

Теория вероятностей и математическая статистика

С. М. Ананьевский *

10 января 2017 г.

Содержание

1	Случайные события. Вероятность	2
2	Свойства вероятности	3
3	Классическое определение вероятности. Схема равновозможных исходов . .	4
4	Условные вероятности	5
5	Формула полной вероятности. Формула Байеса	5
6	Независимые события. Пример Бернулли	6
7	Независимые испытания Бернулли. Формулы Бернулли	8
8	Предельные теоремы в схеме испытаний Бернулли	9
9	Случайная величина. Распределение случайных величин	12
10	Различные типы распределений случайных величин. Случайные величины с дискретным распределением	13
11	Случайные векторы. Распределение случайных векторов	15
12	Независимость случайных величин	15
13	Свёртка распределения	18
14	Числовые характеристики случайных величин. Математическое ожидание .	19
15	Дисперсия	22
16	Моменты случайных величин	24
17	Ковариация. Коэффициент корреляции	24
18	Неравенства Маркова и Чебышёва	26
19	Виды сходимости случайных величин	27
20	Законы больших чисел	28
21	Характеристические функции	29
22	Формула обращения и теорема единственности	31
23	Связь характеристических функций и моментов случайных величин	31
24	Непрерывность соответствия между характеристической функцией и функцией распределения	32
25	Центральная предельная теорема	33
26	Математическая статистика	34
27	Методы построения оценок для неизвестных параметров	39
28	Проверка статистических гипотез	40

*Конспект подготовлен студентом Яскевичем С. В.

1 Случайные события. Вероятность

Рассмотрим эксперимент с подбрасыванием монеты. Мы не можем угадать следующий исход, но легко можем перечислить все исходы: это выпадение «орла» (О) или «решки» (Р). В случае с подбрасыванием кубика это будут числа от 1 до 6; для двух кубиков это упорядоченные пары таких чисел. Во всех этих случаях мы имеем дело с конечным числом исходов. Мы теперь можем рассмотреть эксперимент с подбрасыванием монеты, который прекращается как только выпадает «орёл». Множество исходов будет иметь вид: $\{O, PO, PPO, \dots, P \dots PO, \dots\}$. Ясно, что это множество счётно. Примером несчётного множества исходов служит выбор случайной точки на отрезке $[0, 1]$.

Условимся обозначать Ω множество всех исходов, а ω — исход эксперимента, или *элементарное событие*.

Пусть на Ω определена система подмножеств \mathfrak{F} , являющаяся сигма-алгеброй. Элемент $A \in \mathfrak{F}$ назовём *случайным событием*. Будем говорить, что случайное событие произошло, если $\omega \in A$.

Замечание 1.1. Обычно в случае конечного или счётного числа исходов в качестве \mathfrak{F} выбирается множество всех подмножеств Ω .

Определение 1.2. Пусть P — функция, заданная на \mathfrak{F} и действующая в \mathbb{R} , а также выполнены условия:

1. $\forall A \in \mathfrak{F} \quad P(A) \geq 0$.
2. $P(\Omega) = 1$.
3. Для $A_1, A_2, A_3, \dots \in \mathfrak{F}$ — не более чем счётного набора событий, причём попарно непересекающихся, выполнено $P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$.

Тогда P называется *вероятностью*, а указанные условия — *аксиомами вероятности*.

Замечание 1.3. Вероятность является счётно-аддитивной мерой на \mathfrak{F} .

Пример 1.4. Вспомним подбрасывание монетки: $\Omega = \{O, P\}$, $\mathfrak{F} = \{\{O\}, \{P\}, \{O, P\}, \emptyset\}$. Мы можем сопоставить каждому элементу \mathfrak{F} число разными способами:

$$\left\{ \begin{array}{l} \{O\} \mapsto \frac{1}{2} \\ \{P\} \mapsto \frac{1}{2} \\ \{O, P\} \mapsto 1 \\ \emptyset \mapsto 0 \end{array} \right. \quad \text{или} \quad \left\{ \begin{array}{l} \{O\} \mapsto \frac{1}{3} \\ \{P\} \mapsto \frac{2}{3} \\ \{O, P\} \mapsto 1 \\ \emptyset \mapsto 0 \end{array} \right.$$

Разумеется, то какая вероятность будет правильной, зависит от монетки.

Определение 1.5. Пусть A — случайное событие. Событие $\bar{A} = \Omega \setminus A$ — называется *обратным* к A .

Определение 1.6. Если $A, B \in \mathfrak{F}$ — случайные события и $A \cap B = \emptyset$, то A, B называются *несовместными*. Это означает, что они не могут произойти одновременно.

Определение 1.7. Ω называется *достоверным событием*, а $\bar{\Omega} = \emptyset$ — *невозможным событием*.

Определение 1.8. Тройка $(\Omega, \mathfrak{F}, P)$ называется *вероятностным пространством*.

Вероятностное пространство представляет собой математическую модель, с помощью которой мы описываем эксперимент.

Определение 1.9. Пусть имеется не более чем счётный набор попарно непересекающихся событий $A_1, A_2, A_3, \dots \in \mathfrak{F}$, $A_i \cap A_j = \emptyset$ ($i \neq j$), $\bigcup_i A_i = \Omega$. Тогда этот набор называется *полной системой событий*. То есть, полная система событий образует разбиение Ω .

Определение 1.10. Если $A, B \in \mathfrak{F}$, то $A \cup B$ называется *суммой событий*, а $A \cap B = AB$ — *произведением событий*.

2 Свойства вероятности

Предложение 2.1. Любая вероятность $P: \mathfrak{F} \rightarrow \mathbb{R}$ удовлетворяет следующим свойствам:

1. $\forall A \in \mathfrak{F} \quad P(\bar{A}) = 1 - P(A)$.

2. $\forall A \in \mathfrak{F} \quad 0 \leq P(A) \leq 1$.

3. $P(\emptyset) = 0$.

4. Если $A, B \in \mathfrak{F}$, то $P(A \cup B) = P(A) + P(B) - P(AB)$.

5. Для конечного набора $A_1, A_2, \dots, A_m \in \mathfrak{F}$ верно:

$$P\left(\bigcup_{i=1}^m A_i\right) = \sum_{i=1}^m P(A_i) - \sum_{i \neq j} P(A_i \cap A_j) + \sum_{i \neq j \neq k} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{m+1} P\left(\bigcap_{i=1}^m A_i\right)$$

6. Если $A, B \in \mathfrak{F}$ и $A \subset B$, то $P(A) \leq P(B)$.

7. Если $A_1, A_2, A_3, \dots \in \mathfrak{F}$, то $P\left(\bigcup_i A_i\right) \leq \sum_i P(A_i)$.

8. Если $A_1, A_2, A_3, \dots \in \mathfrak{F}$ — не более чем счётный набор событий, причём $A_1 \subset A_2 \subset A_3 \subset \dots$, и $\bigcup_i A_i = A$, то $P(A) = \lim_{n \rightarrow \infty} P(A_n)$.

9. Если $B_1, B_2, B_3, \dots \in \mathfrak{F}$ — не более чем счётный набор событий, причём $B_1 \supset B_2 \supset B_3 \supset \dots$, и $\bigcap_i B_i = B$, то $P(B) = \lim_{n \rightarrow \infty} P(B_n)$.

Последние два свойства утверждают о непрерывности вероятности.

Доказательство.

1. Заметим, что A и \bar{A} несовместны и $A \cup \bar{A} = \Omega$. По второй аксиоме вероятности $P(\Omega) = 1 = P(A \cup \bar{A})$, откуда по третьей аксиоме $1 = P(A) + P(\bar{A})$.

2. $P(\bar{A}) = 1 - P(A) \geq 0$ — из первой аксиомы.

3. Поскольку $\bar{\Omega} = \emptyset$, то $P(\emptyset) = 1 - P(\Omega) = 1 - 1 = 0$.

4. Рассмотрим $A \cup B = (A \cap \bar{B}) \cup (\bar{A} \cap B) \cup (A \cap B)$. Легко видеть, что слагаемые попарно несовместны. Отсюда:

$$P(A \cup B) = P(A \cap \bar{B}) + P(\bar{A} \cap B) + P(A \cap B) + P(A \cap B) - P(A \cap B)$$

Здесь мы прибавили и вычли $P(A \cap B)$. Так как $A \cap \bar{B}$ и $A \cap B$ несовместны, то по третьей аксиоме будет $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

5. *Упражнение.*

6. Разложим B как $B = A \cup (B \cap \bar{A})$. Заметим, что A и $B \cap \bar{A}$ несовместны. Тогда $P(B) = P(A) + P(B \cap \bar{A}) \geq P(A)$ (так как $P(B \cap \bar{A}) \geq 0$).

