ F(F^{-1}(x)) = x, & 0 \leqslant x < 1 \\ 1, & x \geqslant 1 \end{cases}$$

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:

$$F^{-1}(x) = \min_t \left\{ t \,\middle|\, F(t) = x \right\}$$

Тогда при $0 < x < 1$

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

■

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

$$F^{-1}(x) = \inf_t \left\{ t \,\middle|\, F(t) \geqslant x \right\}$$

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 $F(x)$.

□
$$F_\xi(x) = \mathsf{P}\,(\xi < x) = \mathsf{P}\left(F^{-1}(\eta) < x\right) = \mathsf{P}\,(\eta < F(x)) = \frac{F(x) - 0}{1 - 0} = F(x)$$

■

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.

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

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

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

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

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

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

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

при условии, что данный ряд сходится абсолютно.

2. Если ξ это абсолютно непрерывная случайная величина, то

$$\mathbb{E}\left(g(\xi)\right) = \int_{-\infty}^{\infty} g(x) f_\xi(x)\mathrm{d}x$$

при условии, что данный интеграл сходится абсолютно.

Свойства моментов

<u>Lm</u> 1.9.20.

$$\xi \geqslant 0 \Longrightarrow \mathbb{E}\left(\xi\right) \geqslant 0$$

<u>Lm</u> 1.9.21.

$$\xi \geqslant \eta \Longrightarrow \mathbb{E}\left(\xi\right) \geqslant \mathbb{E}\left(\eta\right)$$

□

$$\xi \geqslant \eta \Longrightarrow \xi - \eta \geqslant 0 \Longrightarrow \mathbb{E}\left(\xi - \eta\right) = \mathbb{E}\left(\xi\right) - \mathbb{E}\left(\eta\right) \geqslant 0 \Longrightarrow \mathbb{E}\left(\xi\right) \geqslant \mathbb{E}\left(\eta\right)$$

■

<u>Lm</u> 1.9.22.

$$|\xi| \geqslant |\eta| \Longrightarrow \mathbb{E}\left(|\xi|^k\right) \geqslant \mathbb{E}\left(|\eta|^k\right)$$

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

□ Заметим, что

$$\left.\begin{array}{ll} x \geqslant 1\colon & |x|^s \leqslant |x|^t \\ x < 1\colon & |x|^s \leqslant 1 \end{array}\right\} \Longrightarrow |x|^s \leqslant \max\left(|x|^t, 1\right) \leqslant |x|^t + 1 \Longrightarrow \mathbb{E}\left(|\xi|^s\right) < \mathbb{E}\left(|\eta|^t\right) + 1$$

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

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

$$\mathbb{E}\left(\xi\right) \geqslant g\left(\mathbb{E}\left(\xi\right)\right)$$

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.

$$\forall x_0 \,\Big|\, \exists k(x_0) \,\Big|\, g(x) \geqslant g(x_0) + k(x_0)(x - x_0)$$

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

$$g(\xi) \geqslant g\left(\mathbb{E}\left(\xi\right)\right) + k\left(\mathbb{E}\left(\xi\right)\right)\left(\xi - \mathbb{E}\left(\xi\right)\right)$$

По 1.9.21 получаем, что

$$\mathbb{E}\left(g(\xi)\right) \geqslant \mathbb{E}\left(g\left(\mathbb{E}\left(\xi\right)\right)\right) + k\left(\mathbb{E}\left(\xi\right)\right)\underbrace{\mathbb{E}\left(\xi - \mathbb{E}\left(\xi\right)\right)}_{0}$$
$$\mathbb{E}\left(g(\xi)\right) \geqslant g\left(\mathbb{E}\left(\xi\right)\right)$$

■

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

$$\begin{aligned} \mathbb{E}\left(e^\xi\right) &\geqslant e^{\mathbb{E}(\xi)} \\ \mathbb{E}\left(\xi^n\right) &\geqslant \left(\mathbb{E}\left(\xi\right)\right)^n \qquad (n > 1) \\ \mathbb{E}\left(\ln\xi\right) &\leqslant \ln\left(\mathbb{E}\left(\xi\right)\right) \end{aligned}$$

1.10. Лекция 23.11.07.

Сходимость случайных величин

I. Сходимость «почти наверное»

Def 1.10.1. Случайная величина ξ имеет свойство Cnd «почти наверное», если $\mathsf{P}\left(\xi \text{ имеет свойство } Cnd\right) = 1$. Иными словами $\mathsf{P}\left(\omega \in \Omega \,\middle|\, \xi(\omega) \text{ не имеет свойства } Cnd\right) = 0$.

Def 1.10.2. Последовательность случайных величин $\{\xi_n\}$ «почти наверное» сходится к случайной величине ξ, если

$$\mathsf{P}\left(\omega \in \Omega \,\middle|\, \xi_n(\omega) \to \xi(\omega)\right) = 1$$

Обозначается $\xi_n \xrightarrow{\text{п.н.}} \xi$.

II. Сходимость по вероятности

Def 1.10.3. Последовательность случайных величин $\{\xi_n\}$ сходится по вероятности к случайной величине ξ, если

$$\forall \varepsilon > 0 \,\Big|\, \mathsf{P}\left(|\xi_n - \xi| < \varepsilon\right) \to 1 \qquad (n \to \infty)$$

Обозначается $\xi_n \xrightarrow{\mathsf{P}} \xi$.

Замечание 1.10.4. Это аналог сходимости по мере, только по вероятностной мере. Можно понимать как ослабленную равномерную сходимость.

$$\mathsf{P}\left(\omega \in \Omega \,\middle|\, \xi_n(\omega) \text{ не укладывается в } \varepsilon\text{-полосу}\right) \to 0$$

Замечание 1.10.5. В общем случае

$$\xi_n \xrightarrow{\mathsf{P}} \xi \not\Longrightarrow \mathbb{E}\left(\xi_n\right) \xrightarrow{n\to\infty} \mathbb{E}\left(\xi\right)$$

Теорема 1.10.6. Пусть $\forall n\colon\ |\xi_n| \leqslant C = const$, тогда

$$\xi_n \xrightarrow{\mathsf{P}} \xi \Longrightarrow \mathbb{E}\left(\xi_n\right) \xrightarrow{n\to\infty} \mathbb{E}\left(\xi\right)$$

III Слабая сходимость (сходимость по функции распределения)

Def 1.10.7. Последовательность случайных величин ξ_n слабо сходится к случайной величине ξ, если

$$F_{\xi_n}(x) \xrightarrow{n\to\infty} F_\xi \qquad \forall x \,\Big|\, F_\xi(x) \text{ непрерывна в точке } x$$

Таким образом, слабая сходимость это сходимость функций распределений во всех точках непрерывности. Обозначается $\xi_n \rightrightarrows \xi$.

Связь между видами сходимости

Теорема 1.10.8.

$$\xi_n \xrightarrow{\text{п.н.}} \xi \Longrightarrow \xi_n \xrightarrow{\mathsf{P}} \xi \Longrightarrow \xi_n \rightrightarrows \xi$$

Теорема 1.10.9. Если $\xi_n \rightrightarrows C = const$, то $\xi_n \xrightarrow{\mathsf{P}} C$.

$\square$

$$\xi_n \rightrightarrows C = const \Longrightarrow F_{\xi_n}(x) \to F_C(x) = \begin{cases} 0, & x < C \\ 1, & x \geqslant C \end{cases} \text{ за исключением точки разрыва } x = C$$

$$\mathsf{P}\left(|\xi_n - C| < \varepsilon\right) = \mathsf{P}\left(C - \varepsilon < \xi_n < C + \varepsilon\right) \geqslant \mathsf{P}\left(C - \frac{\varepsilon}{2} \leqslant \xi_n < C + \varepsilon\right) = F_{\xi_n}(C+\varepsilon) - F_{\xi_n}\left(C - \frac{\varepsilon}{2}\right)$$

$$F_{\xi_n}(C+\varepsilon) - F_{\xi_n}\left(C - \frac{\varepsilon}{2}\right) \xrightarrow{n\to\infty} F_C(C+\varepsilon) - F_C\left(C - \frac{\varepsilon}{2}\right) = 1 - 0 = 1$$

Т.к. $\mathsf{P}\left(|\xi_n - C| < \varepsilon\right) \leqslant 1$, то по теореме о двух милиционерах $\mathsf{P}\left(|\xi_n - C| < \varepsilon\right) \xrightarrow{n\to\infty} 1$. $\blacksquare$

Замечание 1.10.10. В общем случае из слабой сходимость не только не следует сходимость по вероятность, но даже бессмысленно говорить от этом, т.к. слабая сходимость это по существу не сходимость случайных величин, а сходимость их распределений.

Пример 1.10.11. Пусть $\xi \in N(0;1)$, тогда $\eta = -\xi \in N(0;1)$. Получаем, что $F_\xi(x) = F_\eta(x)$. Ясно, что

$$\xi_n \rightrightarrows \xi \not\Longrightarrow \xi_n \xrightarrow{\mathsf{P}} \xi$$

Т.к. в противном случае $\xi_n \xrightarrow{\mathsf{P}} \eta = -\xi$.

Ключевые неравенства

I. Неравенство Маркова

Теорема 1.10.12.

$$\forall \varepsilon > 0 \,\Big|\, \mathsf{P}\left(|\xi| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{E}\left(|\xi|\right)}{\varepsilon}$$

□

$$I_A(\omega) = \begin{cases} 0, \omega \notin A \text{ событие } A \text{ не происходит} \\ 1, \omega \in A \text{ событие } A \text{ происходит} \end{cases}$$

$I_A(\omega)$ имеет распределение Бернулли с параметром p, где $p = \mathsf{P}(A)$ и $\mathbb{E}\left(I_A(\omega)\right) = p = \mathsf{P}(A)$.

$$|\xi| \geqslant |\xi| \cdot I\left(|\xi| \geqslant \varepsilon\right) \geqslant \varepsilon \cdot I\left(|\xi| \geqslant \varepsilon\right)$$

Применим математическое ожидание к обеим частям неравенства.

$$\mathbb{E}\left(|\xi|\right) \geqslant \mathbb{E}\left(\varepsilon \cdot I\left(|\xi| \geqslant \varepsilon\right)\right) = \varepsilon \cdot \mathbb{E}\left(I\left(|\xi| \geqslant \varepsilon\right)\right) = \varepsilon \cdot \mathsf{P}\left(|\xi| \geqslant \varepsilon\right) \Longrightarrow \mathsf{P}\left(|\xi| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{E}\left(|\xi|\right)}{\varepsilon}$$

■

II. Неравенство Чебышева

Теорема 1.10.13.

$$\mathsf{P}\left(|\xi - \mathbb{E}\left(\xi\right)| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{D}\left(\xi\right)}{\varepsilon^2}$$

□ По 1.10.12 имеем

$$\mathsf{P}\left(|\xi - \mathbb{E}\left(\xi\right)| \geqslant \varepsilon\right) = \mathsf{P}\left(|\xi - \mathbb{E}\left(\xi\right)|^2 \geqslant \varepsilon^2\right) \leqslant \frac{\mathbb{E}\left(\left(\xi - \mathbb{E}\left(\xi\right)\right)^2\right)}{\varepsilon^2} = \frac{\mathbb{D}\left(\xi\right)}{\varepsilon^2}$$

■

<u>Lm</u> 1.10.14. Оценка для стандартизованной случайной величины Пусть $\eta = \frac{\xi - \mathbb{E}(\xi)}{\sigma(\xi)}$, тогда

$$\mathsf{P}\left(|\eta| \geqslant \varepsilon\right) \leqslant \frac{1}{\varepsilon^2}$$

III. Правило трех сигм

Теорема 1.10.15.

$$\mathsf{P}\left(|\xi - \mathbb{E}\left(\xi\right)| \geqslant 3\sigma\left(\xi\right)\right) \leqslant \frac{1}{9}$$

□ По 1.10.13 имеем

$$\mathsf{P}\left(|\xi - \mathbb{E}\left(\xi\right)| \geqslant 3\sigma\left(\xi\right)\right) \leqslant \frac{\mathbb{D}\left(\xi\right)}{\left(3\sigma\left(\xi\right)\right)^2} = \frac{\mathbb{D}\left(\xi\right)}{9\sigma\left(\xi\right)^2} = \frac{1}{9}$$

■

Среднее арифметическое случайных величин

Пусть $\xi_1, \ldots, \xi_n$ — независимая, одинаково распределенная случайная величина с конечным вторым моментом. Обозначим $a = \mathbb{E}\left(\xi_i\right)$, $d = \mathbb{D}\left(\xi_i\right)$, $\sigma = \sigma\left(\xi_i\right)$, $S_n = \xi_1 + \ldots + \xi_n$. Тогда

$$\mathbb{E}\left(\frac{S_n}{n}\right) = \mathbb{E}\left(\frac{\xi_1 + \ldots + \xi_n}{n}\right) = \frac{1}{n}\left(\mathbb{E}\left(\xi_1\right) + \ldots + \mathbb{E}\left(\xi_n\right)\right) = \frac{1}{n} \cdot an = a$$

Математическое ожидание осталось прежним.

$$\mathbb{D}\left(\frac{S_n}{n}\right) = \mathbb{D}\left(\frac{\xi_1 + \ldots + \xi_n}{n}\right) = \frac{1}{n^2}\left(\mathbb{D}\left(\xi_1\right) + \ldots + \mathbb{D}\left(\xi_n\right)\right) = \frac{1}{n^2} \cdot dn = \frac{d}{n}$$

Дисперсия уменьшилась в n раз.

$$\sigma\left(\frac{S_n}{n}\right) = \frac{\sigma}{\sqrt{n}}$$

Среднеквадратическое отклонение уменьшилось в $\sqrt{n}$ раз.
Законы больших чисел

I. Закон больших чисел Чебышева

Теорема 1.10.16. Пусть $\xi_1, \ldots, \xi_n, \ldots$ — последовательность независимых одинаково распределенных случайных величин с конечным вторым моментом, тогда

$$\frac{\xi_1 + \ldots + \xi_n}{n} \xrightarrow{\mathsf{P}} \mathbb{E}(\xi_1)$$

□ Обозначим $a = \mathbb{E}(\xi_i)$, $d = \mathbb{D}(\xi_i) < \infty$, $S_n = \sum_{i=1}^{n} \xi_i$. Тогда по 1.10.13 имеем

$$\mathsf{P}\left(\left|\frac{S_n}{n} - a\right| \geqslant \varepsilon\right) = \mathsf{P}\left(\left|\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)\right| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{D}\left(\frac{S_n}{n}\right)}{\varepsilon^2} = \underbrace{\frac{d}{\varepsilon^2}}_{=const} \cdot \underbrace{\frac{1}{n}}_{\to 0} \xrightarrow{n\to\infty} 0 \Longrightarrow \mathsf{P}\left(\left|\frac{S_n}{n} - a\right| \leqslant \varepsilon\right) \xrightarrow{n\to\infty} 1$$

что и означает сходимость по вероятности. ■

Среднее арифметическое большого числа независимых одинаковых случайных величин «стабилизируется» около математического ожидания. Или при больших n случайность переходит в почти закономерность.
Статистический смысл: при большом объеме n экспериментальных данных их среднее арифметическое даст достаточно точную оценку неизвестного математического ожидания.

Замечание 1.10.17. Попутно получили удобную, но грубую оценку:

$$\mathsf{P}\left(\left|\frac{S_n}{n} - a\right| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{D}(\xi_1)}{n\varepsilon^2}$$

II. Закон больших чисел Бернулли

Теорема 1.10.18. Пусть v_n — число появления события A в серии из n независимых испытаний, $p_n = \mathsf{P}(A)$ — вероятность появления A при одном испытании. Тогда

$$\frac{v_n}{n} \xrightarrow{\mathsf{P}} \mathsf{P}(A) \qquad \mathsf{P}\left(\left|\frac{v_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant \frac{p(1-p)}{n\varepsilon^2} \qquad (\forall \varepsilon > 0)$$

□ $v_n = \xi_1 + \ldots + \xi_n$, где ξ_i — число появления события A при i-том испытании, причем ξ_i независимые и имеют одинаковое распределение B_p. Т.к. $\mathbb{E}(\xi_i) = p$, $\mathbb{D}(\xi_i) = pq$, то по 1.10.16

$$\frac{v_n}{n} = \frac{S_n}{n} \xrightarrow{\mathsf{P}} \mathbb{E}(\xi_1) = \mathsf{P}(A)$$
$$\mathsf{P}\left(\left|\frac{v_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant \frac{pq}{n\varepsilon^2}$$

■

III. Закон больших чисел Хинчина

Теорема 1.10.19. Пусть $\xi_1, \ldots, \xi_n, \ldots$ последовательность независимых одинаково распределенных случайных величин с конечным первым моментом $\mathbb{E}(\xi_1) < \infty$, тогда

$$\frac{\xi_1 + \ldots + \xi_n}{n} \xrightarrow{\mathsf{P}} \mathbb{E}(\xi_1)$$

IV. Усиленный закон больших чисел Колмогорова

Теорема 1.10.20. В условиях теоремы 1.10.19

$$\frac{\xi_1 + \ldots + \xi_n}{n} \xrightarrow{\text{п.н.}} \mathbb{E}(\xi_1)$$

V. Закон больших чисел Маркова

Теорема 1.10.21. Пусть имеется последовательность случайных величин $\xi_1, \ldots, \xi_n, \ldots$ таких, что $\mathbb{D}(S_n) = o(n^2)$, и их вторые моменты конечны. Тогда

$$\frac{S_n}{n} \xrightarrow{\mathsf{P}} \frac{\mathbb{E}(\xi_1) + \ldots + \mathbb{E}(\xi_n)}{n}$$

□

$$\mathbb{E}\left(\frac{S_n}{n}\right) = \frac{\mathbb{E}(\xi_1) + \ldots + \mathbb{E}(\xi_n)}{n}$$

И по 1.10.13 получаем

$$\mathsf{P}\left(\left|\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)\right| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{D}\left(\frac{S_n}{n}\right)}{\varepsilon^2} = \frac{\mathbb{D}(S_n)}{n^2\varepsilon^2} = \frac{o(n^2)}{n^2\varepsilon^2} = \underbrace{\frac{o(n^2)}{n^2}}_{\to 0} \cdot \underbrace{\frac{1}{\varepsilon^2}}_{=const} \xrightarrow{n\to\infty} 0$$

■

Центральная предельная теорема

Теорема 1.10.22. (Центральная предельная теорема Ляпунова) Пусть $\xi_1, \ldots, \xi_n$ — последовательность независимых одинаково распределенных случайных величин с конечным вторым моментом $\mathbb{D}(\xi_1) < \infty$ и $S_n = \sum_{i=1}^n \xi_i$. Тогда имеет место слабая сходимость

$$\frac{S_n - n\mathbb{E}(\xi_1)}{\sqrt{n\mathbb{D}(\xi_1)}} \rightrightarrows N(0;1)$$

Т.е. стандартизованная сумма в пределе имеет стандартное нормальное распределение.

Замечание 1.10.23. Теорему можно сформулировать по-другому. Пусть $a = \mathbb{E}(\xi_1)$, $\sigma = \sigma(\xi_1)$. Тогда, как уже было показано $\mathbb{E}\left(\frac{S_n}{n}\right) = a$ и $\mathbb{D}\left(\frac{S_n}{n}\right) = \frac{\sigma}{\sqrt{n}}$, при этом

$$\frac{\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)}{\sigma\left(\frac{S_n}{n}\right)} \rightrightarrows N(0;1)$$

Т.е. стандартизованное среднее арифметическое в пределе имеет стандартное нормальное распределение.

Замечание 1.10.24. В грубой формулировке теорема будет иметь вид

$$\frac{S_n}{n} \rightrightarrows N\left(a; \frac{\sigma^2}{n}\right)$$

1.11. Лекция 23.11.14.

Совместное распределение случайных величин

Def 1.11.1. Случайным вектором $\langle \xi_1, \ldots, \xi_n \rangle$ называется упорядоченный набор случайных величин, заданных на одном вероятностном пространстве $\langle \Omega, \mathcal{F}, \mathsf{P} \rangle$.

Случайный вектор задает отображение $\langle \xi_1, \ldots, \xi_n \rangle(\omega)\colon \Omega \to \mathbb{R}^n$, поэтому его еще называют многомерной случайной величиной, а соответствующее распределение называют многомерным распределением.

$$\mathsf{P}(B) = \mathsf{P}\left(\omega \in \Omega \,\middle|\, \langle \xi_1, \ldots, \xi_n \rangle(\omega) \in B\right) \qquad B \in \mathcal{B}(\mathbb{R}^n)$$

Таким образом мы получаем новое вероятностное пространство $\langle \mathbb{R}^n, \mathcal{B}, \mathsf{P}(B) \rangle$.
Функция распределения

Def 1.11.2. Функцией совместного распределения случайных величин $\xi_1, \ldots, \xi_n$ (или случайного вектора) называется функция

$$F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = \mathsf{P}(\xi_1 < x_1, \ldots, \xi_n < x_n)$$

Замечание 1.11.3. Зная функцию распределения можно найти вероятность попадания случайного вектора в любое борелевское множество, а значит функция распределения полностью задает распределение.

Замечание 1.11.5. Далее рассматриваем только системы из двух случайных величин.

$$F_{\xi,\eta}(x, y) = \mathsf{P}(\xi < x, \eta < y)$$

Иллюстрация приведена на рис. 1.11.4.

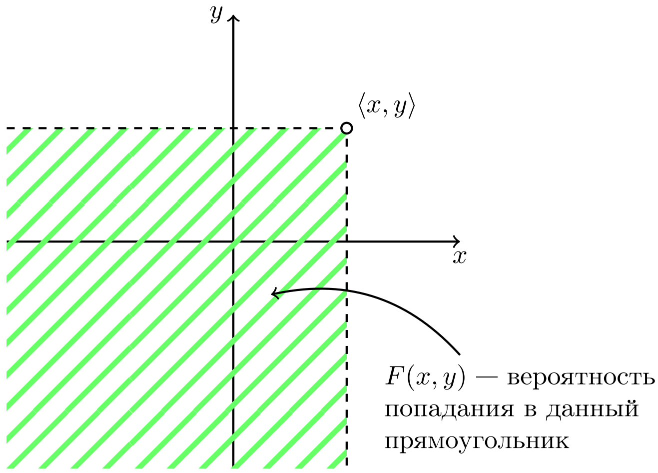


Рис. 1.11.4: Функция совместного распределения двух случайных величин

Свойства функции распределения

$$0 \leqslant F_{\xi,\eta}(x,y) \leqslant 1$$
$$F_{\xi,\eta}(x,y) \text{ неубывающая по каждому аргументу}$$
$$\lim_{x\to-\infty} F_{\xi,\eta}(x,y) = 0$$
$$\lim_{y\to-\infty} F_{\xi,\eta}(x,y) = 0$$
$$\lim_{\substack{x\to\infty\\ y\to\infty}} F_{\xi,\eta}(x,y) = 1$$
$$F_{\xi,\eta}(x,y) \text{ непрерывна слева по каждому аргументу}$$
$$\lim_{y\to+\infty} F_{\xi,\eta}(x,y) = F_\xi(x)$$
$$\lim_{x\to+\infty} F_{\xi,\eta}(x,y) = F_\eta(y)$$
$$\mathsf{P}\left(x_1 \leqslant \xi < x_2, y_1 \leqslant \eta < y_2\right) = F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) + F(x_1,y_1)$$

Независимость случайных величин

Def 1.11.6. Случайные величины $\xi_1,\dots,\xi_n$ независимы в совокупности, если

$$\mathsf{P}\left(\xi_1 \in B_1,\dots,\xi_n \in B_n\right) = \mathsf{P}\left(\xi_1 \in B_1\right)\dots\mathsf{P}\left(\xi_n \in B_n\right) \qquad \forall B_1,\dots,B_n \in \mathcal{B}\left(\mathbb{R}^n\right)$$

Т.е. случайные величины принимают значения независимо друг от друга.

Def 1.11.7. Случайные величины $\xi_1,\dots,\xi_n$ попарно независимы, если независимы любые две из них.

Замечание 1.11.8. Из независимости в совокупности следует попарная независимость. Пусть $\xi_1,\dots,\xi_n$ независимы в совокупности, тогда $\forall i,j$ возьмем $B_k = \mathbb{R}$, $k \neq i, k \neq j$. Получим

$$\mathsf{P}\left(\xi_i \in B_i, \xi_j \in B_j\right) = 1\cdot\ldots\cdot\mathsf{P}\left(\xi_i \in B_i\right)\cdot\ldots\cdot\mathsf{P}\left(\xi_j \in B_j\right)\cdot\ldots\cdot 1 = \mathsf{P}\left(\xi_i \in B_i\right)\mathsf{P}\left(\xi_j \in B_j\right)$$

Обратное в общем случае неверно

Т.к. вероятности $\mathsf{P}\left(\xi_1 \in B_1,\dots,\xi_n \in B_n\right)$ полностью определяются функцией совместного распределения, то получаем эквивалентное определение:

Def 1.11.9. $\xi_1,\dots,\xi_n$ независимы в совокупности, если

$$F_{\xi_1,\dots,\xi_n}(x_1,\dots,\xi_n) = F_{\xi_1}(x)\dots F_{\xi_n}(x)$$

Замечание 1.11.10. В дальнейшем под «независимостью» будем понимать независимость в совокупности.

Дискретная система двух случайных величин

Def 1.11.11. Случайные величины ξ и η имеют дискретное совместное распределение, если случайный вектор $\langle\xi,\eta\rangle$ принимает не более чем счетное число значений. Т.е. существует конечный или счетный набор точек x_i, y_j таких, что

$$\forall i,j \,\Big|\, \mathsf{P}\left(\xi = x_i, \eta = y_j\right) > 0 \qquad \sum_{i,j}\mathsf{P}\left(\xi = x_i, \eta = y_j\right) = 1$$

Таким образом двумерная дискретная случайная величина задается законом распределения — таблицей вероятностей $\mathsf{P}_{i,j} = \mathsf{P}(\xi = x_i, \eta = y_j)$.