7. $\bigcup_i A_i = A_1 \cup (\bar{A}_1 \cap A_2) \cup (\bar{A}_1 \cap \bar{A}_2 \cap A_3) \cup \dots$ — сумма попарно несовместных событий. Отсюда:

$$P\left(\bigcup_i A_i\right) = P(A_1) + P(\underbrace{\bar{A}_1 \cap A_2}_{\subset A_2}) + P(\underbrace{\bar{A}_1 \cap \bar{A}_2 \cap A_3}_{\subset A_3}) + \dots$$

По предыдущему свойству получаем $P\left(\bigcup_i A_i\right) \leq P(A_1) + P(A_2) + P(A_3) + \dots$

8. Рассмотрим $A = \bigcup_{i=1}^n A_i \cup \bigcup_{i>n} A_i = A_n \cup \bigcup_{i>n} A_i = A_n \cup (\bar{A}_n \cap A_{n+1}) \cup (\bar{A}_{n+1} \cap A_{n+2}) \cup \dots$

Отсюда $P(A) = P(A_n) + \sum_{k=n}^{\infty} P(\bar{A}_k \cap A_{k+1})$. Так как $P(A) \leq 1$, то можно утверждать, что ряд сходится и $\sum_{k=n}^{\infty} P(\bar{A}_k \cap A_{k+1}) \xrightarrow{n \rightarrow \infty} 0$.

9. Рассмотрим $\bar{B}_1 \subset \bar{B}_2 \subset \bar{B}_3 \subset \dots$. Заметим, что $\bar{B} = \bigcup_i \bar{B}_i$. По предыдущему свойству $P(\bar{B}) = \lim_{n \rightarrow \infty} P(\bar{B}_n)$ и $1 - P(B) = \lim_{n \rightarrow \infty} (1 - P(B_n)) = 1 - \lim_{n \rightarrow \infty} P(B_n)$.

□

3 Классическое определение вероятности. Схема равновероятных исходов

Рассмотрим вероятностное пространство некоторого эксперимента: $(\Omega, \mathfrak{F}, P)$. Пусть $\Omega = \{\omega_1, \dots, \omega_n\}$. Введём следующее предположение: $P(\{\omega_1\}) = \dots = P(\{\omega_n\})$. Такую ситуацию будем называть *схемой равновероятных исходов*. Тогда:

$$1 = P(\Omega) = P\left(\bigcup_{i=1}^n \{\omega_i\}\right) = \sum_{i=1}^n P(\{\omega_i\}) = n \cdot P(\{\omega_j\}) \quad \forall j = 1 \dots n$$

Мы воспользовались тем, что элементарные события, конечно, попарно несовместны, и получили, что $P(\{\omega_j\}) = \frac{1}{n}$.

Рассмотрим теперь событие $A \in \mathfrak{F}$, $A = \{\omega_{i_1}, \dots, \omega_{i_k}\}$, $k \leq n$. Тогда:

$$P(A) = \sum_{j=1}^k P(\{\omega_{i_j}\}) = \frac{k}{n}$$

Таким образом мы получили формулу классического определения вероятности:

$$P(A) = \frac{\text{число исходов, благоприятствующих событию } A}{\text{число всех возможных исходов}}$$

Пример 3.1. Представим, что мы одновременно подбрасываем два кубика, которые будем считать честными: $\Omega = \{(i, j) \mid 1 \leq i, j \leq 6\}$. Рассмотрим событие A :

$$A = \{\text{сумма выпавших очков больше или равна 10}\}$$

$$P(A) = \frac{6}{36} = \frac{1}{6}$$

Замечание 3.2. События {сумма очков равна 2} и {сумма очков равна 3} не равновероятны.

4 Условные вероятности

Пусть у нас есть вероятностное пространство $(\Omega, \mathfrak{F}, P)$ и два случайных события $A, B \in \mathfrak{F}$, причём будем считать, что $P(B) \neq 0$.

Определение 4.1. Условной вероятностью события A при условии события B называется число $P(A/B) = \frac{P(A \cap B)}{P(B)}$ (иногда обозначается $P_B(A)$).

Пример 4.2. Если у нас есть игральный кубик, то вероятность выпадения нечётной грани при условии, что количество очков не превосходит 3, равна $\frac{2}{3}$.

Теорема 4.3 (Свойства условной вероятности). Условная вероятность является вероятностью.

Доказательство. Проверим аксиомы вероятности:

1. $P_B(A) = \frac{P(A \cap B)}{P(B)}$. Так как числитель и знаменатель неотрицательны, то и дробь неотрицательна.
2. $P_B(\Omega) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$
3. Пусть $A_1, A_2, \dots \in \mathfrak{F}$; $A_i \cap A_j = \emptyset$ (при $i \neq j$). Тогда:

$$\begin{aligned} P_B\left(\bigcup_i A_i\right) &= \frac{P\left(\bigcup_i B \cap A_i\right)}{P(B)} = \frac{\sum_i P(B \cap A_i)}{P(B)} = \\ &= \sum_i \frac{P(B \cap A_i)}{P(B)} = \sum_i P_B(A_i) \end{aligned}$$

□

Следствие 4.4. Все свойства вероятности для условной вероятности выполнены.

5 Формула полной вероятности. Формула Байеса

Теорема 5.1 (Формула полной вероятности). Пусть A_1, A_2, \dots — полная система событий и $\forall i P(A_i) > 0$. Тогда вероятность любого случайного события $B \in \mathfrak{F}$ можно вычислить по формуле:

$$P(B) = \sum_i P(B/A_i) \cdot P(A_i)$$

Доказательство.

$$\begin{aligned} P(B) &= P(B \cap \Omega) = P\left(B \cap \left(\bigcup_i A_i\right)\right) = P\left(\bigcup_i (B \cap A_i)\right) = \sum_i P(B \cap A_i) = \sum_i \frac{P(B \cap A_i) \cdot P(A_i)}{P(A_i)} = \\ &= \sum_i P(B/A_i) \cdot P(A_i) \end{aligned}$$

□

Теорема 5.2 (Формула Байеса). Пусть A_1, A_2, \dots — полная система событий, $\forall i P(A_i) > 0$, $P(B) > 0$. Тогда $\forall k \geq 1$ $P(A_k/B) = \frac{P(B/A_k) \cdot P(A_k)}{\sum_i P(B/A_i) \cdot P(A_i)}$

Доказательство. Правая часть равна:

$$\frac{P(B/A_k) \cdot P(A_k)}{P(B)} = \frac{P(B \cap A_k) \cdot P(A_k)}{P(A_k) \cdot P(B)} = \frac{P(B \cap A_k)}{P(B)} = P(A_k/B)$$

□

Примеры 5.3. 1. Пусть с завода №1 поставлено 5 ящиков деталей, с завода №2 — 3 ящика, а с завода №3 — 2 ящика. Предположим также, что завод №1 допускает 2% брака, завод №2 — 5% брака, а завод №3 — 10%. Какова вероятность выбрать хорошую деталь? Какова вероятность того, что деталь изготовлена на заводе №1 при условии, что она хорошая?

Ящики считаем одинаковыми. Рассмотрим события C_1, C_2, C_3 , где C_i означает выбрать ящик с завода № i , и событие B , означающее выбор хорошей детали. Ясно, что C_1, C_2, C_3 — полная система событий. Чтобы вычислить вероятность события B , можно воспользоваться формулой полной вероятности:

$$P(B) = \sum_{i=1}^3 P(B/C_i) \cdot P(C_i) = 0.98 \cdot 0.5 + 0.95 \cdot 0.3 + 0.9 \cdot 0.2 = 0.955$$

Чтобы ответить на второй вопрос, мы можем воспользоваться формулой Байеса:

$$P(C_1/B) = \frac{0.98 \cdot 0.5}{0.98 \cdot 0.5 + 0.95 \cdot 0.3 + 0.9 \cdot 0.2} \approx 0.51$$

2. Представим, что у нас имеется ящик с шестью белыми и четырьмя чёрными шариками. Сначала мы потеряли один шарик из этого ящика (какой — неизвестно), а затем из оставшихся мы вытащили два шарика. Какова вероятность вытащить два белых шарика? Какова вероятность того, что был потерян чёрный шар, при условии, что мы вытащили два белых шара?

Введём два события, описывающие первый этап эксперимента: $C_б, C_ч$ — потеря белого и чёрного шаров соответственно. $C_б, C_ч$ — полная система событий. Пусть B означает «вытащить два белых шарика».

$$P(B) = P(B/C_б) \cdot P(C_б) + P(B/C_ч) \cdot P(C_ч) = \frac{C_5^2}{C_9^2} \cdot \frac{6}{10} + \frac{C_6^2}{C_9^2} \cdot \frac{4}{10}$$

$$P(C_ч/B) = \frac{C_6^2 \cdot 4}{C_5^2 \cdot 6 + C_6^2 \cdot 4}$$

6 Независимые события. Пример Бернштейна

Важно: нельзя путать понятия *независимости* событий и *несовместности*.

Пусть имеется эксперимент, описываемый с помощью вероятностного пространства $(\Omega, \mathfrak{F}, P)$, и даны случайные события $A, B \in \mathfrak{F}$.

Независимость событий можно было бы рассматривать как выполнение равенств

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

Однако здесь нарушена симметрия - логично, что если событие A независимо от B , то и обратное тоже верно — B независимо от A .

Определение 6.1. A и B независимы, если $P(A \cap B) = P(A)P(B)$.

Таким образом, если $P(B) > 0$, то независимость A и B равносильна $P(A/B) = P(A)$.

Предложение 6.2 (Свойства независимых событий). 1. A, B независимы $\iff A, \bar{B}$ независимы $\iff \bar{A}, B$ независимы $\iff \bar{A}, \bar{B}$ независимы.