$\xi \backslash \eta$	y_1	y_2	$\dots$	y_m
x_1	$p_{1,1}$	$p_{1,2}$	$\dots$	$p_{1,m}$
x_2	$p_{2,1}$	$p_{2,2}$	$\dots$	$p_{2,m}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_n	$p_{n,1}$	$p_{n,2}$	$\dots$	$p_{n,m}$

x_i — значения случайной величины ξ, y_j — значения случайной величины η, $p_{i,j}$ — вероятность появления точки (x_i, y_j).
По условию нормировки $\sum_{i,j} p_{i,j} = 1$.
Зная общий закон распределения можно найти частные законы распределения ξ и η по формулам

$$p_i = \mathsf{P}(\xi = x_i) = \sum_{j=1}^{m} p_{i,j} \qquad p_j = \mathsf{P}(\eta = y_j) = \sum_{i=1}^{n} p_{i,j}$$

Def 1.11.12. Дискретные случайные величины $\xi_1, \dots, \xi_n$ независимы, если

$$\mathsf{P}(\xi_1 = x_1, \dots \xi_n = x_n) = \mathsf{P}(\xi_1 = x_1) \dots \mathsf{P}(\xi_n = x_n) \qquad \forall x_i \in \mathbb{R}$$

Замечание 1.11.13. Т.е. в случае двух случайных величин ξ и η независимы, если $\forall i, j \,\Big|\, p_{i,j} = p_i q_j$.

Абсолютно непрерывная система двух случайных величин

Def 1.11.14. Случайные величины ξ и η имеют абсолютно непрерывное совместное распределение, если

$$\exists f_{\xi,\eta}(x,y) \,\Big|\, \mathsf{P}(\langle \xi, \eta \rangle \in B) = \iint_B f_{\xi,\eta}(x,y) \mathrm{d}x \mathrm{d}y$$

Def 1.11.15. Функция $f_{\xi,\eta}(x,y)$ называется плотностью совместного распределения случайных величин ξ и η.

Замечание 1.11.16. В n-мерном случае определение плотности аналогично:

$$\mathsf{P}(\langle \xi_1, \dots, \xi_n \rangle \in B) = \int \!\!\cdots\!\! \int_B f_{\xi_1,\dots,\xi_n}(x_1, \dots, x_n) \mathrm{d}x_1 \dots \mathrm{d}x_n$$

Свойства плотности

$$\begin{aligned} f_{\xi,\eta}(x,y) &\geqslant 0 \\ \text{Условие нормировки } \iint_{\mathbb{R}^2} f_{\xi,\eta}(x,y) \mathrm{d}x \mathrm{d}y &= 1 \\ F_{\xi,\eta}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{\xi,\eta}(x,y) \mathrm{d}x \mathrm{d}y \\ f_{\xi,\eta}(x,y) = \frac{\partial F_{\xi,\eta}(x,y)}{\partial x \partial y} \end{aligned}$$

__Lm__ **1.11.17.** Если случайные величины ξ и η имеют абсолютные непрерывные распределения с плотностью $f_{\xi,\eta}(x,y)$, то маргинальные распределения случайных величин ξ и η также абсолютно непрерывны с плотностями

$$f_\xi(x) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y) \mathrm{d}y \qquad f_\eta(y) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y) \mathrm{d}x$$

□

$$\left.\begin{aligned} F_\xi(x) &= \lim_{y \to \infty} F_{\xi,\eta}(x,y) = \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} f_{\xi,\eta}(x,y) \mathrm{d}y \right) \mathrm{d}x \\ F_\xi(x) &= \int_{-\infty}^{x} f_\xi(x) \mathrm{d}x \end{aligned}\right\} \Longrightarrow f_\xi(x) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y) \mathrm{d}y$$

Доказательство для $F_\eta(y)$ аналогично. ■

Теорема 1.11.18. Абсолютно непрерывные случайные величины $\xi_1, \dots, \xi_n$ независимы в совокупности тогда и только тогда, когда

$$f_{\xi_1,\dots,\xi_n}(x_1, \dots, x_n) = f_{\xi_1}(x) \dots f_{\xi_n}(x)$$

□ *(только при $n=2$)*
Пусть случайные величины ξ и η независимы, тогда

$$\left.\begin{aligned} F_{\xi,\eta}(x,y) &= F_\xi(x)F_\eta(y) = \int_{-\infty}^{x} f_\xi(x)\mathrm{d}x \int_{-\infty}^{y} f_\eta(y)\mathrm{d}y = \int_{-\infty}^{x}\int_{-\infty}^{y} f_\xi(x)f_\eta(y)\mathrm{d}x\mathrm{d}y \\ F_{\xi,\eta}(x,y) &= \int_{-\infty}^{x}\int_{-\infty}^{y} f_{\xi,\eta}(x,y)\mathrm{d}x\mathrm{d}y \end{aligned}\right\} \Longrightarrow f_\xi(x)f_\eta(y) = f_{\xi,\eta}(x,y)$$

■

Замечание 1.11.19. Совместное распределение абсолютно непрерывных случайных величин не обязано быть абсолютно непрерывным. Оно может быть сингулярным.

Пример 1.11.20. Бросаем наугад точку на отрезок прямой $y=2x$, $0 \leqslant x \leqslant 1$. Пусть случайная величина $\xi \in U(0;1)$ это первая координата точки, а $\eta \in U(0;2)$ — вторая. Обе случайные величины имеют абсолютно непрерывное равномерное распределение, однако случайный вектор $\langle \xi, \eta \rangle$ не имеет абсолютно непрерывного распределения, т.к. мера Лебега отрезка равна нулю и случайный вектор распределен на несчетном множестве нулевой меры. Таким образом имеем сингулярное распределение.

Пример 1.11.21. Случайному вектору $\langle \xi, \eta \rangle$ сопоставляем его направление. Направление задается углом, т.е. значение случайного вектора — точки на единичной окружности, которая имеет меру ноль.

Многомерное равномерное распределение

Def 1.11.22. Пусть область $D \subset \mathbb{R}^n$ это борелевское множество с конечной лебеговой мерой $\lambda(D) > 0$. Случайный вектор $\langle \xi_1, \dots, \xi_n \rangle$ имеет равномерное распределение в области D, если плотность совместного распределения постоянна в области D и равна нулю вне этой области.

$$f_{\xi_1,\dots,\xi_n}(x_1,\dots,x_n) = \begin{cases} 0, & (x_1,\dots,x_n) \notin D \\ \frac{1}{\lambda(D)}, & (x_1,\dots,x_n) \in D \end{cases}$$

Смысл: случайный вектор $\langle \xi_1, \dots, \xi_n \rangle$ это координаты наугад брошенной точки в области D. В этом и только в этом случае применима формула геометрической вероятности.
Многомерное нормальное распределение (кратко)

Def 1.11.23. Случайный вектор $\langle \xi_1, \dots, \xi_n \rangle$ имеет многомерное нормальное распределение, если каждая его координата имеет нормальное распределение (но их параметры могут отличаться) и существует плотность совместного распределения.

Многомерное нормальное распределение задается через вектор математических ожиданий и матрицу ковариаций.

$$\mathbb{E}\left(\langle \xi_1, \dots, \xi_n \rangle\right) = \begin{pmatrix} \mathbb{E}(\xi_1) \\ \vdots \\ \mathbb{E}(\xi_n) \end{pmatrix} \qquad \operatorname{cov}(\xi_i, \xi_j) = \begin{pmatrix} \mathbb{D}(\xi_1) & \operatorname{cov}(\xi_1,\xi_2) & \dots & \operatorname{cov}(\xi_1,\xi_n) \\ \operatorname{cov}(\xi_1,\xi_2) & \mathbb{D}(\xi_2) & \dots & \operatorname{cov}(\xi_2,\xi_n) \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{cov}(\xi_1,\xi_n) & \operatorname{cov}(\xi_2,\xi_n) & \dots & \mathbb{D}(\xi_n) \end{pmatrix}$$

1.12. Лекция 23.11.21.

Функции от двух случайных величин

Если известно совместное распределение двух или нескольких случайных величин, то можно найти распределение их суммы, произведения и иных функций от этих случайных величин.
Пусть ξ_1 и ξ_2 — случайные величины с плотностью совместного распределения $f_{\xi_1,\xi_2}(x,y)$. Дана $g(x,y)\colon \mathbb{R}^2 \to \mathbb{R}$ — борелевская функция. Требуется найти функцию распределения случайной величины $\eta = g(\xi_1, \xi_2)$.
Теорема 1.12.1. Функция распределения случайной величины η

$$F_\eta(z) = \iint_{D_z} f_{\xi_1,\xi_2}(x,y)\mathrm{d}x\mathrm{d}y \qquad D_z = \left\{ \langle x, y \rangle \;\middle|\; g(x,y) < z \right\}$$

□

$$F_\eta(z) = \mathsf{P}(\eta < z) = \mathsf{P}(g(\xi_1,\xi_2) < z) = \mathsf{P}((x,y) \in D_z) = \iint_{D_z} f_{\xi_1,\xi_2}(x,y)\mathrm{d}x\mathrm{d}y$$

■

Пример 1.12.2. Двое договариваются встретиться между 12:00 и 13:00. Случайная величина η — время ожидания. Найти ее функцию распределения.

Решение: Пусть случайные величины ξ_1, ξ_2 — это время прихода первого и второго человека соответственно. Тогда $\eta = |\xi_1 - \xi_2|$. Имеем

$$\xi_1, \xi_2 \in U\,(0;1) \Longrightarrow \begin{cases} f_{\xi_1}(x) = \begin{cases} 1, & x \in [0;1] \\ 0, & x \notin [0;1] \end{cases} \\ f_{\xi_2}(y) = \begin{cases} 1, & y \in [0;1] \\ 0, & y \notin [0;1] \end{cases} \end{cases}$$

Т.к. ξ_1 и ξ_2 независимы, то

$$f_{\xi_1,\xi_2}(x,y) = f_{\xi_1}(x) f_{\xi_2}(y) = 1 \qquad 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1$$

Далее применяем теорему 1.12.1 и получаем, что

$$F_\eta(z) = \iint_{D_z} f_{\xi_1,\xi_2}(x,y)\mathrm{d}x\mathrm{d}y \qquad D_z = \left\{ \langle x,y \rangle \,\middle|\, |x-y| < z \right\}$$

$$F_\eta(z) = \iint_{D_z} \mathrm{d}x\mathrm{d}y = S_{D_z} \xlongequal{\text{рис. } 1.12.3} 1 - 2 \cdot \frac{1}{2}(1-z)^2 = 2z - z^2 \qquad z \in [0;1]$$

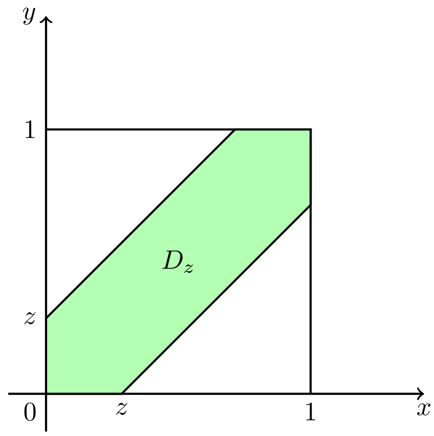


Рис. 1.12.3: Иллюстрация к задаче о встрече

Формула свертки

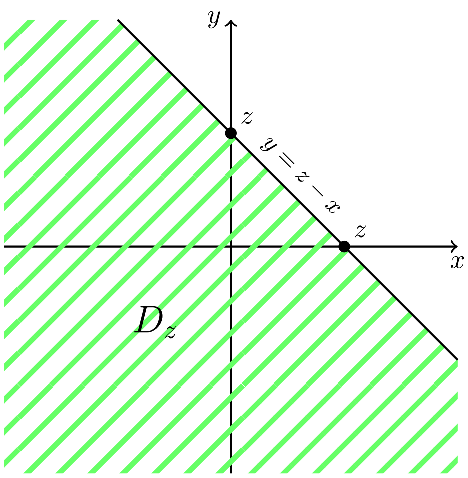


Рис. 1.12.4: Иллюстрация области D_z в формуле свертки

Теорема 1.12.5. Пусть ξ_1 и ξ_2 независимые абсолютно непрерывные случайные величины с плотностями $f_{\xi_1}(x)$ и $f_{\xi_2}(y)$ тогда плотность суммы $\xi_1 + \xi_2$ существует и равна свертке плотностей $f_{\xi_1}(x)$ и $f_{\xi_2}(y)$:

$$f_{\xi_1+\xi_2}(t) = \int_{-\infty}^{\infty} f_{\xi_1}(x) f_{\xi_2}(t-x)\mathrm{d}x$$

$\square$ Т.к. ξ_1 и ξ_2 независимые, то $f_{\xi_1,\xi_2}(x,y) = f_{\xi_1}(x)f_{\xi_2}(y)$. Применим 1.12.1.

$$g(\xi_1, \xi_2) = \xi_1 + \xi_2 \Longrightarrow D_z = \left\{\langle x, y\rangle \in \mathbb{R}^2 \,\middle|\, x + y < z\right\}$$

Для перехода к кратному интегралу смотрим на рис. 1.12.4

$$F_{\xi_1+\xi_2}(z) = \iint_{D_z} f_{\xi_1}(x)f_{\xi_2}(y)\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{z-x} f_{\xi_1}(x)f_{\xi_2}(y)\mathrm{d}y$$

$$\begin{bmatrix} y = t - x & \mathrm{d}y = \mathrm{d}t & t = y + x \\ t(-\infty) = -\infty & t(z-x) = z & \end{bmatrix}$$

$$F_{\xi_1+\xi_2}(z) = \int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{z} f_{\xi_1}(x)f_{\xi_2}(t-x)\mathrm{d}t$$

Поменяем порядок интегрирования

$$F_{\xi_1+\xi_2}(z) = \int_{-\infty}^{z}\left(\int_{-\infty}^{\infty} f_{\xi_1}(x)f_{\xi_2}(t-x)\mathrm{d}x\right)\mathrm{d}t \Longrightarrow f_{\xi_1+\xi_2}(t) = \int_{-\infty}^{\infty} f_{\xi_1}(x)f_{\xi_2}(t-x)\mathrm{d}x$$

$\blacksquare$

Замечание 1.12.6. Сумма независимых абсолютно непрерывных случайных величин также имеет абсолютно непрерывное распределение. Для зависимых случайных величин это в общем случае неверно. Например, $\eta = -\xi$. Тогда $\eta + \xi = 0$ — вырожденное распределение.

Замечание 1.12.7. Если ξ и η независимые случайные величины, ξ имеет абсолютно непрерывное распределение, а η — дискретное, то их сумма также имеет абсолютно непрерывное распределение.

Суммы стандартных распределений. Устойчивость относительно суммирования

Def 1.12.8. Если сумма двух независимых случайных величин одного типа распределений также будет этого типа, то говорят, что распределение устойчиво относительно суммирования.

Пример 1.12.9. Пусть $\xi \in B_{n,p}$ и $\eta \in B_{m,p}$ — независимые случайные величины. Тогда $\xi + \eta \in B_{n+m,p}$. Т.к. ξ и η это число появления события с вероятностью p в сериях из n и m испытаний, то их сумма это число успехов в серии из $n + m$ испытаний.

Замечание 1.12.10. Если второй параметр p разный, то распределение не будет устойчивым.

Lm 1.12.11. Пусть $\xi \in \Pi_\alpha$ и $\eta \in \Pi_\mu$ — независимые случайные величины. Тогда $\xi + \eta \in \Pi_{\lambda+\mu}$.

$\square$ Пусть $k \geqslant 0$. По формуле Пуассона $\mathsf{P}(\xi = k) = \frac{\lambda^k}{k!}e^{-\lambda}$. Получаем

$$\begin{aligned}
\mathsf{P}(\xi + \eta = k) &= \sum_{i=0}^{k}\mathsf{P}(\xi = i, \eta = k - i) \\
&= \sum_{i=0}^{k}\mathsf{P}(\xi = i)\,\mathsf{P}(\eta = k - i) \\
&= \sum_{i=0}^{k}\frac{\lambda^i}{i!}e^{-\lambda}\cdot\frac{\mu^{k-i}}{(k-i)!}e^{-\mu} \\
&= e^{-\lambda-\mu}\sum_{i=0}^{k}\frac{\lambda^i\mu^{k-i}}{i!(k-i)!} \\
&= e^{-\lambda-\mu}\frac{1}{k!}\sum_{i=0}^{k}\frac{k!}{i!(k-i)!}\cdot\lambda^i\mu^{k-i} \\
&= e^{-\lambda-\mu}\frac{1}{k!}\sum_{i=0}^{k}C_i^k\cdot\lambda^i\mu^{k-i} \\
&= \frac{(\lambda+\mu)^k}{k!}e^{-(\lambda+\mu)}
\end{aligned}$$

Таким образом $\xi + \eta \in \Pi_{\lambda+\mu}$. $\blacksquare$

Lm 1.12.12. Пусть $\xi \in N(0;1)$ и $\eta \in N(0;1)$ — независимые случайные величины. Тогда $\xi + \eta \in N(0;2)$.

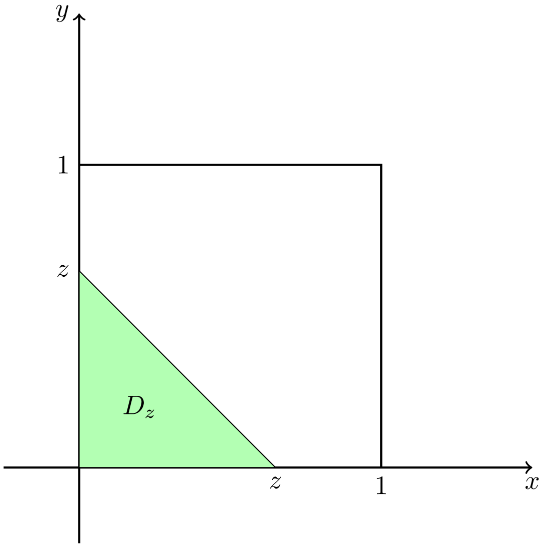
□

$$\left.\begin{aligned} f_\xi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \\ f_\eta(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}} \end{aligned}\right\} \Longrightarrow f_{\xi+\eta}(t) = \int_{-\infty}^{\infty} f_\xi(x) f_\eta(t-x)\mathrm{d}x = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-(x^2-tx+\frac{t^2}{2})}\mathrm{d}x = \frac{1}{2\pi}e^{-\frac{t^2}{4}}\int_{-\infty}^{\infty} e^{-(x^2-tx+\frac{t^2}{4})}\mathrm{d}x$$

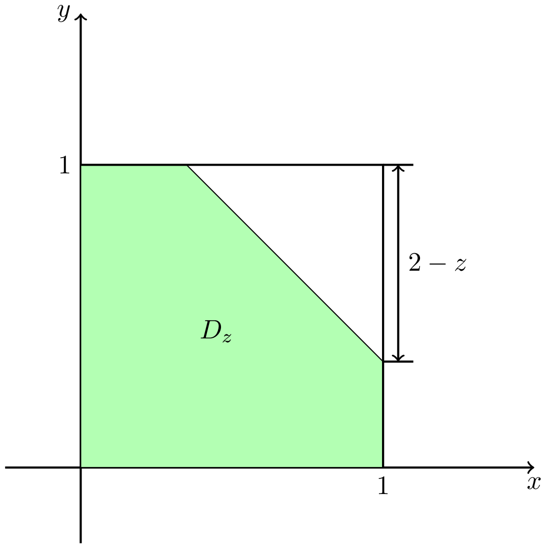
$$f_{\xi+\eta}(t) = \frac{1}{2\pi}e^{-\frac{t^2}{4}} \underbrace{\int_{-\infty}^{\infty} e^{-\left(x-\frac{t}{2}\right)^2}\mathrm{d}\left(x-\frac{t}{2}\right)}_{\text{Интеграл Пуассона }=\sqrt{\pi}} = \frac{1}{\sqrt{2}\sqrt{2\pi}}e^{-\frac{t^2}{2(\sqrt{2})^2}}$$

Таким образом $\xi+\eta \in N(0;(\sqrt{2})^2)$. ■

Замечание 1.12.13. Пусть $\xi \in N(a_1;\sigma_1)$ и $\eta \in N(a_2;\sigma_2)$ — независимые случайные величины. Тогда $\xi+\eta \in N(a_1 + a_2;\sigma_1+\sigma_2)$. Это можно доказать аналогично доказательству выше, но далее будет приведен более простой способ.



(a) $0 \leqslant z \leqslant 1$



(b) $1 \leqslant z \leqslant 2$

Рис. 1.12.14: Площадь области D_z

Lm 1.12.15. Равномерное распределение не является устойчивым относительно суммирования.

□ Пусть $\xi \in U\,(0;1)$ и $\eta \in U\,(0;1)$ — независимые случайные величины.

$$\left.\begin{aligned} f_\xi(x) = 1, x\in[0;1] \\ f_\eta(y) = 1, y \in [0;1] \end{aligned}\right\} \Longrightarrow f_{\xi,\eta}(x,y) = f_\xi(x)f_\eta(y) = 1$$

$$F_{\xi+\eta}(z) = \iint_{D_z} f_{\xi,\eta}(x,y)\mathrm{d}x\mathrm{d}y \qquad D_z = \left\{\langle x,y\rangle \;\middle|\; x+y<z\right\}$$

Для вычисления площади области D_z смотрим на рис. 1.12.14

$$F_{\xi+\eta}(z) = \iint_{D_z} \mathrm{d}x\mathrm{d}y = S_{D_z} = \begin{cases} 0, & z<0 \\ \frac{1}{2}z^2, & 0 \leqslant z < 1 \\ 1-\frac{1}{2}(2-z)^2, & 1 \leqslant z < 2 \\ 1, & z > 2 \end{cases}$$

$$f_{\xi+\eta}(z) = F'_{\xi+\eta}(z) = \begin{cases} 0, & z<0 \\ z, & 0 \leqslant z < 1 \\ 2-z, & 1 \leqslant z < 2 \\ 0, & z > 2 \end{cases}$$

Полученная плотность очевидно не является константой, значит полученное распределение не является равномерным. Оно называется треугольным распределением Симпсона. ■

Условное распределение

Def 1.12.16. Условным распределением случайной величины из системы случайных величин $\langle \xi, \eta \rangle$ называется ее распределение найденное при условии, что другая случайная величина приняла определенное значение.
Обозначается $\xi \mid \eta = y$, $\eta \mid \xi = x$.

Def 1.12.17. Условным математическим ожиданием (УМО) $\mathbb{E}\left(\xi \mid \eta = y\right)$ называется математическое ожидание случайной величины ξ при соответствующем условном распределении.

I. Условное распределение в дискретной системе двух случайных величин

Пусть двумерная случайная величина $\langle \xi, \eta \rangle$ задана законом распределения:

$\xi \backslash \eta$	y_1	y_2	$\dots$	y_m
x_1	$p_{1,1}$	$p_{1,2}$	$\dots$	$p_{1,m}$
x_2	$p_{2,1}$	$p_{2,2}$	$\dots$	$p_{2,m}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_n	$p_{n,1}$	$p_{n,2}$	$\dots$	$p_{n,m}$

x_i — значения случайной величины ξ, y_j — значения случайной величины η, $p_{i,j}$ — вероятность появления точки (x_i, y_j).
При составлении условных распределений используем формулы, вытекающие из формулы условной вероятности:

$$\xi \mid \eta = y_j: \qquad p_i = \mathsf{P}\left(\xi = x_i \mid \eta = y_j\right) = \frac{\mathsf{P}\left(\xi = x_i, \eta = y_j\right)}{\mathsf{P}\left(\eta = y_j\right)} = \frac{p_{i,j}}{q_j} \qquad \text{вероятности } i\text{-той строки делим на их сумму}$$

$$\eta \mid \xi = x_i: \qquad q_j = \mathsf{P}\left(\eta = y_j \mid \xi = x_i\right) = \frac{\mathsf{P}\left(\eta = y_j, \xi = x_i\right)}{\mathsf{P}\left(\eta = x_i\right)} = \frac{p_{i,j}}{p_i} \qquad \text{вероятности } j\text{-ого столбца делим на их сумму}$$

II. Условное распределение в непрерывной системе двух случайных величин

Пусть имеется двумерная абсолютно непрерывная случайная величина $\langle \xi, \eta \rangle$, которая задана плотностью совместного распределения $f_{\xi,\eta}(x, y)$. Тогда плотность условного распределения $\xi \mid \eta = y$ будет равна

$$f(x \mid y) = \frac{f_{\xi,\eta}(x, y)}{\int_{-\infty}^{\infty} f_{\xi,\eta}(x, y = const)\mathrm{d}x} = \frac{f_{\xi,\eta}(x, y)}{f_\eta(y)}$$

Аналогично

$$f(y \mid x) = \frac{f_{\xi,\eta}(x, y)}{f_\xi(x)}$$

Def 1.12.18. Функция

$$f(x \mid y) = \frac{f_{\xi,\eta}(x, y)}{f_\eta(y)}$$

называется условной плотностью.

Замечание 1.12.19. Условное математическое ожидание вычисляется по формуле

$$\mathbb{E}\left(\xi \mid \eta = y\right) = \int_{-\infty}^{\infty} x f(x \mid y)\mathrm{d}x \qquad \mathbb{E}\left(\eta \mid \xi = x\right) = \int_{-\infty}^{\infty} y f(y \mid x)\mathrm{d}y$$

Это можно рассматривать как обобщение формулы Байеса.

Замечание 1.12.20. При фиксированном значении переменной x условная плотность зависит только от y, а условное математическое ожидание $\mathbb{E}\left(\eta \mid \xi = x\right)$ — число. Если считать x переменной, то условное математическое ожидание является функцией, зависящей от x, и называется функцией регрессии η на ξ.

Замечание 1.12.21. Т.к. η случайная величина, то условное математическое ожидание $\mathbb{E}\left(\xi \mid \eta\right)$ также можно считать случайной величиной.