2. $\forall A \in \mathfrak{F} \quad A$ и Ω независимы.

3. $\forall A \in \mathfrak{F} \quad A$ и \emptyset независимы.

Упражнение 6.3. Пусть A и B независимы, A и C независимы. Верно ли, что A и $B \cup C$ независимы? Верно ли, что A и $B \cap C$ независимы?

Можем ли мы определить понятие независимости для числа событий, большего 2? Для событий A_1, A_2, \dots, A_n мы можем выделить попарную независимость:

$$\forall i \neq j \quad A_i, A_j \text{ независимы.}$$

Или же независимость в совокупности (совместную): A_1, \dots, A_n независимы в совокупности, если:

1. $\forall i \neq j \quad A_i, A_j$ — независимы;

2. $\forall i_1 < i_2 < i_3 \quad P\left(\bigcap_{j=1}^3 A_{i_j}\right) = \prod_{j=1}^3 P(A_{i_j})$, и так далее;

3. $P\left(\bigcap_{j=1}^n A_{i_j}\right) = \prod_{j=1}^n P(A_{i_j})$.

Это равносильно:

$$\forall 2 \leq k \leq n \quad \forall i_1 < i_2 < \dots < i_k \quad P\left(\bigcap_{j=1}^k A_{i_j}\right) = \prod_{j=1}^k P(A_{i_j})$$

Пример 6.4 (Берштейна). Рассмотрим эксперимент: будем подбрасывать тетраэдр с белой, синей, красной и разноцветной (бело-сине-красной) гранями. Введём три события:

$$B = \{\text{внизу присутствует белый цвет}\}$$

$$C = \{\text{внизу присутствует синий цвет}\}$$

$$K = \{\text{внизу присутствует красный цвет}\}$$

Проверим, что эти события попарно независимы. Верно ли, что $P(B \cap C) = P(B)P(C)$? Очевидно, что да. Значит, попарная независимость есть. Проверим теперь совместную независимость:

$$P(B \cap C \cap K) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = P(B)P(C)P(K)$$

Это доказывает, что попарная независимость и совместная независимость неравносильны.

7 Независимые испытания Бернулли. Формулы Бернулли

Пусть у нас есть вероятностное пространство $(\Omega, \mathfrak{F}, P)$. Определим *испытание* A_1, \dots, A_m как набор событий, являющийся полной системой событий.

Пример 7.1. Представим, что мы одновременно подбрасываем монетку и кубик. Каким будет вероятностное пространство? $\Omega = \{O1, O2, \dots, O6, P1, \dots, P6\}$ Зададим испытания $A_1 = \{A_{11}, A_{12}\}$, где $A_{11} = \{\text{на монете О}\}$, $A_{12} = \{\text{на монете Р}\}$; $A_2 = \{A_{21}, \dots, A_{26}\}$, где $A_{2i} = \{\text{на кубике цифра } i\}$.

$$P(A_{11}) \cap P(A_{23}) = \frac{1}{12}, \quad P(A_{11}) = \frac{1}{2}, \quad P(A_{23}) = \frac{1}{2}$$

Определение 7.2. Испытания A_1, \dots, A_m будем называть *независимыми*, если для любого набора $A_{1i_1}, A_{2i_2}, \dots, A_{mi_m}$ составляющие его события являются совместно независимыми.

Испытаниями Бернулли называются набор из n независимых испытаний с двумя исходами в каждом из них, условно называемыми успехом и неудачей, и с постоянной вероятностью успеха во всех испытаниях. Будем обозначать такой набор $(Y_1, Y_2, H_3, \dots, Y_n)$, а вероятность успеха $P(Y_k) = p$.

Пусть $A = \{\text{в } n \text{ испытаниях Бернулли успех произошёл ровно } k \text{ раз}\}$. Какова вероятность A ? Заметим, что $A = (\underbrace{Y_1 \dots Y_k}_{k} \underbrace{H_{k+1} \dots H_n}_{n-k}) \cup (\underbrace{Y_1 \dots Y_{k-1}}_{k-1} \underbrace{Y_k H_{k+1} \dots H_n}_{n-k+1}) \cup \dots$

$$P(Y_1 \dots Y_k H_{k+1} \dots H_n) = P(Y_1 \cap Y_2 \cap \dots \cap Y_k \cap H_{k+1} \cap \dots \cap H_n) = p^k (1-p)^{n-k}$$

Ясно, что вероятность любой цепочки, содержащей k успехов и $n-k$ неудач, равна $p^k (1-p)^{n-k}$

Получаем, что в нашем примере $P(A) = C_n^k p^k (1-p)^{n-k} = P_n(k)$ — эта формула носит имя Бернулли.

Следствие 7.3.

1. $P_n(n) = p^n$
2. $P_n(0) = (1-p)^n$
3. $P_n(\text{хотя бы один успех}) = 1 - P_n(0) = 1 - (1-p)^n$

Пример 7.4. Пусть мы подбрасываем монету 10 раз. Какова вероятность того, что все десять раз выпал орёл? По формуле Бернулли:

$$P_{10}(10) = \left(\frac{1}{2}\right)^{10} = \frac{1}{1024}$$

А вероятность того, что орёл выпал ровно пять раз, равна

$$P_{10}(5) = C_{10}^5 \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^5 = \frac{252}{1024}$$

Возникает вопрос, каково наиболее вероятное число выпадений орла? $P_{10}(k) = C_{10}^k \left(\frac{1}{2}\right)^{10}$, и ясно, что максимум достигается при $k = 5$

Обобщим последний вопрос примера. Пусть имеется n испытаний Бернулли. Вероятность успеха в каждом испытании равна p . Чему равно наиболее вероятное число появления успехов?

Рассмотрим неравенство:

$$P_n(k) < P_n(k+1)$$

Его можно переписать:

$$C_n^k p^k (1-p)^{n-k} < C_n^{k+1} p^{k+1} (1-p)^{n-k-1}$$

$$\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} < \frac{n!}{(k+1)!(n-k-1)!} p^{k+1} (1-p)^{n-k-1}$$

Полученное равносильно:

$$(k+1)(1-p) < p(n-k)$$

$$k < (n+1)p - 1$$

Если мы теперь посмотрим на обратное неравенство $P_n(k) > P_n(k+1)$, то увидим, что оно равносильно $k > (n+1)p - 1$.

Рассмотрим два случая:

1. $(n+1)p \notin \mathbb{Z}$. Обозначим за k_n^* наиболее вероятное число успехов в n испытаниях. Тогда $k_n^* = [(n+1)p]$
2. $(n+1)p \in \mathbb{Z}$. В этом случае $k_n^{*1} = (n+1)p - 1$ и $k_n^{*2} = (n+1)p$ — наиболее вероятные числа успехов.

Пример 7.5. Пусть $p = \frac{1}{2}$, $n = 10$. Тогда $k_{10}^* = 5$. Если же $n = 11$, то $k_{11}^{*1} = 5$ и $k_{11}^{*2} = 6$, так как $C_{11}^5 = C_{11}^6$.

8 Предельные теоремы в схеме испытаний Бернулли

Представим, что мы подбрасываем монету 10000 раз. Ясно, что наиболее вероятное число выпадений орла равна 5000. Чему же равна вероятность такого исхода?

$P_{10000}(5000) = C_{10000}^{5000} \left(\frac{1}{2}\right)^{10000}$. Мы хотели бы оценить это число.

Рассмотрим функции:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \Phi(x) = \int_{-\infty}^x \varphi(t) dt$$

Ясно, что $\varphi(x)$ — чётная, а $\Phi(-x) = 1 - \Phi(x)$.

Пусть n — число испытаний, k — число успехов, p — вероятность успеха, $q = 1 - p$. Введём обозначение: $x_{n,k} = \frac{k-np}{\sqrt{npq}}$.

Теорема 8.1 (Локальная теорема Муавра-Лапласа). *Справедливо следующее соотношение:*

$$\frac{P_n(k)}{\frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x_{n,k}^2}{2}}} \xrightarrow{n \rightarrow \infty} 1 \quad \text{равномерно по всем } k: \quad |x_{n,k}| \leq C n^{\frac{1}{6}-\varepsilon} \quad \forall C > 0, \varepsilon > 0$$

Таким образом, $P_n(k) \approx \frac{1}{\sqrt{npq}} \varphi(x_{n,k})$.

Лемма 8.2 (Формула Стирлинга).

$$n! = n^n e^{-n} \sqrt{2\pi n} \cdot (1 + o(1)) \quad (n \rightarrow \infty)$$

Лемма 8.3.

$$\ln(1+x) = x - \frac{x^2}{2} + \theta x^3, \text{ где } |\theta| \leq 3 \quad \forall |x| < \frac{1}{2}$$

Доказательство теоремы. По формуле Бернулли:

$$\begin{aligned} P_n(k) &= \frac{n!}{k!(n-k)!} p^k q^{n-k} = \\ &= \frac{n^n e^{-n} \sqrt{2\pi n} (1+o(1))}{k^k e^{-k} \sqrt{2\pi k} (1+o(1)) (n-k)^{n-k} e^{-n+k} \sqrt{2\pi(n-k)} (1+o(1))} p^k q^{n-k} \end{aligned}$$

Но это верно при $n \rightarrow \infty$, $k \rightarrow \infty$, $n-k \rightarrow \infty$.

$$k = np + x_{n,k} \sqrt{npq} \xrightarrow{n \rightarrow \infty} \infty$$

$$n-k = nq - x_{n,k} \sqrt{npq} \xrightarrow{n \rightarrow \infty} \infty$$

Продолжая вычисления:

$$P_n(k) = \left(\frac{k}{np}\right)^{-k-\frac{1}{2}} \left(\frac{n-k}{nq}\right)^{-n+k-\frac{1}{2}} \frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}} (1+o(1))$$

Пусть $\sqrt{npq} P_n(k) = T_{n,k}$. Тогда:

$$\begin{aligned} T_{n,k} &= \left(\frac{k}{np}\right)^{-k-\frac{1}{2}} \left(\frac{n-k}{nq}\right)^{-n+k-\frac{1}{2}} \frac{1}{\sqrt{2\pi}} (1+o(1)) \\ \ln T_{n,k} &= \left(-k-\frac{1}{2}\right) \ln \frac{k}{np} + \left(-n+k-\frac{1}{2}\right) \ln \frac{n-k}{nq} + \ln \frac{1}{\sqrt{2\pi}} + o(1) \end{aligned}$$

С учётом

$$\begin{aligned} \frac{k}{np} &= 1 + x_{n,k} \sqrt{\frac{q}{np}} \\ \frac{n-k}{nq} &= 1 - x_{n,k} \sqrt{\frac{p}{nq}} \end{aligned}$$

Применив лемму 8.3, получим:

$$\begin{aligned} \ln T_{n,k} &= \left(-np - x_{n,k} \sqrt{npq} - \frac{1}{2}\right) \left(x_{n,k} \sqrt{\frac{q}{np}} - \frac{x_{n,k}^2}{2} \frac{q}{np} + \theta_1 \frac{x_{n,k}^3 q \sqrt{q}}{np \sqrt{np}}\right) + \\ &+ \left(-np + x_{n,k} \sqrt{npq} - \frac{1}{2}\right) \left(-x_{n,k} \sqrt{\frac{p}{nq}} - \frac{x_{n,k}^2}{2nq} p + \theta_2 x_{n,k}^3 \frac{p}{nq} \frac{\sqrt{p}}{\sqrt{nq}}\right) + \ln \frac{1}{\sqrt{2\pi}} + o(1) = \\ &= -x_{n,k} \sqrt{npq} + \frac{x_{n,k}^2}{2} q + o(1) - x_{n,k}^2 q + x_{n,k} \sqrt{npq} + \frac{x_{n,k}^2}{2} p - x_{n,k}^2 p + \ln \frac{1}{\sqrt{2\pi}} = \\ &= -\frac{x_{n,k}^2}{2} + \ln \frac{1}{\sqrt{2\pi}} + o(1) \end{aligned}$$

Тогда само $T_{n,k}$ равно:

$$T_{n,k} = e^{-\frac{x_{n,k}^2}{2}} \frac{1}{\sqrt{2\pi}} e^{o(1)}$$

Отсюда:

$$P_n(k) = \frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_{n,k}^2}{2}} e^{o(1)}$$

Теорема доказана. □

Попробуем ответить на вопрос: как найти $P_n(a < \text{число успехов} < b)$ при $a < b$ и больших n ? Применение локальной теоремы Муавра-Лапласа может давать слишком высокие погрешности, поэтому необходимо использовать иное решение.

Теорема 8.4 (Интегральная теорема Муавра-Лапласа). Пусть p — вероятность успеха, $0 < p < 1$. Тогда:

$$\sup_{a < b} \left| P_n(a < \text{число успехов} < b) - \int_{\frac{a-np}{\sqrt{npq}}}^{\frac{b-np}{\sqrt{npq}}} \varphi(t) dt \right| \xrightarrow{n \rightarrow \infty} 0$$

Мы докажем эту теорему позже — как частный случай более общей теоремы.

Следствие 8.5.

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = 1$$

Доказательство. Примем $a = -\infty$, $b = +\infty$. □

Если мы возьмём функцию $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$, то теорему можно сформулировать так:

$$P_n(a < \text{число успехов} < b) \approx \Phi\left(\frac{b-np}{\sqrt{npq}}\right) - \Phi\left(\frac{a-np}{\sqrt{npq}}\right)$$

Пример 8.6. Подсчитаем вероятность того, что при подбрасывании монетки 10000 раз «орёл» выпадет 5000 раз. Для этого применим локальную теорему Муавра-Лапласа.

$$P_{10000}(5000) \approx \frac{1}{\sqrt{10000 \cdot \frac{1}{2} \cdot \frac{1}{2}}} \varphi(0) = \frac{1}{50} \cdot 0.39894 \approx 0,00798$$

Если мы захотим подсчитать вероятность того, что «орёл» выпадет как минимум 4900 раз и как максимум 5100, то необходимо будет применить интегральную теорему:

$$P_{10000}(4900 < \text{число успехов} < 5100) \approx \Phi(2) - \Phi(-2) = 2\Phi(2) - 1 \approx 0.9544$$

Теорема 8.7 (Пуассона). Будем рассматривать схему серий испытаний Бернулли. Допустим, первая серия состоит из одного испытания такого, что $p_1 = P_1(Y)$. Вторая серия состоит из двух испытаний и $p_2 = P_2(Y)$. n -я серия испытаний состоит из n испытаний и $p_n = P_n(Y)$. (Здесь $P_i(Y)$ — вероятность успеха в одном испытании для каждой серии соответственно.) Пусть также $np_n = \text{const} = \lambda > 0$. Тогда:

$$P_n(k) \xrightarrow{n \rightarrow \infty} \frac{\lambda^k}{k!} e^{-\lambda}$$

Доказательство.

$$\begin{aligned} P_n(k) &= C_n^k p_n^k (1-p_n)^{n-k} = \frac{n!}{k!(n-k)!} p_n^k (1-p_n)^{n-k} = \\ &= \frac{1}{k!} \frac{n(n-1)\dots(n-k+1)}{n^k} n^k p_n^k (1-p_n)^n (1-p_n)^{-k} = \\ &= \frac{\lambda^k}{k!} \underbrace{\left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right)}_{\rightarrow 1} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{\rightarrow e^{-\lambda}} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-k}}_{\rightarrow 1} \end{aligned}$$

□

Теорема 8.8 (Закон больших чисел Бернулли).

$$\forall \varepsilon > 0 \quad P_n \left(\left| \frac{k_n}{n} - p \right| > \varepsilon \right) \xrightarrow{n \rightarrow \infty} 0$$

Здесь $\frac{k_n}{n}$ называется частотой успеха.

Доказательство. Рассмотрим вероятность:

$$\begin{aligned} P_n \left(\left| \frac{k_n}{n} - p \right| > \varepsilon \right) &= 1 - P_n \left(\left| \frac{k_n}{n} - p \right| \leq \varepsilon \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt - P_n(np - n\varepsilon \leq k_n \leq n\varepsilon + np) = \\ &= \frac{1}{\sqrt{2\pi}} \int_{|t| \leq \varepsilon \sqrt{\frac{n}{pq}}} e^{-\frac{t^2}{2}} dt + \frac{1}{\sqrt{2\pi}} \int_{|t| > \varepsilon \sqrt{\frac{n}{pq}}} e^{-\frac{t^2}{2}} dt - P_n(np - n\varepsilon \leq k_n \leq n\varepsilon + np) \end{aligned}$$

Разность первого и третьего слагаемого стремится к нулю, второе слагаемое тоже стремится к нулю. Теорема доказана. \square

9 Случайная величина. Распределение случайных величин

Будем рассматривать вероятностное пространство $(\Omega, \mathfrak{F}, P)$.

Определение 9.1. Функция $\xi : \Omega \rightarrow \mathbb{R}$ такая, что $\forall B \in \mathfrak{B} \quad \xi^{-1}(B) \in \mathfrak{F}$, называется *случайной величиной*. (Здесь \mathfrak{B} обозначает борелевскую сигма-алгебру.)

Примеры 9.2. 1. Пусть мы подбрасываем игральный кубик. $\Omega = \{\omega_1, \dots, \omega_6\}$. Пусть

$$\xi_1(\omega_i) = i, \quad \xi_2(\omega_i) = \begin{cases} 0, & i \neq 6 \\ 1, & i = 6 \end{cases} \quad . \quad \xi_1, \xi_2 \text{ — случайные величины.}$$

2. Пусть $\Omega = [0, 1]$, $\mathfrak{F} = \left\{ \left[0, \frac{1}{2}\right], \left(\frac{1}{2}, 1\right], [0, 1], \emptyset \right\}$, $\eta(x) = x$. η не будет случайной величиной, так как $\eta^{-1}\left(\left[0, \frac{1}{3}\right]\right) = \left[0, \frac{1}{3}\right] \notin \mathfrak{F}$.

Рассмотрим функцию $\mathfrak{P}_\xi : \mathfrak{B} \rightarrow \mathbb{R}$ такую, что $\forall B \in \mathfrak{B} \quad \mathfrak{P}_\xi(B) = P(\xi^{-1}(B))$.

Определение 9.3. \mathfrak{P}_ξ называется *распределением случайной величины* ξ .

Теорема 9.4. \mathfrak{P}_ξ является вероятностью.