Пример 1.12.22. Пусть даны независимые случайные величины $\xi_1, \dots, \xi_n \in \Pi_\lambda$. Обозначим $S_n = \xi_1 + \ldots + \xi_n$. Найти $\mathbb{E}\left(\xi_1 \mid S_n = m\right)$.
Решение: Найдем условное распределение. Пусть $S_n = m$, тогда при $k \leqslant m$

$$\mathsf{P}\left(\xi_1 = k \mid S_n = m\right) = \frac{\mathsf{P}\left(\xi_1 = k, S_n = m\right)}{\mathsf{P}\left(S_n = m\right)} = \frac{\mathsf{P}\left(\xi_1 = k, \xi_2 + \ldots + \xi_n = m - k\right)}{\mathsf{P}\left(S_n = m\right)} = \frac{\mathsf{P}\left(\xi_1 = k\right)\mathsf{P}\left(\xi_2 + \ldots + \xi_n = m - k\right)}{\mathsf{P}\left(S_n = m\right)}$$

Распределение Пуассона устойчиво относительно суммирования, поэтому $\begin{cases} S_n \in \Pi_{n\lambda} \\ \xi_2 + \ldots + \xi_n \in \Pi_{(n-1)\lambda} \end{cases}$

$$\begin{aligned}
\mathsf{P}(\xi_1 = k \mid S_n = m) &= \frac{\lambda^k}{k!}e^{-\lambda} \cdot \frac{((n-1)\lambda)^{m-k}}{(m-k)!}e^{-(n-1)\lambda} \cdot \frac{m!}{(n\lambda)^m}e^{n\lambda} \\
&= \frac{\lambda^k}{k!}e^{-\lambda} \cdot \frac{(n-1)^{m-k}\lambda^{m-k}}{(m-k)!}e^{-n\lambda+\lambda} \cdot \frac{m!}{n^m\lambda^m}e^{n\lambda} \\
&= \frac{m!}{k!(m-k)!} \cdot \frac{(n-1)^{m-k}}{n^m} \\
&= C_m^k \left(\frac{1}{n}\right)^k \left(1-\frac{1}{n}\right)^{m-k}
\end{aligned}$$

Таким образом

$$\xi_1 \mid S_n = m \in B_{m,\frac{1}{n}} \quad \Longrightarrow \quad \mathbb{E}(\xi_1 \mid S_n = m) = \frac{m}{n}$$

Решение: Приведем другое решение. В силу симметрии

$$\mathbb{E}(\xi_1 \mid S_n = m) = \mathbb{E}(\xi_2 \mid S_n = m) = \ldots = \mathbb{E}(\xi_n \mid S_n = m)$$

По формуле полной вероятности получаем

$$n\mathbb{E}(\xi_1 \mid S_n = m) = \mathbb{E}(\xi_1 \mid S_n = m) + \ldots + \mathbb{E}(\xi_n \mid S_n = m) = \mathbb{E}(\xi_1 + \ldots + \xi_n \mid S_n = m) = \mathbb{E}(S_n \mid S_n = m) = m$$

Откуда $\mathbb{E}(\xi_1 \mid S_n = m) = \frac{m}{n}$.

1.13. Лекция 23.11.28.

Пространство случайных величин

Замечание 1.13.1. Если случайные величины ξ и η равны «почти наверное» ($\mathsf{P}(\xi = \eta) = 1$), то будем считать, что $\xi = \eta$.

Пусть дано вероятностное пространство $\langle \Omega, \mathcal{F}, \mathsf{P} \rangle$. Введем пространство L_2 как множество случайных величин с конечным вторым моментом.

$$L_2(\Omega, \mathcal{F}, \mathsf{P}) = \left\{ \xi \,\middle|\, \mathbb{D}(\xi) < \infty \right\}$$

Ясно, что это линейное пространство.

Def 1.13.2. Скалярным произведением случайных величин $\xi, \eta \in L_2$ называется $(\xi; \eta) = \mathbb{E}(\xi\eta)$.

Замечание 1.13.3. Если $\langle \xi, \eta \rangle$ имеет дискретное распределение, то $\mathbb{E}(\xi\eta) = \sum_{i,j} x_i y_j p_{i,j}$.
Если $\langle \xi, \eta \rangle$ имеет абсолютно непрерывное распределение с плотностью $f_{\xi,\eta}(x,y)$, то $\mathbb{E}(\xi\eta) = \iint_{\mathbb{R}^2} xy f_{\xi,\eta}(x,y) \mathrm{d}x\mathrm{d}y$.

Замечание 1.13.4. Т.к. дисперсии конечны, то данное скалярное произведение существует и конечно.

$$\left.\begin{array}{l}
(\xi; \eta) = (\eta; \xi) \\
(C\xi; \eta) = C(\xi; \eta) \\
(\xi_1 + \xi_2; \eta) = (\xi_1; \eta) + (\xi_2; \eta) \\
(\xi; \xi) = \mathbb{E}(\xi^2) \geqslant 0 \\
(\xi; \xi) = 0 \Longrightarrow \xi = 0 \text{ п.н.}
\end{array}\right\} \Longrightarrow \mathbb{E}(\xi\eta) \text{ действительно скалярное произведение}$$

Def 1.13.5. Норма случайной величины определяется стандартно как корень из скалярного квадрата: $\|\xi\| = \sqrt{(\xi; \xi)}$.

Замечание 1.13.6. Расстояние между случайными величинами (метрику) определим как норму разности: $\|\xi - \eta\|$.

Замечание 1.13.7. Таким образом рассматриваемое пространство является Гильбертовым (и даже Банаховым).

Теорема 1.13.8. (Неравенство Коши-Буняковского-Шварца) Пусть случайные величины ξ и η имеют конечные вторые моменты, тогда

$$|\mathbb{E}(\xi\eta)| \leqslant \sqrt{\mathbb{E}(\xi^2)\,\mathbb{E}(\eta^2)}$$

Причем равенство достигается тогда и только тогда, когда $\xi = C\eta$.

□ Доказательство аналогично приведенному ранее в курсе линейной алгебры. Рассмотрим квадратный трехчлен:

$$P_2(x) = \mathbb{E}\left((x\xi - \eta)^2\right) = x^2\mathbb{E}(\xi^2) - 2x\mathbb{E}(\xi\eta) + \mathbb{E}(\eta^2)$$

Т.к. $\forall x$ это трехчлен неотрицательный, то его дискриминант неположительный. Имеем

$$D = 4\mathbb{E}\left(\xi\eta\right) - 4\mathbb{E}\left(\xi^2\right)\mathbb{E}\left(\eta^2\right) \leqslant 0 \Longrightarrow \left|\mathbb{E}\left(\xi\eta\right)\right| \leqslant \sqrt{\mathbb{E}\left(\xi^2\right)\mathbb{E}\left(\eta^2\right)}$$

Далее покажем, когда достигается равенство.

$$\left|\mathbb{E}\left(\xi\eta\right)\right| = \sqrt{\mathbb{E}\left(\xi^2\right)\mathbb{E}\left(\eta^2\right)} \Longleftrightarrow D = 0 \Longrightarrow \exists C = const \,\Big|\, \mathbb{E}\left((C\xi - \eta)^2\right) = 0 \Longrightarrow C\xi = \eta \text{ п.н.}$$

■

Условное математическое ожидание

Рассмотрим линейное подпространство, образованное множеством случайных величин $g(\eta)$ (функция g — борелевская) с конечным вторым моментом.

$$L = L(\eta) = \left\{g(\eta) \,\Big|\, \mathbb{D}\left(g(\eta)\right) < \infty\right\} \subset L_2(\Omega, \mathcal{F}, \mathsf{P})$$

Def 1.13.9. Условным математическим ожиданием (УМО) $\mathbb{E}\left(\xi \mid \eta\right)$ называется ортогональная проекция случайной величины ξ на подпространство $L(\eta)$.

Далее будем обозначать $\hat{\xi} = \mathbb{E}\left(\xi \mid \eta\right)$.

Lm 1.13.10 (Тождество ортопроекции)**.** Пусть $\hat{\xi} \in L(\eta)$, тогда

$$\hat{\xi} = \mathbb{E}\left(\xi \mid \eta\right) \Longleftrightarrow \mathbb{E}\left(\xi g(\eta)\right) = \mathbb{E}\left(\hat{\xi} g(\eta)\right) \qquad \forall g(\eta) \in L(\eta)$$

□

$$\hat{\xi} = \mathbb{E}\left(\xi \mid \eta\right) \Longleftrightarrow \xi - \hat{\xi} \perp g(\eta) \in L(\eta) \qquad \forall g(\eta)$$

$$\left(\xi - \hat{\xi}; g(\eta)\right) = \mathbb{E}\left((\xi - \hat{\xi})g(\eta)\right) = 0 \Longleftrightarrow \mathbb{E}\left(\xi g(\eta)\right) = \mathbb{E}\left(\hat{\xi} g(\eta)\right) \qquad \forall g(\eta) \in L(\eta)$$

■

Lm 1.13.11 (Формула полного математического ожидания)**.**

$$\mathbb{E}\left(\xi\right) = \mathbb{E}\left(\mathbb{E}\left(\xi \mid \eta\right)\right) = \mathbb{E}\left(\hat{\xi}\right)$$

□ Искомое равенство можно легко получить, если в 1.13.10 положить $g(\eta) \equiv 1$. ■

Замечание 1.13.12. В случае распределения Бернулли эта формула является уже знакомой формулой полной вероятности.

Lm 1.13.13 (Линейность)**.**

$$\mathbb{E}\left(C_1\xi_1 + C_2\xi_2 \mid \eta\right) = C_1\mathbb{E}\left(\xi_1 \mid \eta\right) + C_2\mathbb{E}\left(\xi_2 \mid \eta\right)$$

Lm 1.13.14. Если случайные величины ξ и η независимы, то

$$\mathbb{E}\left(\xi \mid \eta\right) = \mathbb{E}\left(\xi\right)$$

□ Проверим тождество ортопроекции.

$$\mathbb{E}\left(\mathbb{E}\left(\xi\right) g(\eta)\right) = \mathbb{E}\left(\xi\right)\mathbb{E}\left(g(\eta)\right) \overset{\text{независимость}}{=\!=\!=\!=\!=} \mathbb{E}\left(\xi g(\eta)\right)$$

Тождество выполняется, значит $\mathbb{E}\left(\xi\right) = \mathbb{E}\left(\xi \mid \eta\right)$. ■

Lm 1.13.15. Если случайные величины ξ и η независимы, то $\xi - \mathbb{E}\left(\xi\right) \perp L(\eta)$ и, в частности, $\xi - \mathbb{E}\left(\xi\right) \perp \eta$.

Покажем, что данное определение УМО согласуется с введенным ранее понятием математического ожидания условного распределения, т.е. $\mathbb{E}\left(\xi \mid \eta\right) = h(\eta)$, где $h(y) = \mathbb{E}\left(\xi \mid \eta = y\right)$. Рассмотрим случай абсолютно непрерывной системы случайных величин $\langle\xi, \eta\rangle$ с плотностью $f_{\xi,\eta}(x, y)$. Тогда

$$h(y) = \int_{-\infty}^{\infty} x f(x \mid y)\mathrm{d}x \qquad f(x \mid y) = \frac{f_{\xi,\eta}(x, y)}{f_\eta(y)}$$

Следует проверить, что $h(y)$ удовлетворяет тождеству ортопроекции, т.е. то, что $\mathbb{E}\left(\xi g(\eta)\right) = \mathbb{E}\left(h(\eta)g(\eta)\right)$.

$$\mathbb{E}\left(\xi g(\eta)\right) = \iint_{\mathbb{R}^2} x g(y) f_{\xi,\eta}(x, y)\mathrm{d}x\mathrm{d}y$$

$$\mathbb{E}\left(h(\eta)g(\eta)\right) = \int_{-\infty}^{\infty} h(y)g(y)f_\eta(y)\mathrm{d}y = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} x\frac{f_{\xi,\eta}(x, y)}{f_\eta(y)}\mathrm{d}x\right)g(y)f_\eta(y)\mathrm{d}y = \iint_{\mathbb{R}^2} x g(y) f_{\xi,\eta}(x, y)\mathrm{d}x\mathrm{d}y$$

Условная дисперсия

Def 1.13.16. Условной дисперсией случайной величины ξ относительно случайной величины η называется случайная величина

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

Т.е. это дисперсия соответствующего условного распределения.

Lm 1.13.17.

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

□ Доказательство аналогично доказательству такой же формулы для обычной дисперсии. ■

Теорема 1.13.18. (Закон полной дисперсии)

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

□ Известно, что $\mathbb{D}(\xi) = \mathbb{E}\left(\xi^2\right) - (\mathbb{E}(\xi))^2$. Применим к каждого слагаемому формулу полного математического ожидания.

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

Из 1.13.17 можно получить, что

$$\mathbb{E}\left(\xi^2 \mid \eta\right) = \mathbb{D}(\xi \mid \eta) + (\mathbb{E}(\xi \mid \eta))^2 \tag{2}$$

Подставим (2) в (1).

$$\begin{aligned}
\mathbb{D}(\xi) &= \mathbb{E}\left(\mathbb{D}(\xi \mid \eta) + (\mathbb{E}(\xi \mid \eta))^2\right) - (\mathbb{E}(\mathbb{E}(\xi \mid \eta)))^2 \\
&= \mathbb{E}(\mathbb{D}(\xi \mid \eta)) + \underbrace{\mathbb{E}\left((\mathbb{E}(\xi \mid \eta))^2\right) - (\mathbb{E}(\mathbb{E}(\xi \mid \eta)))^2}_{\mathbb{D}(\mathbb{E}(\xi \mid \eta))} \\
&= \mathbb{E}(\mathbb{D}(\xi \mid \eta)) + \mathbb{D}(\mathbb{E}(\xi \mid \eta))
\end{aligned}$$

■

Замечание 1.13.19. Пусть ξ и η независимы, тогда

$$\mathbb{D}(\mathbb{E}(\xi \mid \eta)) = \mathbb{D}(\mathbb{E}(\xi)) = 0 \Longrightarrow \mathbb{D}(\xi) = \mathbb{E}(\mathbb{D}(\xi \mid \eta))$$

Замечание 1.13.20. Если имеется функциональная зависимость $\xi = g(\eta)$, то

$$\mathbb{D}(\mathbb{E}(\xi \mid \eta)) = \mathbb{D}(\mathbb{E}(g(\eta) \mid \eta)) = \mathbb{D}(g(\eta)) = \mathbb{D}(\xi)$$

Таким образом доля

$$0 \leqslant R^2 = \frac{\mathbb{D}(\mathbb{E}(\xi \mid \eta))}{\mathbb{D}(\xi)} \leqslant 1$$

отражает уровень зависимости случайной величины ξ от случайной величины η.

Def 1.13.21. Величина R называется корреляционным отношением.

Числовые характеристики зависимости случайных величин

Если ξ и η независимы, то $\mathbb{E}(\xi\eta) = \mathbb{E}(\xi)\,\mathbb{E}(\eta)$, поэтому в качестве индикатора наличия связи логично взять величину $\mathbb{E}(\xi\eta) - \mathbb{E}(\xi)\,\mathbb{E}(\eta) = \operatorname{cov}(\xi, \eta)$.

Def 1.13.22. Ковариацией случайных величин ξ и η называется число

$$\operatorname{cov}(\xi, \eta) = \mathbb{E}((\xi - \mathbb{E}(\xi))(\eta - \mathbb{E}(\eta)))$$

Lm 1.13.23 (Равносильность определений)**.**

$$\operatorname{cov}(\xi, \eta) = \mathbb{E}((\xi - \mathbb{E}(\xi))(\eta - \mathbb{E}(\eta))) = \mathbb{E}(\xi\eta) - \mathbb{E}(\xi)\,\mathbb{E}(\eta)$$

□

$$\begin{aligned}
\operatorname{cov}(\xi, \eta) &= \mathbb{E}((\xi - \mathbb{E}(\xi))(\eta - \mathbb{E}(\eta))) \\
&= \mathbb{E}(\xi\eta - \eta\mathbb{E}(\xi) - \xi\mathbb{E}(\eta) + \mathbb{E}(\xi)\,\mathbb{E}(\eta)) \\
&= \mathbb{E}(\xi\eta) - \mathbb{E}(\xi)\,\mathbb{E}(\eta) - \mathbb{E}(\xi)\,\mathbb{E}(\eta) + \mathbb{E}(\xi)\,\mathbb{E}(\eta) \\
&= \mathbb{E}(\xi\eta) - \mathbb{E}(\xi)\,\mathbb{E}(\eta)
\end{aligned}$$

■

Свойства ковариации

1. $\operatorname{cov}(\xi, \xi) = \mathbb{E}\left(\xi^2\right) - \left(\mathbb{E}(\xi)\right)^2 = \mathbb{D}(\xi)$
2. $\operatorname{cov}(\xi, \eta) = \operatorname{cov}(\eta, \xi)$
3. $\operatorname{cov}(C\xi, \eta) = C \operatorname{cov}(\xi, \eta)$
4. $\operatorname{cov}(\xi_1 + \xi_2, \eta) = \operatorname{cov}(\xi_1, \eta) + \operatorname{cov}(\xi_2, \eta)$
5. $\mathbb{D}(\xi_1 + \ldots + \xi_n) = \mathbb{D}(\xi_1) + \cdots + \mathbb{D}(\xi_n) + 2\sum\limits_{i<j} \operatorname{cov}(\xi_i, \xi_j) = \sum\limits_{i,j} \operatorname{cov}(\xi_i, \xi_j)$
6. Если ξ и η независимы, то $\operatorname{cov}(\xi, \eta) = 0$.
7. Если $\operatorname{cov}(\xi, \eta) \neq 0$, то ξ и η — зависимы.
8. Если $\operatorname{cov}(\xi, \eta) = 0$, то нельзя сделать никаких выводов.
9. Если $\operatorname{cov}(\xi, \eta) > 0$, то зависимость прямая.
10. Если $\operatorname{cov}(\xi, \eta) < 0$, то зависимость обратная.

Замечание 1.13.24. По величине ковариации нельзя судить о силе связи, т.к. она зависит от единиц измерения. Ее следует «нормировать».

Коэффициент линейной корреляции

Def 1.13.25. Коэффициентом линейной корреляции $\rho_{\xi,\eta}$ случайных величин ξ и η с конечным вторым моментом называется число

$$\rho_{\xi,\eta} = \frac{\operatorname{cov}(\xi, \eta)}{\sqrt{\mathbb{D}(\xi)}\sqrt{\mathbb{D}(\eta)}} = \frac{\mathbb{E}(\xi\eta) - \mathbb{E}(\xi)\,\mathbb{E}(\eta)}{\sigma(\xi)\,\sigma(\eta)}$$

Замечание 1.13.26. Эта безразмерная величина, поэтому по ней можно судить о силе связи.

Замечание 1.13.27.

$$\rho_{\xi,\eta} = \frac{\mathbb{E}\left(\left(\xi - \mathbb{E}(\xi)\right)\left(\eta - \mathbb{E}(\eta)\right)\right)}{\sqrt{\mathbb{E}\left(\xi - \mathbb{E}(\xi)^2\right)}\sqrt{\mathbb{E}\left(\eta - \mathbb{E}(\eta)^2\right)}} = \frac{\left(\xi - \mathbb{E}(\xi)\,; \eta - \mathbb{E}(\eta)\right)}{\|\xi - \mathbb{E}(\xi)\| \cdot \|\eta - \mathbb{E}(\eta)\|} = \cos\angle\Big(\xi - \mathbb{E}(\xi)\,, \eta - \mathbb{E}(\eta)\Big)$$

Свойства корреляции

1. $\rho_{\xi,\eta} = \rho_{\eta,\xi}$
2. Если ξ и η независимы, то $\rho_{\xi,\eta} = 0$.
3. Если $\rho_{\xi,\eta} \neq 0$, то ξ и η — зависимы.
4. Если $\rho_{\xi,\eta} = 0$, то нельзя сделать никаких выводов.

Lm 1.13.28.

$$|\rho_{\xi,\eta}| \leqslant 1$$

□ По 1.13.8 имеем

$$\left|\mathbb{E}\left(\left(\xi - \mathbb{E}(\xi)\right)\left(\eta - \mathbb{E}(\eta)\right)\right)\right| \leqslant \sqrt{\left(\xi - \mathbb{E}(\xi)\right)^2\left(\eta - \mathbb{E}(\eta)\right)^2}$$

Подставив это в определение корреляции получаем искомую оценку. ■

Lm 1.13.29.

$$|\rho_{\xi,\eta}| = 1 \Longleftrightarrow \eta = C\xi + b$$

□ По 1.13.8 имеем

$$\begin{aligned}|\rho_{\xi,\eta}| = 1 \Longleftrightarrow \left|\mathbb{E}\left(\left(\xi - \mathbb{E}(\xi)\right)\left(\eta - \mathbb{E}(\eta)\right)\right)\right| &= \sqrt{\left(\xi - \mathbb{E}(\xi)\right)^2\left(\eta - \mathbb{E}(\eta)\right)^2} \\ \eta - \mathbb{E}(\eta) = C\left(\xi - \mathbb{E}(\xi)\right) \Longleftrightarrow \eta &= C\xi + \underbrace{\mathbb{E}(\eta) - C\mathbb{E}(\xi)}_{b}\end{aligned}$$

■

Lm 1.13.30.

$$\rho_{\xi,\eta} = 1 \Longrightarrow \eta = a\xi + b \qquad a > 0 \text{ (прямая зависимость)}$$
$$\rho_{\xi,\eta} = -1 \Longrightarrow \eta = a\xi + b \qquad a < 0 \text{ (обратная зависимость)}$$

□ По 1.13.29 получаем, что $\eta = a\xi + b$. Подставим это в определение корреляции.

$$\begin{aligned}
\rho_{\xi,\eta} &= \frac{\mathbb{E}(\xi\eta) - \mathbb{E}(\xi)\mathbb{E}(\eta)}{\sqrt{\mathbb{D}(\xi)}\sqrt{\mathbb{D}(\eta)}} \\
&= \frac{\mathbb{E}(\xi(a\xi+b)) - \mathbb{E}(\xi)\mathbb{E}(a\xi+b)}{\sqrt{\mathbb{D}(\xi)}\sqrt{\mathbb{D}(a\xi+b)}} \\
&= \frac{a\mathbb{E}(\xi^2) + b\mathbb{E}(\xi) - a(\mathbb{E}(\xi))^2 - b\mathbb{E}(\xi)}{\sqrt{\mathbb{D}(\xi)}\sqrt{a^2\mathbb{D}(\xi)}} \\
&= \frac{a\left(\mathbb{E}(\xi^2) - (\mathbb{E}(\xi))^2\right)}{|a|\,\mathbb{D}(\xi)} \\
&= \frac{a\mathbb{D}(\xi)}{|a|\,\mathbb{D}(\xi)} \\
&= \operatorname{sign} a
\end{aligned}$$

■

Def 1.13.31. Если $\rho_{\xi,\eta} \neq 0$, то говорят, что случайные величины ξ и η коррелированны друг с другом. Если $\rho_{\xi,\eta}$ больше нуля, то имеется прямая корреляция между случайными величинами ξ и η, а если меньше нуля, то обратная корреляция.

Замечание 1.13.32. Коэффициент корреляции отражает преимущественно силу линейной части связи, а саму силу лучше отражает корреляционное отношение.

Теорема 1.13.33.

$$R \geqslant \rho_{\xi,\eta}$$

Замечание 1.13.34. $R = \rho_{\xi,\eta}$ тогда и только тогда, когда уравнение регрессии является линейным.

Пример 1.13.35. Пусть между случайными величинами ξ_1 и ξ_2 прямая корреляция, между ξ_2 и ξ_3 также прямая корреляция. Однако из этого **не следует**, что между ξ_1 и ξ_3 будет прямая корреляция.

1.14. Лекция 23.12.05.

Характеристические функции

Пусть $\xi + i\eta$ — комплексная случайная величина, причем ξ и η это обычные случайные величины с конечным первым моментом.

Def 1.14.1. Математическое ожидание комплекснозначной случайной величины определим как

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

Def 1.14.2. Характеристической функцией случайной величины ξ называется функция

$$\varphi_\xi(t) = \mathbb{E}\left(e^{i\xi t}\right) \qquad t \in \mathbb{R}$$

Lm 1.14.3. Характеристическая функция существует для любой случайной величины ξ, причем $|\varphi_\xi(t)| \leqslant 1$.

□ Заметим, что

$$\mathbb{D}(\xi) = \mathbb{E}\left(\xi^2\right) - (\mathbb{E}(\xi))^2 \geqslant 0 \Longrightarrow (\mathbb{E}(\xi))^2 \leqslant \mathbb{E}\left(\xi^2\right) \tag{1}$$

Оценим модуль характеристической функции.

$$\begin{aligned}
|\varphi_\xi(t)|^2 &= \left|\mathbb{E}\left(e^{i\xi t}\right)\right|^2 \\
&= |\mathbb{E}(\cos(\xi t) + i\sin(\xi t))|^2 \\
&= |\mathbb{E}(\cos(\xi t)) + i\mathbb{E}(\sin(\xi t))|^2 \\
&= |\mathbb{E}(\cos(\xi t))|^2 + |\mathbb{E}(\sin(\xi t))|^2
\end{aligned} \tag{2}$$

Применим к полученному в (2) неравенство из (1), получим

$$|\varphi_\xi(t)|^2 \leqslant \mathbb{E}\left(\cos^2(\xi t)\right) + \mathbb{E}\left(\sin^2(\xi t)\right) = \mathbb{E}\left(\cos^2(\xi t) + \sin^2(\xi t)\right) = \mathbb{E}(1) = 1$$

■

Lm 1.14.4. Пусть $\varphi_\xi(t)$ это характеристическая функция случайной величины ξ. Тогда характеристическая функция случайной величины $\eta = a + b\xi$ будет иметь вид

$$\varphi_{a+b\xi}(t) = e^{ita} \cdot \varphi_\xi(bt)$$

□

$$\varphi_{a+b\xi}(t) = \mathbb{E}\left(e^{i(a+b\xi)t}\right) = \mathbb{E}\left(e^{iat} \cdot e^{ib\xi t}\right) = e^{iat}\mathbb{E}\left(e^{i(bt)\xi}\right) = e^{ita} \cdot \varphi_\xi(bt)$$

■

Lm 1.14.5. Характеристическая функция суммы независимых случайных величин равна произведению их характеристических функций.

□ Пусть случайные величины ξ и η независимы, тогда

$$\varphi_{\xi+\eta}(t) = \mathbb{E}\left(e^{i(\xi+\eta)t}\right) = \mathbb{E}\left(e^{i\xi t} \cdot e^{i\eta t}\right) = \mathbb{E}\left(e^{i\xi t}\right) \cdot \mathbb{E}\left(e^{i\eta t}\right) = \varphi_\xi(t)\varphi_\eta(t)$$

■

Lm 1.14.6. Пусть $\mathbb{E}\left(\xi^k\right) < \infty$. Тогда

$$\varphi_\xi(t) = 1 + it \cdot \mathbb{E}(\xi) - \frac{t^2}{2}\mathbb{E}\left(\xi^2\right) + \ldots + \frac{i^k \mathbb{E}\left(\xi^k\right)}{k!} t^k + o\left(|t|^k\right)$$

□ Разложим экспоненту в ряд Маклорена и воспользуемся свойствами математического ожидания.

$$\varphi_\xi(t) = \mathbb{E}\left(e^{i\xi t}\right) = \mathbb{E}\left(1 + i\xi t + \frac{(i\xi t)^2}{2!} + \ldots + \frac{(i\xi t)^k}{k!} + o\left(|t|^k\right)\right) = 1 + it \cdot \mathbb{E}(\xi) - \frac{t^2}{2}\mathbb{E}\left(\xi^2\right) + \ldots + \frac{i^k \mathbb{E}\left(\xi^k\right)}{k!} t^k + o\left(|t|^k\right)$$

■

Lm 1.14.7. Пусть $\mathbb{E}\left(\xi^k\right) < \infty$. Тогда характеристическая функция непрерывно дифференцируема k раз, причем

$$\varphi_\xi^{(k)}(0) = i^k \mathbb{E}\left(\xi^k\right)$$

□ Это следует из существования k-ого члена разложения характеристической функции в ряд Маклорена и равенства коэффициентов разложения:

$$\frac{\varphi_\xi^{(k)}(0)}{k!} = \frac{i^k \mathbb{E}\left(\xi^k\right)}{k!} \Longrightarrow \varphi_\xi^{(k)}(0) = i^k \mathbb{E}\left(\xi^k\right)$$

■

Замечание 1.14.8. Существует взаимнооднозначное соответствие между распределениями и характеристическими функциями. По характеристической функции можно восстановить распределение. В частности, если ξ абсолютно непрерывная случайная величина, то плотность можно найти по формуле

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-it\xi} \varphi_\xi(t) \mathrm{d}t$$

Это обратное преобразование Фурье.