Доказательство. Проверим аксиомы вероятности.

1. $\mathfrak{P}_\xi(B) = P(\xi^{-1}(B)) \geq 0$
2. $\mathfrak{P}_\xi(\mathbb{R}) = P(\xi^{-1}(\mathbb{R})) = P(\Omega) = 1$
3. Пусть $B_1, B_2, \dots \in \mathfrak{B}$, $B_i \cap B_j = \emptyset$ ($i \neq j$).

$$\begin{aligned} \mathfrak{P}_\xi \left(\bigcup_{i=1}^{\infty} B_i \right) &= P \left(\xi^{-1} \left(\bigcup_{i=1}^{\infty} B_i \right) \right) = \\ &= P \left(\bigcup_{i=1}^{\infty} \underbrace{\xi^{-1}(B_i)}_{\text{попарно несовместны}} \right) = \sum_i P(\xi^{-1}(B_i)) = \sum_i \mathfrak{P}_\xi(B_i) \end{aligned}$$

\square

Теперь мы можем перейти к использованию вероятностного пространства $(\mathbb{R}, \mathfrak{B}, \mathfrak{P}_\xi)$.

Определение 9.5. Функция $F_\xi : \mathbb{R} \rightarrow \mathbb{R}$ такая, что $\forall x \in \mathbb{R} F_\xi(x) = \mathfrak{P}_\xi((-\infty, x)) = P(\xi^{-1}((-\infty, x))) = P(\omega : \xi(\omega) < x) = P(\xi < x)$, называется *функцией распределения случайной величины ξ* .

Рассмотрим значение $F_\xi(b) = \mathfrak{P}_\xi((-\infty, b)) = \mathfrak{P}_\xi((-\infty, a) \cup [a, b)) = \mathfrak{P}_\xi((-\infty, a)) + \mathfrak{P}_\xi([a, b)) = F_\xi(a) + \mathfrak{P}_\xi([a, b))$. Получается, что $\mathfrak{P}_\xi([a, b)) = F_\xi(b) - F_\xi(a)$, то есть, существует взаимно-однозначное соответствие между \mathfrak{P}_ξ и F_ξ .

Предложение 9.6 (Свойства функции распределения). 1. $\forall x \ 0 \leq F_\xi(x) \leq 1$;

2. F_ξ неубывает;

3. F_ξ непрерывна слева во всех точках.

4. $F_\xi(x) \xrightarrow{x \rightarrow -\infty} 0, F_\xi(x) \xrightarrow{x \rightarrow \infty} 1$

Доказательство. Докажем свойство 3. Пусть $x_1 < x_2 < \dots$ и $x_n \xrightarrow{n \rightarrow \infty} x$. Мы хотим доказать, что $F_\xi(x_n) \xrightarrow{n \rightarrow \infty} F_\xi(x)$. $(-\infty, x_n) \subset (-\infty, x_{n+1})$ и $\bigcup_{n=1}^{\infty} (-\infty, x_n) = (-\infty, x)$. Отсюда, по свойству вероятности, $\underbrace{\mathfrak{P}_\xi((-\infty, x_n))}_{F_\xi(x_n)} \xrightarrow{n \rightarrow \infty} \underbrace{\mathfrak{P}_\xi((-\infty, x))}_{F_\xi(x)}$. Всё доказано.

Докажем свойство 4. $\forall n \geq 1 \ [-n, n] \subset [-(n+1), n+1]$ и $\bigcup_{n \geq 1} [-n, n] = \mathbb{R}$. Тогда:

$$\underbrace{\mathfrak{P}_\xi([-n, n])}_{F_\xi(n) - F_\xi(-n)} \xrightarrow{n \rightarrow \infty} \mathfrak{P}_\xi(\mathbb{R}) = 1$$

Отсюда получаем, что $\lim_{n \rightarrow \infty} F_\xi(n) = 1$ и $\lim_{n \rightarrow \infty} F_\xi(-n) = 0$. Свойство 4 доказано. \square

Мы доказали, что функция распределения удовлетворяет указанным свойствам, однако верно и обратное.

Предложение 9.7. Пусть $G : \mathbb{R} \rightarrow \mathbb{R}$ — функция, удовлетворяющая свойствам функции распределения. Тогда G является функцией распределения некоторой случайной величины.

Доказательство. Упражнение. \square

10 Различные типы распределений случайных величин. Случайные величины с дискретным распределением

Определение 10.1. Случайная величина ξ называется случайной величиной с дискретным распределением, если существует не более чем счётное подмножество вещественных чисел A такое, что $P(\xi \in A) = 1$. Другими словами, у этой случайной величины конечное или счётное число значений.

Занумеруем все элементы множества A : $A = \{a_1, a_2, \dots\}$, обозначим $p_i = P(\xi = a_i)$. Правило, сопоставляющее каждому значению a_i случайной величины ξ вероятность p_i , называется *законом распределения*.

Примеры 10.2. 1. $\xi = c, p = 1$ — случайная величина с вырожденным в точке c распределением. График функции распределения $F_\xi(x) = P(\xi < x)$ представляет собой ступенчатую функцию, принимающую значение 0 при $x \leq c$ и 1 при $x > c$.

2. $(\xi, p) : (0, 1 - p), (1, p)$ — случайная величина с распределением Бернулли.
3. $(\xi, p) : (a_1, p_1), (a_2, p_2), \dots, (a_n, p_n)$.
4. $\xi : 0, 1, 2, \dots, n$ и $p_k = P(\xi = k) = C_n^k p^k (1 - p)^{n-k}$ — случайная величина с биномиальным законом распределения и параметрами n и p .
5. $\xi : 0, 1, 2, \dots$ и $P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ — случайная величина с распределением Пуассона с параметром $\lambda > 0$.
6. $\xi : 0, 1, 2, \dots$ и $P(\xi = k) = p^k (1 - p)$ — случайная величина с геометрическим распределением с параметром p ($0 \leq p \leq 1$).

Определение 10.3. Случайная величина ξ называется случайной величиной с абсолютно непрерывным распределением, если существует функция $p(x)$ такая, что для любого $y \in \mathbb{R}$ $F_\xi(y) = \int_{-\infty}^y p(x) dx$. В этом случае $p(x) = p_\xi(x)$ называется *плотностью распределения случайной величины ξ* .

Предложение 10.4 (Свойства плотности распределения).

1. $p(x) \geq 0$.
2. $\int_{-\infty}^{+\infty} p(x) dx = 1$.

Примеры 10.5. 1. ξ — случайная величина с равномерным распределением на отрезке $[a, b]$. (Обозначается $\xi \in U_{[a, b]}$).

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

2. ξ — случайная величина с экспоненциальным распределением с параметром α .

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

3. (Важный пример!) ξ — случайная величина с нормальным распределением с параметрами $a \in \mathbb{R}$ и $\sigma^2 > 0$.

$$p_\xi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

В случае, когда $a = 0$, $\sigma^2 = 1$, говорят о *стандартном нормальном распределении*. В этом случае $p_\xi = \varphi$ — функция Гаусса.

4. ξ — случайная величина с распределением Коши. $p_\xi(x) = \frac{1}{\pi(1+x^2)}$.

Заметим, что $P(a < \xi < b) = P(a \leq \xi < b) = P(a \leq \xi \leq b) = F_\xi(b) - F_\xi(a)$. Почему строгость неравенства не имеет значения? В силу непрерывности функции распределения $P(a < \xi < b) = F_\xi(b) - F_\xi(a) \xrightarrow{b \rightarrow a} 0$.

11 Случайные векторы. Распределение случайных векторов

Определение 11.1. Пусть ξ_1, \dots, ξ_d — случайные величины, определённые на одном вероятностном пространстве. Тогда $\xi = (\xi_1, \dots, \xi_d)$ будет называться случайным d -мерным вектором, $\xi : \Omega \rightarrow \mathbb{R}^d$. Иначе говоря, $\forall B \in \mathfrak{B}^d \quad \xi^{-1}(B) \in \mathfrak{F}$.

Определение 11.2. Функция $F_{(\xi_1, \dots, \xi_d)} : \mathbb{R}^d \rightarrow \mathbb{R}$ такая, что $\forall x_1, \dots, x_d \quad F_{(\xi_1, \dots, \xi_d)}(x_1, \dots, x_d) = P(\xi_1 < x_1, \dots, \xi_d < x_d)$, называется функцией распределения случайного вектора.

ξ_1, \dots, ξ_d — с дискретным распределением, если существует $A \subset \mathbb{R}^d$ такое, что A не более чем счётно и $P((\xi_1, \dots, \xi_d) \in A) = 1$.

ξ_1, \dots, ξ_d — с абсолютно непрерывным распределением, если существует $p(x_1, \dots, x_d)$ такое, что $F_{(\xi_1, \dots, \xi_d)}(y_1, \dots, y_d) = \int_{-\infty}^{y_1} \dots \int_{-\infty}^{y_d} p(x_1, \dots, x_d) dx_1 \dots dx_d$.

Предложение 11.3 (Свойства функции распределения случайного вектора).

1. $0 \leq F_{(\xi_1, \dots, \xi_d)}(x_1, \dots, x_d) \leq 1$
2. Пусть $x_1 \leq y_1, \dots, x_d \leq y_d$. Тогда $F_{(\xi_1, \dots, \xi_d)}(x_1, \dots, x_d) \leq F_{(\xi_1, \dots, \xi_d)}(y_1, \dots, y_d)$
3. $F_{(\xi_1, \dots, \xi_d)}$ непрерывна слева по каждой координате.
4. $F_{(\xi_1, \dots, \xi_d)}(x_1, \dots, x_d) \xrightarrow{x_k \rightarrow -\infty} 0$.
5. $F_{(\xi_1, \dots, \xi_d)}(x_1, \dots, x_d) \xrightarrow{x_k \rightarrow +\infty} F_{(\xi_1, \dots, \xi_{k-1}, \xi_{k+1}, \dots, \xi_d)}(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_d)$.
6. $F_{(\xi_1, \dots, \xi_d)}(x_1, \dots, x_d) \xrightarrow{x_1 \rightarrow +\infty, \dots, x_d \rightarrow +\infty} 1$
7. Пусть $d = 2$. Тогда $P((\xi_1, \xi_2) \in [a_1, b_1] \times [a_2, b_2]) = F_{(\xi_1, \xi_2)}(b_1, b_2) - F_{(\xi_1, \xi_2)}(a_1, b_2) - F_{(\xi_1, \xi_2)}(b_1, a_2) + F_{(\xi_1, \xi_2)}(a_1, a_2)$. То есть, вероятность попадания точки в произвольный параллелепипед однозначно выражается через значения функции распределения в вершинах этого параллелепипеда.

12 Независимость случайных величин

Вспомним, что события $A, B \in \mathfrak{F}$ называются независимыми, если $P(A \cap B) = P(A)P(B)$. Введём теперь понятие независимости для случайных величин.

Определение 12.1. Пусть ξ, η — случайные величины, определённые на одном вероятностном пространстве. Будем говорить, что ξ и η — независимые случайные величины, если для любых вещественных $x, y \quad F_{(\xi, \eta)}(x, y) = F_\xi(x)F_\eta(y)$.

Определение 12.2. ξ и η — независимые случайные величины, если $\forall a_1 < b_1, a_2 < b_2 \quad P((\xi, \eta) \in [a_1, b_1] \times [a_2, b_2]) = P(a_1 \leq \xi < b_1) \cdot P(a_2 \leq \eta < b_2)$.

Предложение 12.3. Определения 12.1 и 12.2 эквивалентны.

Доказательство.

1. $(12.2 \Rightarrow 12.1)$

$$P((\xi, \eta) \in [a_1, b_1] \times [a_2, b_2]) = P(a_1 \leq \xi < b_1) \cdot P(a_2 \leq \eta < b_2)$$

При $a_1, a_2 \rightarrow -\infty$ левая часть равенства стремится к $F_{(\xi, \eta)}(b_1, b_2)$, а правая часть — к $F_\xi(b_1) \cdot F_\eta(b_2)$.

2. (12.1 \Rightarrow 12.2)

$$\begin{aligned} P((\xi, \eta) \in [a_1, b_1] \times [a_2, b_2]) &= F_{(\xi, \eta)}(b_1, b_2) - F_{(\xi, \eta)}(a_1, b_2) - F_{(\xi, \eta)}(b_1, a_2) + \\ &+ F_{(\xi, \eta)}(a_1, a_2) = F_{\xi}(b_1)F_{\eta}(b_2) - F_{\xi}(a_1)F_{\eta}(b_2) - F_{\xi}(b_1)F_{\eta}(a_2) + \\ &+ F_{\xi}(a_1)F_{\eta}(a_2) = (F_{\xi}(b_1) - F_{\xi}(a_1))(F_{\eta}(b_2) - F_{\eta}(a_2)) = \\ &P(a_1 \leq \xi < b_1) \cdot P(a_2 \leq \eta < b_2) \end{aligned}$$

□

Определение 12.4. Пусть для любых $B_1, \dots, B_d \in \mathfrak{B}$ $P(\xi_1 \in B_1, \dots, \xi_d \in B_d) = \prod_{i=1}^d P(\xi_i \in B_i)$. Тогда ξ_1, \dots, ξ_d называются независимыми случайными величинами.

Очевидно, что это определение также эквивалентно двум предыдущим.

Пример 12.5. Рассмотрим случайный вектор $\xi = (\xi_1, \xi_2)$ с плотностью распределения

$$p_{\xi}(x, y) = \begin{cases} 1, & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{иначе} \end{cases}$$

Здесь:

$$F_{\xi_1}(x) = P(\xi_1 < x) = \begin{cases} 0, & x \leq 0 \\ x, & 0 < x \leq 1 \\ 1, & x > 1 \end{cases}$$

$$F_{\xi_2}(y) = P(\xi_2 < y) = \begin{cases} 0, & y \leq 0 \\ y, & 0 < y \leq 1 \\ 1, & y > 1 \end{cases}$$

$$F_{(\xi_1, \xi_2)}(x, y) = \begin{cases} 0, & x \leq 0 \text{ или } y \leq 0 \\ xy, & 0 < x \leq 1 \text{ и } 0 < y \leq 1 \\ x, & 0 < x \leq 1 \text{ и } y > 1 \\ y, & 0 < y \leq 1 \text{ и } x > 1 \\ 1, & x > 1 \text{ и } y > 1 \end{cases}$$

Видим, что $F_{(\xi_1, \xi_2)}(x, y) = F_{\xi_1}(x)F_{\xi_2}(y)$, значит, ξ_1 и ξ_2 независимы.

Заметим, что для того, чтобы доказать, что события не являются независимыми, достаточно найти хотя бы одну точку, в которой определение нарушается.

Мы хотим найти более простой критерий независимости событий.

Рассмотрим для начала случай дискретного распределения, ограничимся двумерным случаем.

Предложение 12.6. $\xi = (\xi_1, \xi_2)$ — случайный вектор с дискретным распределением тогда и только тогда, когда его компоненты — с дискретным распределением.

Доказательство. Очевидно. □

Теорема 12.7. Пусть $\xi = (\xi_1, \xi_2)$ — с дискретным распределением. Тогда независимость случайных величин ξ_1 и ξ_2 равносильна тому, что для любых вещественных a и b выполняется $P(\xi_1 = a, \xi_2 = b) = P(\xi_1 = a)P(\xi_2 = b)$.

Доказательство. Если ξ_1, ξ_2 независимы, то $\forall a, b \in \mathbb{R}$ и $\forall \varepsilon > 0$ $P(a \leq \xi_1 \leq a + \varepsilon, b \leq \xi_2 < b + \varepsilon) = P(a \leq \xi_1 \leq a + \varepsilon)P(b \leq \xi_2 < b + \varepsilon)$ — по определению 12.2. Устремим ε к нулю.

Обратно:

$$\begin{aligned} F_{(\xi_1, \xi_2)}(y_1, y_2) &= \sum_{\substack{a_i < y_1 \\ b_j < y_2}} P(\xi_1 = a_i, \xi_2 = b_j) = \sum_{\substack{a_i < y_1 \\ b_j < y_2}} P(\xi_1 = a_i)P(\xi_2 = b_j) = \\ &= \sum_{b_j < y_2} \left(\sum_{a_i < y_1} P(\xi_1 = a_i)P(\xi_2 = b_j) \right) = \sum_{a_i < y_1} P(\xi_1 = a_i) \cdot \sum_{b_j < y_2} P(\xi_2 = b_j) = F_{\xi_1}(y_1)F_{\xi_2}(y_2) \end{aligned}$$

□

Теперь рассмотрим случай непрерывного распределения.

Предложение 12.8. Если $\xi = (\xi_1, \dots, \xi_d)$ — с абсолютным непрерывным распределением, то $\forall k = 1, 2, \dots, d$ ξ_k с абсолютно непрерывным распределением.

Заметим, что обратное, вообще говоря, неверно (доказательство этого факта можно найти, например, в пособии «Контрпримеры в теории вероятностей», Й. Стоянов, гл. 2 §5.6).

Таким образом, ξ — случайный вектор с абсолютно непрерывным распределением тогда и только тогда, когда для любых y_1, \dots, y_d

$$F_{\xi}(y_1, \dots, y_d) = \int_{-\infty}^{y_1} \dots \int_{-\infty}^{y_d} p_{\xi}(x_1, \dots, x_d) dx_1 \dots dx_d$$

При $y_2, \dots, y_d \rightarrow \infty$ левая часть стремится к $F_{\xi_1}(y_1)$, а в правой будет

$$\int_{-\infty}^{y_1} \left(\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} p_{\xi}(x_1, \dots, x_d) dx_2 \dots dx_d \right) dx_1$$

Выражение в скобках равно $p_{\xi_1}(x_1)$ — плотности распределения ξ_1 .

Помимо этого, $\xi = (\xi_1, \dots, \xi_d)$ — с абсолютно непрерывным распределением \iff для любых $a_1 < b_1, \dots, a_d < b_d$ выполнено

$$P(a_1 \leq \xi_1 < b_1, \dots, a_d \leq \xi_d < b_d) = \int_{a_1}^{b_1} \dots \int_{a_d}^{b_d} p_{\xi}(x_1, \dots, x_d) dx_1 \dots dx_d$$

Что также равносильно

$$\forall B \in \mathfrak{B}^d \quad P(\xi \in B) = \int_B p_{\xi}(x_1, \dots, x_d) dx_1 \dots dx_d$$

Как нам проверить, будут ли случайные величины, составляющие случайный вектор с абсолютно непрерывным распределением, независимыми?