Теорема 1.14.9. (О непрерывном соответствии) Последовательность случайных величин ξ_n слабо сходится к случайной величине ξ тогда и только тогда, когда соответствующая последовательность характеристических функций поточечно сходится к характеристической функции случайной величины ξ.

$$\xi_n \rightrightarrows \xi \Longleftrightarrow \varphi_{\xi_n}(t) \to \varphi_\xi(t) \qquad \forall t \in \mathbb{R}$$

Характеристические функции стандартных распределений

I. Распределение Бернулли

Пусть $\xi \in B_p$, тогда

$$\varphi_\xi(t) = \mathbb{E}\left(e^{i\xi t}\right) = (1-p)\mathbb{E}\left(e^{it\cdot 0}\right) + p\mathbb{E}\left(e^{it\cdot 1}\right) = 1 - p + pe^{it}$$

II. Биномиальное распределение

Пусть $\xi \in B_{n,p}$. Напомним, что

$$\mathsf{P}(\xi = k) = C_n^k p^k q^{n-k} \qquad k = 0, 1, \ldots, n$$

Тогда воспользуемся 1.14.5 и получим

$$\xi = \xi_1 + \ldots + \xi_n \qquad \xi_i \in B_p$$
$$\varphi_\xi(t) = (\varphi_{\xi_1}(t))^n = \left(1 - p + pe^{it}\right)^n$$

III. Распределение Пуассона

Пусть $\xi \in \Pi_\lambda$. Напомним, что

$$\mathsf{P}(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad k = 0, 1, \ldots$$

Тогда получим

$$\varphi_\xi(t) = \mathbb{E}\left(e^{i\xi t}\right) = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\left(\lambda e^{it}\right)^k}{k!} = e^{-\lambda} \cdot e^{\lambda e^{it}} = \exp\left(\lambda\left(e^{it} - 1\right)\right)$$

Lm 1.14.10. Распределение Пуассона устойчиво относительно суммирования.

□ Пусть даны независимые случайные величины $\xi \in \Pi_\lambda$ и $\eta \in \Pi_\mu$. Тогда по 1.14.5

$$\varphi_{\xi+\eta}(t) = \varphi_\xi(t)\varphi_\eta(t) = \exp\left(\lambda\left(e^{it} - 1\right)\right) \exp\left(\mu\left(e^{it} - 1\right)\right) = \exp\left((\lambda + \mu)\left(e^{it} - 1\right)\right)$$

Получили характеристическую функцию распределения $\Pi_{\xi+\eta}$. ■

IV. Показательное распределение

Пусть $\xi \in E_\alpha$. Напомним, что

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

Тогда получим

$$\varphi_\xi(t) = \mathbb{E}\left(e^{i\xi t}\right) = \int_0^{+\infty} e^{ixt} \alpha e^{-\alpha x} \mathrm{d}x = \alpha \int_0^{+\infty} e^{(it-\alpha)x} \mathrm{d}x = \frac{\alpha}{it - \alpha} e^{(it-\alpha)x}\Big|_0^{+\infty} = \frac{\alpha}{it - \alpha}(0 - 1) = \frac{\alpha}{\alpha - it}$$

V. Стандартное нормальное распределение

Пусть $\xi \in N(0; 1)$. Напомним, что

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \qquad x \in \mathbb{R}$$

Тогда получим

$$\begin{aligned}
\varphi_\xi(t) &= \mathbb{E}\left(e^{i\xi t}\right) \\
&= \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(x^2 - 2itx)\right) \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(x^2 - 2itx - t^2)\right) \exp\left(-\frac{t^2}{2}\right) \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \underbrace{\textstyle\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(x - it)^2\right) \mathrm{d}(x - it)}_{\text{Интеграл Пуассона}} \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \sqrt{2\pi} \\
&= \exp\left(-\frac{t^2}{2}\right)
\end{aligned}$$

VI. Нормальное распределение

Пусть $\xi \in N(a;\sigma^2)$. Нормируем ее и обозначим $\eta = \frac{\xi-a}{\sigma} \in N(0;1)$. Характеристическая функция для нее уже найдена ранее:

$$\varphi_\eta(t) = \exp\left(-\frac{t^2}{2}\right)$$

Т.к. $\xi = a + \eta\sigma$, то по 1.14.4 получаем, что

$$\varphi_\xi(t) = e^{iat}\exp\left(-\frac{(\sigma t)^2}{2}\right) = e^{iat}\exp\left(-\frac{1}{2}t^2\sigma^2\right)$$

<u>Lm</u> 1.14.11.

$$\left.\begin{array}{l} \xi \in N(a_1,\sigma_1^2) \\ \eta \in N(a_2,\sigma_2^2) \\ \xi \text{ и } \eta \text{ независимы} \end{array}\right\} \Longrightarrow \xi+\eta \in N(a_1+a_2,\sigma_1^2+\sigma_2^2)$$

□

$$\varphi_{\xi+\eta}(t) = \varphi_\xi(t)\varphi_\eta(t) = e^{ia_1t}\exp\left(-\frac{1}{2}t^2\sigma_1^2\right)\cdot e^{ia_2t}\exp\left(-\frac{1}{2}t^2\sigma_2^2\right) = e^{i(a_1+a_2)t}\exp\left(-\frac{1}{2}t^2(\sigma_1^2+\sigma_2^2)\right)$$

Получили характеристическую функцию нормального распределения с параметрами $a = a_1 + a_2$ и $\sigma^2 = \sigma_1^2 + \sigma_2^2$. ■

<u>Lm</u> 1.14.12.

$$\left(1 + \frac{x}{n} + o\left(\frac{1}{n}\right)\right)^n \xrightarrow{n\to\infty} e^x$$

□

$$\begin{aligned}\left(1 + \frac{x}{n} + o\left(\frac{1}{n}\right)\right)^n &= \exp\left(n\ln\left(1 + \frac{x}{n} + o\left(\frac{1}{n}\right)\right)\right) \\ &= \exp\left(n\left(\frac{x}{n} + o\left(\frac{1}{n}\right) + o\left(\frac{x}{n} + o\left(\frac{1}{n}\right)\right)\right)\right) \\ &= \exp\left(n\left(\frac{x}{n} + o\left(\frac{1}{n}\right)\right)\right) \\ &= \exp\left(x + n\cdot o\left(\frac{1}{n}\right)\right) \xrightarrow{n\to\infty} e^x\end{aligned}$$

■

Теорема 1.14.13. (Закон больших чисел Хинчина) Пусть $\xi_1,\ldots,\xi_n$ независимые одинаково распределенные случайные величины с конечным первым моментом $\mathbb{E}(\xi) = a < \infty$. Тогда

$$\frac{S_n}{n} = \frac{\xi_1 + \ldots + \xi_n}{n} \xrightarrow{\mathsf{P}} a$$

□ Сходимость по вероятности к константе эквивалентна слабой сходимости (1.10.9), поэтому достаточно доказать, что $\frac{S_n}{n} \rightrightarrows a$. По 1.14.9 эта сходимость имеет место, если

$$\varphi_{\frac{S_n}{n}}(t) \to \varphi_a(t) = \mathbb{E}\left(e^{ita}\right) = e^{ita} \qquad \forall t \in \mathbb{R}$$

По 1.14.4 и 1.14.5 имеем

$$\varphi_{\frac{S_n}{n}}(t) = \varphi_{S_n}\left(\frac{t}{n}\right) = \left(\varphi_{\xi_1}\left(\frac{t}{n}\right)\right)^n$$

Т.к. первый момент существует, то по 1.14.6 получаем

$$\varphi_{\xi_1} = 1 + it\mathbb{E}(\xi_1) + o(t) = 1 + ita + o(t)$$

Собираем все полученные формулы и применяем 1.14.12.

$$\varphi_{\frac{S_n}{n}}(t) = \left(1 + i\frac{t}{n}a + o\left(\frac{t}{n}\right)\right)^n = \left(1 + \frac{ita}{n} + o\left(\frac{1}{n}\right)\right)^n \xrightarrow{n\to\infty} e^{ita}$$

Получили характеристическую функцию случайной величины a. ■

Теорема 1.14.14. (Центральная предельная теорема) Пусть $\xi_1,\ldots,\xi_n$ последовательность независимых одинаково распределенных случайных величин с конечным вторым моментом. Обозначим $\mathbb{E}(\xi_i) = a$ и $\mathbb{D}(\xi_i) = \sigma^2$. Тогда

$$\frac{S_n - na}{\sigma\sqrt{n}} \rightrightarrows N(0;1)$$

□ Пусть $\eta_i = \frac{\xi_i - a}{\sigma}$ — последовательность соответствующих стандартизованных случайных величин, причем $\mathbb{E}\left(\eta_i\right) = 0$ и $\mathbb{D}\left(\eta_i\right) = 1$. Тогда

$$z_n = \eta_1 + \ldots + \eta_n = \frac{\sum \xi_i - na}{\sigma} = \frac{S_n - na}{\sigma}$$

Теперь надо доказать, что

$$\frac{z_n}{\sqrt{n}} = \frac{S_n - na}{\sigma\sqrt{n}} \rightrightarrows N(0;1)$$

По 1.14.4 и 1.14.5 имеем

$$\varphi_{\frac{z_n}{\sqrt{n}}}(t) = \varphi_{z_n}\left(\frac{t}{\sqrt{n}}\right) = \left(\varphi_{\eta_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n$$

Т.к. второй момент существует, то по 1.14.6 получаем

$$\varphi_{\eta_1} = 1 + it\underbrace{\mathbb{E}\left(\eta_1\right)}_{=0} - \frac{t^2}{2}\underbrace{\mathbb{E}\left(\eta_1^2\right)}_{=1} + o\left(t^2\right) = 1 - \frac{t^2}{2} + o\left(t^2\right)$$

Подставим это в предыдущую формулу

$$\varphi_{\frac{z_n}{\sqrt{n}}}(t) = \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n = \left(1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n$$

Применяем 1.14.12.

$$\varphi_{\frac{z_n}{\sqrt{n}}}(t) \xrightarrow{n\to\infty} \exp\left(-\frac{t^2}{2}\right)$$

Получили характеристическую функцию стандартного нормального распределения и по 1.14.9 имеем

$$\frac{z_n}{\sqrt{n}} = \frac{S_n - na}{\sigma\sqrt{n}} \rightrightarrows N(0;1)$$

■

Теорема 1.14.15. (Предельная теорема Муавра—Лапласа) Пусть $v_n(A)$ — число появления события A при n независимых испытаниях, p — вероятность успеха при одном испытании, а $q = 1 - p$. Тогда

$$\frac{v_n(A) - np}{\sqrt{npq}} \rightrightarrows N(0;1) \qquad (n \to \infty)$$

□

$$v_n(A) = \xi_1 + \ldots + \xi_n \qquad \xi_i \in B_p \text{ это число успехов при } i\text{-ом испытании}$$
$$\mathbb{E}\left(\xi_1\right) = p \qquad \mathbb{D}\left(\xi_1\right) = pq \qquad \sigma\left(\xi_1\right) = \sqrt{pq}$$

По 1.14.14 имеем

$$\frac{v_n(A) - np}{\sqrt{pq}\sqrt{n}} = \frac{v_n(A) - np}{\sqrt{npq}} \rightrightarrows N(0;1)$$

■

Теорема 1.14.16. (Интегральная формула Муавра—Лапласа)

$$\mathsf{P}(k_1 \leqslant v_n(A) \leqslant k_2) \approx \Phi(x_2) - \Phi(x_1)$$

□

$$\mathsf{P}(k_1 \leqslant v_n(A) \leqslant k_2) = \mathsf{P}\left(\frac{k_1 - np}{\sqrt{npq}} \leqslant \frac{v_n(A) - np}{\sqrt{npq}} \leqslant \frac{k_2 - np}{\sqrt{npq}}\right) = F_\eta\left(\frac{k_2 - np}{\sqrt{npq}}\right) - F_\eta\left(\frac{k_1 - np}{\sqrt{npq}}\right) \qquad \eta = \frac{v_n(A) - np}{\sqrt{npq}}$$

По 1.14.15 получаем, что

$$F_\eta\left(\frac{k_2 - np}{\sqrt{npq}}\right) - F_\eta\left(\frac{k_1 - np}{\sqrt{npq}}\right) \rightrightarrows F_0\left(\frac{k_2 - np}{\sqrt{npq}}\right) - F_0\left(\frac{k_1 - np}{\sqrt{npq}}\right)$$

где

$$F_0 = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) \mathrm{dt}$$

это функция стандартного нормального распределения. ■

Замечание 1.14.17. Аналогичным образом Центральную Предельную Теорему применяют для приближенного вычисления вероятностей связанных с суммами большого числа независимых одинаково распределенных случайных величин, заменяя стандартизованную сумму на стандартное нормальное распределение. Но какова ошибка этого приближения? Для этого используют следующую теорему.