Теорема 12.9. Пусть $\xi = (\xi_1, \xi_2)$ — с абсолютно непрерывным распределением и плотностью $p_{\xi}(x, y)$. Тогда независимость ξ_1, ξ_2 равносильна тому, что $p_{\xi}(x, y) = p_{\xi_1}(x)p_{\xi_2}(y)$.

Доказательство. Проверим необходимость. Рассмотрим функцию распределения в произвольной точке: $F_{(\xi_1, \xi_2)}(x, y) = F_{\xi_1}(x)F_{\xi_2}(y)$. Дважды продифференцировав левую часть, получим

$$\frac{d^2 F_{(\xi_1, \xi_2)}(x, y)}{dx dy} = p_{\xi}(x, y)$$

$$\text{Но } \frac{d^2 F_{\xi_1}(x)F_{\xi_2}(y)}{dx dy} = \frac{dF_{\xi_1}(x)}{dx} \frac{dF_{\xi_2}(y)}{dy} = p_{\xi_1}(x)p_{\xi_2}(y).$$

Обратно. Для любых x, y

$$\int_{-\infty}^x \int_{-\infty}^y p_{\xi}(s, t) ds dt = \int_{-\infty}^x p_{\xi_1}(s) ds \int_{-\infty}^y p_{\xi_2}(t) dt = F_{\xi_1}(x)F_{\xi_2}(y)$$

□

Теорема 12.10. Пусть ξ, η — независимые случайные величины, f, g — измеримые функции. Тогда $f(\xi)$ и $g(\eta)$ — тоже независимые случайные величины.

Доказательство. Рассмотрим $B_1, B_2 \in \mathfrak{B}$.

$$P(f(\xi) \in B_1, g(\eta) \in B_2) = P(\xi \in f^{-1}(B_1), \eta \in g^{-1}(B_2))$$

Но так как функции измеримы, то прообразы измеримых подмножеств будут также измеримыми. По определению 12.4 это равно:

$$P(\xi \in f^{-1}(B_1))P(\eta \in g^{-1}(B_2)) = P(f(\xi) \in B_1)P(g(\eta) \in B_2)$$

□

13 Свёртка распределения

Пусть имеется две независимые случайные величины ξ, η . Мы хотим узнать распределение $\xi + \eta$, зная функции распределения F_{ξ} и F_{η} .

Рассмотрим функцию распределения $\xi + \eta$ в произвольной точке x :

$$\begin{aligned} F_{\xi+\eta}(x) &= P(\xi + \eta < x) = P((\xi, \eta) \in A_x) = \iint_{A_x} dF_{(\xi, \eta)}(s, t) = \iint_{A_x} d(F_{\xi}(s)F_{\eta}(t)) = \\ &= \iint_{A_x} dF_{\xi}(s) dF_{\eta}(t) = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{x-s} dF_{\eta}(t) \right) dF_{\xi}(s) \end{aligned}$$

Здесь $A_x = \{(s, t) \mid s + t < x\}$. Таким образом, доказана следующая теорема.

Теорема 13.1 (О свёртке распределения).

$$F_{\xi+\eta}(x) = \int_{-\infty}^{+\infty} F_{\eta}(x-s) dF_{\xi}(s) = \int_{-\infty}^{+\infty} F_{\xi}(x-t) dF_{\eta}(t)$$

Следствие 13.2. Если ξ, η — независимые случайные величины и ξ — с абсолютно непрерывным распределением, то $\xi + \eta$ — тоже с абсолютно непрерывным распределением

и, кроме того, $p_{\xi+\eta}(x) = \int_{-\infty}^{+\infty} p_{\xi}(x-t) dF_{\eta}(t)$. Если же обе случайные величины ξ, η с

абсолютно непрерывным распределением, то $p_{\xi+\eta}(x) = \int_{-\infty}^{+\infty} p_{\xi}(x-t)p_{\eta}(t) dt$.

14 Числовые характеристики случайных величин. Математическое ожидание

Как обычно, будем рассматривать вероятностное пространство $(\Omega, \mathfrak{F}, P)$, случайная величина $\xi : \Omega \rightarrow \mathbb{R}$ с функцией распределения $F_\xi(x) = P(\xi \leq x)$ и распределением $\mathfrak{P}_\xi(A) = P(\xi^{-1}(A)) = P(\xi \in A)$.

Определение 14.1. Математическим ожиданием случайной величины ξ называется число $E\xi = \int_{\Omega} \xi(\omega) dP(\omega)$, если таковое существует.

Интеграл можно записать следующим образом: $\int_{\Omega} \xi^+(\omega) dP(\omega) + \int_{\Omega} \xi^-(\omega) dP(\omega)$. Здесь:

$$\xi^+ = \begin{cases} \xi, & \xi \geq 0 \\ 0, & \xi < 0 \end{cases} \quad \xi^- = \begin{cases} 0, & \xi \geq 0 \\ \xi, & \xi < 0 \end{cases}$$

Тогда $|\xi| = \xi^+ + \xi^-$. Ясно, что $E\xi$ и $E|\xi|$ существуют одновременно.

Замечание 14.2. В литературе математическое ожидание также может обозначаться как $M\xi$.

Для того, чтобы понять смысл этой величины, рассмотрим следующий пример.

Пример 14.3. Пусть ξ — случайная величина с дискретным распределением и законом распределения $(\xi, p) : (a_1, p_1), (a_2, p_2), \dots, (a_n, p_n)$. Для такой величины:

$$E\xi = \int_{\Omega} \xi(\omega) dP$$

Рассмотрим события $A_k = (\xi = a_k)$, причём $A_k \cap A_l = \emptyset$ при $k \neq l$ и $\bigcup_k A_k = \Omega$. Тогда:

$$E\xi = \sum_k \int_{A_k} \xi(\omega) dP = \sum_k a_k \underbrace{\int_{A_k} dP}_{P(\xi=a_k)} = \sum_k a_k p_k$$

То есть, математическое ожидание является средневзвешенным значением случайной величины. В частности, если $p_1 = p_2 = \dots = p_n = \frac{1}{n}$, то $E\xi = \frac{a_1 + \dots + a_n}{n}$.

Предложение 14.4 (Свойства математического ожидания). Пусть ξ, η — случайные величины. Верно:

1. $P(\xi = c) = 1 \implies E\xi = c$
2. $E(a\xi + b\eta) = aE\xi + bE\eta$, где $a, b \in \mathbb{R}$
3. $P(\xi \geq 0) = 1 \implies E\xi \geq 0$
4. $P(\xi \geq \eta) = 1 \implies E\xi \geq E\eta$
5. $|E\xi| \leq E|\xi|$

Доказательство.

1. $E\xi = \int_{\Omega} \xi(\omega) dP = c \int_{\Omega} dP = c \cdot P(\Omega) = c$
2. Очевидно из линейности интеграла.

3. Очевидно.

$$4. P(\xi \geq \eta) = P(\xi - \eta \geq 0) = 1 \implies E(\xi - \eta) = E\xi - E\eta \geq 0$$

5. Так как $|\xi| \geq \xi$ и $|\xi| \geq -\xi$, то $E|\xi| \geq E\xi$ и $E|\xi| \geq -E\xi$.

□

Следующая теорема даёт правило вычисления математического ожидания случайных величин.

Теорема 14.5. Пусть $\xi_1, \xi_2, \dots, \xi_n$ — случайные величины, определённые на одном вероятностном пространстве. Обозначим через ξ случайный вектор $\xi = (\xi_1, \dots, \xi_n)$. Рассмотрим также измеримую функцию $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Тогда:

$$Ef(\xi) = \int_{\mathbb{R}^n} f(x_1, \dots, x_n) \mathfrak{P}_\xi(dx_1, \dots, dx_n) = \int_{\mathbb{R}^n} f(x_1, \dots, x_n) dF_\xi(x_1, \dots, x_n)$$

Следствие 14.6. Если ξ — случайная величина с функцией распределения F_ξ , $f: \mathbb{R}^n \rightarrow \mathbb{R}$ — измеримая функция, то

$$Ef(\xi) = \int_{\mathbb{R}} f(x) dF_\xi(x)$$

$$E\xi = \int_{\mathbb{R}} x dF_\xi(x)$$

Если известна плотность p_ξ , то $E\xi = \int_{\mathbb{R}} xp_\xi(x) dx$

Доказательство теоремы. Рассмотрим два случая.