Теорема 1.14.18. (Неравенство Берри—Эссеена) В условиях Центральной Предельной Теоремы для случайных величин с конечным третьим моментом справедливо

$$\left|\mathsf{P}\left(\frac{S_n - n\mathbb{E}\left(\xi_1\right)}{\sqrt{n\mathbb{D}\left(\xi_1\right)}} < x\right) - F_0(x)\right| \leqslant C\frac{\mathbb{E}\left(\left|\xi_1 - \mathbb{E}\left(\xi_1\right)\right|^3\right)}{\sqrt{n}\left(\sqrt{\mathbb{D}\left(\xi_1\right)}\right)^3} \qquad \forall x \in \mathbb{R}$$

Замечание 1.14.19. По теории $C < 0.9$, однако на практике обычно можно брать 0.4.

1.15. Лекция 23.12.12.

Производящая функция моментов

Def 1.15.1. Производящей функцией моментов случайной величины ξ называется функция

$$M_\xi(t) = \mathbb{E}\left(e^{\xi t}\right)$$

Замечание 1.15.2. Если $\mathbb{E}\left(\xi^k\right) < \infty$, то

$$M_\xi(t) = \mathbb{E}\left(e^{\xi t}\right) = \mathbb{E}\left(1 + \xi t + \frac{\xi^2 t^2}{2} + \ldots + \frac{\xi^k t^k}{k!} + o\left(t^k\right)\right) = 1 + t\mathbb{E}\left(\xi\right) + \frac{t^2}{2}\mathbb{E}\left(\xi^2\right) + \ldots + \frac{t^k}{k!}\mathbb{E}\left(\xi^k\right) + o\left(t^k\right)$$

Причем

$$\mathbb{E}\left(\xi^k\right) = M_\xi^{(k)}(0)$$

Lm 1.15.3. Если случайные величины ξ и η независимы, то

$$M_{\xi+\eta}(t) = M_\xi(t) \cdot M_\eta(t)$$

□

$$M_{\xi+\eta}(t) = \mathbb{E}\left(e^{(\xi+\eta)t}\right) = \mathbb{E}\left(e^{\xi t} e^{\eta + t}\right) = \mathbb{E}\left(e^{\xi t}\right)\mathbb{E}\left(e^{\eta t}\right) = M_\xi(t) M_\eta(t)$$

■

Замечание 1.15.4. Если $M_\xi(t) = M_\eta(t)$, то случайные величины ξ и η имеют одинаковое распределение.

Lm 1.15.5 (Об ограниченности)**.** Пусть $\xi \in B_p$, тогда

$$M_{\xi t} \leqslant \exp\left(p\left(e^t - 1\right)\right) \qquad \forall t \in \mathbb{R}$$

□

$$M_\xi(t) = (1-p)e^{0\cdot t} + pe^{1\cdot t} = 1 - p + pe^t = 1 + p\left(e^t - 1\right) \leqslant \exp\left(p\left(e^t - 1\right)\right) \qquad \left(e^x \geqslant x + 1\right)$$

■

Оценка хвостов биномиального распределения. Граница Чернова (неравенство Хевдинга)

Хотим дать оценки отклонения случайной величины с биномиальным распределением от ее среднего значения.

Замечание 1.15.6. Можно дать наивную оценку по формуле Бернулли.

$$\mathsf{P}\left(\xi \geqslant m\right) \sum_{k=m}^{n} C_k^n p^k q^{n-k}$$

Однако это неудобно и неинформативно, хочется получить аналитическую оценку.

Замечание 1.15.7. Эту оценку также можно дать по неравенству Чебышева.

$$\mathsf{P}\left(\left|\frac{v_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant \frac{pq}{n\varepsilon^2}$$

Однако это слишком грубая оценка.

Замечание 1.15.8. Если n велико, а вероятность p при этом не слишком мала, то можно дать оценку при помощи Центральной Предельной Теоремы (формула Муавра—Лапласа) и неравенства Берри—Эссеена. Однако это плохо работает при недостаточно больших n, и поправка Берри—Эссеена ведет себя примерно как $\frac{1}{\sqrt{n}}$, что также не достаточно.

Общий случай

Используем неравенство Маркова (1.10.12)

$$\mathsf{P}(\xi \geqslant a) = \mathsf{P}\left(e^{\xi t} \geqslant e^{at}\right) \leqslant \frac{\mathbb{E}\left(e^{\xi t}\right)}{e^{at}} \qquad \forall t > 0$$

Пусть $\xi = \xi_1 + \ldots + \xi_n$ — сумма независимых случайных величин. Тогда

$$\begin{aligned} \mathsf{P}(\xi \geqslant a) &\leqslant e^{-at} \cdot \Pi_{i=1}^n \mathbb{E}\left(e^{t\xi_i}\right) \qquad \forall t > 0 \\ \mathsf{P}(\xi \geqslant a) &\leqslant \min_{t>0} e^{-at} \cdot \Pi_{i=1}^n \mathbb{E}\left(e^{t\xi_i}\right) \end{aligned} \tag{BTG}$$

Аналогично $\forall t > 0$ справедливо

$$\mathsf{P}(\xi \leqslant a) = \mathsf{P}\left(e^{-t\xi} \geqslant e^{-at}\right) \leqslant \min_{t>0} e^{at} \cdot \Pi_{i=1}^n \mathbb{E}\left(e^{-t\xi_i}\right)$$

Замечание 1.15.9. Идея состоит в том, чтобы для данного типа распределения случайных величин ξ_i подобрать значение переменной t при которых, эти минимумы достигаются.

Основной случай

Пусть $S_n = \xi_1 + \ldots + \xi_n, S_n \in B_{n,p}$ и $\xi_i \in B_p$. Тогда $\mathbb{E}(\xi_i) = p, \mathbb{D}(\xi_i) = pq, M_{\xi_i}(t) = 1 - p + pe^t$.

I. Мультипликативная форма

Теорема 1.15.10. (Граница Чернова)

$$\mathsf{P}(S_n \geqslant (1+\delta)np) \leqslant \exp\Big(np(\delta - (1+\delta)\ln(1+\delta))\Big) \leqslant \exp\Big(-\frac{\delta^2}{2+\delta} \cdot np\Big) \qquad \delta > 0$$

$$\mathsf{P}(S_n \leqslant (1-\delta)np) \leqslant \exp\Big(np(-\delta - (1-\delta)\ln(1-\delta))\Big) \leqslant \exp\Big(-\frac{\delta^2}{2} \cdot np\Big) \qquad 0 < \delta < 1$$

□ Согласно BTG имеем

$$\mathsf{P}(S_n \geqslant (1+\delta)np) \leqslant e^{-t(1+\delta)np} \cdot \Pi_{i=1}^n (1 - p + pe^t) \tag{1}$$

Применим оценку 1.15.5 для производящей функции моментов.

$$\begin{aligned} \mathsf{P}(S_n \geqslant (1+\delta)np) &\leqslant e^{-t(1+\delta)np} \cdot \Pi_{i=1}^n \exp\left(p(e^t - 1)\right) \\ &= e^{-t(1+\delta)np} \cdot \exp\left(np(e^t - 1)\right) \\ &= \exp\left(np\left(e^t - 1 - t - t\delta\right)\right) \end{aligned} \tag{2}$$

Обозначим эту функцию $f(t), t \geqslant 0$ и найдем ее минимум.

$$\begin{aligned} (\ln f(t))' = np\left(e^t - 1 - \delta\right) &= 0 \\ e^t &= 1 + \delta \\ t &= \ln(1+\delta) \end{aligned} \tag{3}$$

Подставим найденное $t_{\min}$ в (2).

$$\mathsf{P}(S_n \geqslant (1+\delta)np) \leqslant \exp\left(np\left(e^{\ln(1+\delta)} - 1 - \ln(1+\delta) - \ln(1+\delta)\delta\right)\right) = \exp\left(np\left(\delta - (1+\delta)\ln(1+\delta)\right)\right)$$

Более простую оценку получаем, используя известное из курса матанализа неравенство $\ln(1+\delta) \geqslant \dfrac{2\delta}{2+\delta}$.
Второе неравенство доказываем аналогично. ■

II. Аддитивная форма

Теорема 1.15.11. (Чернова—Хевдинга)

$$\mathsf{P}\left(\frac{S_n}{n} \geqslant p + \varepsilon\right) \leqslant \left(\left(\frac{p}{p+\varepsilon}\right)^{p+\varepsilon} \cdot \left(\frac{1-p}{1-p-\varepsilon}\right)^{1-p-\varepsilon}\right)^n \qquad \forall \varepsilon > 0$$

$$\mathsf{P}\left(\frac{S_n}{n} \leqslant p - \varepsilon\right) \leqslant \left(\left(\frac{p}{p-\varepsilon}\right)^{p-\varepsilon} \cdot \left(\frac{1-p}{1-p+\varepsilon}\right)^{1-p+\varepsilon}\right)^n \qquad \forall \varepsilon > 0$$

□ Доказательство аналогично 1.15.10, но чуть более громоздкое в плане вычислений, поэтому мы его опустим. ■

Замечание 1.15.12. Эту оценку можно сделать грубее.

$$\mathsf{P}\left(\left|\frac{S_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant 2\exp\left(-\frac{n\varepsilon^2}{2pq}\right)$$

Или т.к. $pq < \frac{1}{4}$, можно получить еще более грубую оценку

$$\mathsf{P}\left(\left|\frac{S_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant 2\exp\left(-2n\varepsilon^2\right)$$

Элементы теории принятия решений

TODO: Дублирование с 07ой лекцией (конец)
Пусть требуется выбрать оптимальное решение из $A_1, \ldots, A_m$. Результат решения зависит от возможной при этом ситуации $B_1, \ldots, B_n$. Эти результаты известны. Обозначим a_{ij} — результат при принятии i-ого решения и при случившийся при этом j-той ситуации. Матрица, составленная из этих результатов, называется платежной матрицей. Если вероятности данных ситуаций B_i неизвестны, то говорят, что решение принимается в условиях полной неопределенности. Если эти вероятности известны, то говорят, что решение принимается в условиях риска.

I. Принятие решения в условиях полной неопределенности

Пример 1.15.13. На пляже есть ларек, и мы выбираем, что продавать: A_1 — полотенца, A_2 — солнечные очки, A_3 — дождевики и A_4 — надувные матрасы. Продажи зависят от четырех состояний погоды. Имеем платежную матрицу.

	B_1	B_2	B_3	B_4
A_1	1	4	5	9
A_2	3	8	5	2
A_3	4	6	6	2
A_4	2	5	5	2

Решение A_4 сразу исключаем, т.к. оно доминируется решением A_3 (во всех ситуациях решение A_3 не хуже). Далее исследуем ситуацию по критериям

1. Критерий максимакса (критерий крайнего оптимизма)

   Обозначим $M = \max_i \max_j a_{ij}$. Выбираем решение, содержащее M. В нашем случае $M = 9 \Longrightarrow$ выбираем решение A_1.

2. Критерий максимина (критерий крайнего пессимизма)

   Обозначим $W = \max_i \min_j a_{ij}$. Выбираем решение, содержащее W. В нашем случае $W = 2 \Longrightarrow$ выбираем решение A_2 или A_3.

3. Критерий Лапласа

   Считаем вероятности всех ситуаций одинаковыми, равными $\frac{1}{n}$, и принимаем то решение, где ожидаемая прибыль больше. Обозначим $L = \max_i \overline{a_i}$. В нашем случае $L = \frac{19}{4} \Longrightarrow$ выбираем решение A_1.

II. Принятие решения в условиях риска

Пусть известны вероятности p_j ситуаций B_j. Тогда рассматриваем следующие критерии

1. Критерий Байеса (максимизация выигрыша)

   В качестве оптимального выбираем то решение, где средний ожидаемый выигрыш больше. На этот критерий лучше всего ориентироваться в первую очередь. В примере ниже выбираем решение A_1 или A_2.

	0.1	0.4	0.3	0.2	$\overline{a_i}$
A_1	-5	2	4	10	3.5
A_2	-10	0	5	15	3.5
A_3	1	-2	3	5	1.2

2. Принцип минимизации риска

   В качестве оптимального выбираем то решение, где дисперсия меньше. В примере ниже выбираем решение A_1, т.к. у него дисперсия меньше. Дисперсию решения A_3 не рассматриваем, т.к. выше показали, что у него меньше математическое ожидание.

	0.1	0.4	0.3	0.2	$\mathbb{D}(a_i)$
A_1	-5	2	4	10	16.65
A_2	-10	0	5	15	50.25
A_3	1	-2	3	5	–

3. Обобщенный принцип

   Обозначим $K_i = \overline{a_i} - \lambda\sigma_i, 0 < \lambda < 3$. Выбираем решение с наибольшим K_i, константу λ нужно выбрать самим исходя из условий конкретной задачи.

4. Выбираем такое решение, чтобы прибыль была не меньше приемлемого уровня, а риск при этом был наименьшим.

Пример 1.15.14.

	Бог есть	Бога нет	
p_i	0.01	0.99	
верить	∞	-10^3	$\overline{a_1} = \infty$
не верить	0	10^6	$\overline{a_2} = const$