1. Функция f принимает не более чем счётное число значений. Обозначим через f_1, f_2, \dots все значения функции f . Введём множества:

$$B_k = f^{-1}(f_k) \subset \mathbb{R}^n, \text{ при этом } B_k \cap B_l = \emptyset \ (k \neq l), \bigcup_k B_k = \mathbb{R}^n;$$

$$A_k = \xi^{-1}(B_k) \in \mathfrak{F}, A_k \cap A_l = \emptyset \ (k \neq l), \bigcup_k A_k = \Omega.$$

Рассмотрим математическое ожидание $Ef(\xi)$. По определению:

$$\begin{aligned} Ef(\xi) &= \int_{\Omega} f(\xi) dP = \sum_k \int_{A_k} f(\xi) dP = \sum_k \int_{A_k} f_k dP = \sum_k f_k \int_{A_k} dP = \sum_k f_k P(A_k) = \\ &= \sum_k f_k \mathfrak{P}_\xi(B_k) = \sum_k f_k \int_{B_k} \mathfrak{P}_\xi(dx_1, \dots, dx_n) = \sum_k \int_{B_k} f_k \mathfrak{P}_\xi(dx_1, \dots, dx_n) = \\ &= \sum_k \int_{B_k} f(x_1, \dots, x_n) \mathfrak{P}_\xi(dx_1, \dots, dx_n) = \int_{\mathbb{R}^n} f(x_1, \dots, x_n) \mathfrak{P}_\xi(dx_1, \dots, dx_n) \end{aligned}$$

2. Функция f — произвольная. Возьмём $m \geq 1$. Определим f^m следующим образом:

$$f^m(x_1, \dots, x_n) = \frac{k}{m}, \text{ если } \frac{k}{m} \leq f(x_1, \dots, x_n) \leq \frac{k+1}{m}$$

То есть, мы будем приближать функцию f ступенчатой функцией:

$$|f(x_1, \dots, x_n) - f^m(x_1, \dots, x_n)| < \frac{1}{m}$$

Отсюда:

$$\begin{aligned} |Ef(\xi) - Ef^m(\xi)| &= |E(f(\xi) - f^m(\xi))| \leq E|f(\xi) - f^m(\xi)| \leq \frac{1}{m} \\ \left| \int_{\mathbb{R}^n} f(x_1, \dots, x_n) \mathfrak{P}_\xi(dx_1, \dots, dx_n) - \int_{\mathbb{R}^n} f^m(x_1, \dots, x_n) \mathfrak{P}_\xi(dx_1, \dots, dx_n) \right| &\leq \\ \leq \int_{\mathbb{R}^n} |f(x_1, \dots, x_n) - f^m(x_1, \dots, x_n)| \mathfrak{P}_\xi(dx_1, \dots, dx_n) &\leq \frac{1}{m} \int_{\mathbb{R}^n} \mathfrak{P}_\xi(dx_1, \dots, dx_n) \leq \frac{1}{m} \end{aligned}$$

В последнем неравенстве мы пользуемся тем, что интеграл на \mathbb{R}^n от единицы по вероятностной мере равен единице (теорема 9.4). Рассмотрим модуль:

$$\left| Ef(\xi) - \int_{\mathbb{R}^n} f(x_1, \dots, x_n) \mathfrak{P}_\xi(dx_1, \dots, dx_n) \right|$$

Прибавим и вычтем в выражении под знаком модуля

$$\int_{\mathbb{R}^n} f^m(x_1, \dots, x_n) \mathfrak{P}_\xi(dx_1, \dots, dx_n)$$

Получим:

$$\begin{aligned} &\left| Ef(\xi) - \int_{\mathbb{R}^n} f(x_1, \dots, x_n) \mathfrak{P}_\xi(dx_1, \dots, dx_n) \right| \leq \\ &\leq \left| Ef(\xi) - \int_{\mathbb{R}^n} f^m(x_1, \dots, x_n) \mathfrak{P}_\xi(dx_1, \dots, dx_n) \right| + \\ &+ \left| \int_{\mathbb{R}^n} f^m(x_1, \dots, x_n) \mathfrak{P}_\xi(dx_1, \dots, dx_n) - \int_{\mathbb{R}^n} f(x_1, \dots, x_n) \mathfrak{P}_\xi(dx_1, \dots, dx_n) \right| \leq \frac{1}{m} + \frac{1}{m} = \frac{2}{m} \end{aligned}$$

Поскольку это выполнено для любого сколь угодно большого m , исходная разность равна нулю и теорема доказана. □

Теорема 14.7 (Следствие для произведения независимых случайных величин). Пусть ξ, η — независимые случайные величины и, кроме того, существуют $E\xi$ и $E\eta$. Тогда $E(\xi\eta) = E\xi \cdot E\eta$.

Доказательство. Рассмотрим случайные вектор (ξ, η) .

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

$$\mathfrak{P}_{(\xi, \eta)} = \mathfrak{P}_\xi \cdot \mathfrak{P}_\eta$$

Возьмём функцию $f: \mathbb{R}^2 \rightarrow \mathbb{R}$: $f(x, y) = xy$.

$$\begin{aligned} Ef(\xi, \eta) &= E(\xi\eta) = \iint_{\mathbb{R}^2} x_1 x_2 \mathfrak{P}_{(\xi, \eta)}(dx_1, dx_2) = \iint_{\mathbb{R}^2} x_1 x_2 \mathfrak{P}_\xi(dx_1) \mathfrak{P}_\eta(dx_2) = \\ &= \int_{\mathbb{R}} x_1 \mathfrak{P}_\xi(dx_1) \cdot \int_{\mathbb{R}} x_2 \mathfrak{P}_\eta(dx_2) = E\xi \cdot E\eta \end{aligned}$$

□

15 Дисперсия

Определение 15.1. Дисперсией случайной величины ξ называется число $D\xi = E(\xi - E\xi)^2$, если оно существует.

Дисперсия представляет собой среднеквадратичное отклонение случайной величины от своего среднего значения.

Предложение 15.2 (Свойства дисперсии).

1. $D\xi \geq 0$. $D\xi = 0 \iff \exists c : P(\xi = c) = 1$
2. Если $a, b \in \mathbb{R}$, то $D(a\xi + b) = a^2 D\xi$
3. Если ξ, η — случайные величины, то $D(\xi + \eta) = D\xi + D\eta + 2E((\xi - E\xi)(\eta - E\eta))$.
4. Если ξ, η — независимые случайные величины, то $D(\xi + \eta) = D\xi + D\eta$. Обратное, вообще говоря, неверно.
5. $D\xi = E\xi^2 - (E\xi)^2$

Определение 15.3. Величина $E((\xi - E\xi)(\eta - E\eta))$ называется *ковариацией* случайных величин ξ и η . Обозначение: $\text{cov}(\xi, \eta)$.

Доказательство предложения.

1. Неотрицательность следует из свойств математического ожидания. Далее, $D\xi = \int_{\Omega} (\xi(\omega) - E\xi)^2 dP(\omega) = 0$. Это означает, что $P(\xi - E\xi = 0) = 1$, возьмём $c = E\xi$.
2. $D(a\xi + b) = E(a\xi + b - E(a\xi + b))^2 = E(a\xi + b - aE\xi - b)^2 = a^2 E(\xi - E\xi)^2 = a^2 D\xi$.
3. $D(\xi + \eta) = E(\xi + \eta - E\xi - E\eta)^2 = E(\xi - E\xi)^2 + E(\eta - E\eta)^2 + 2E((\xi - E\xi)(\eta - E\eta))$
4. По теореме 12.10 $\xi - E\xi$ и $\eta - E\eta$ будут независимыми, отсюда $E((\xi - E\xi)(\eta - E\eta)) = E(\xi - E\xi) \cdot E(\eta - E\eta) = 0$. Подставим полученный результат в предыдущее свойство.
5. $D\xi = E(\xi - E\xi)^2 = E(\xi^2 - 2\xi E\xi + (E\xi)^2) = E\xi^2 - 2(E\xi)^2 + E(E\xi)^2 = E\xi^2 - (E\xi)^2$.

□

Примеры 15.4.

1. $(\xi, p) : (c, 1)$ (вырожденное распределение). Тогда $E\xi = c$, $D\xi = 0$.
2. Если $(\xi, p) : (0, 1 - p), (1, p)$, то $E\xi = p$, $D\xi = p(1 - p)$.
3. $\xi \in B(n, p)$ — случайная величина с биномиальным законом распределения, то есть $p(\xi = k) = C_n^k p^k (1 - p)^{n-k}$. Для этой величины:

$$\begin{aligned} E\xi &= \sum_{k=0}^n k \cdot C_n^k p^k (1-p)^{n-k} = \sum_{k=1}^n k \cdot C_n^k p^k (1-p)^{n-k} = \sum_{k=1}^n \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} = \\ &= np \sum_{k=1}^n \frac{(n-1)!}{(k-1)!(n-1-(k-1))!} p^{k-1} (1-p)^{n-1-(k-1)} \end{aligned}$$

Взяв $l = k - 1$, получим $E\xi = np \sum_{l=0}^{n-1} C_{n-1}^l p^l (1-p)^{n-1-l} = np(p + (1-p))^{n-1} = np$ (здесь мы использовали бином Ньютона). $D\xi = np(1-p)$.

4. Рассмотрим $\xi \in \Pi(\lambda)$ — случайную величину с пуассоновским законом распределения:
 $P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}$.

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

Для вычисления дисперсии вычислим сначала $E\xi^2$:

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

С учётом $(E\xi)^2 = \lambda^2$ получаем $D\xi = \lambda$.

5. Для $\xi \in U_{[a,b]}$ математическое ожидание будет:

$$E\xi = \int_a^b x \frac{1}{b-a} dx = \frac{a+b}{2}$$

$$E\xi^2 = \int_a^b x^2 \frac{1}{b-a} dx = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}$$

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

6. $\xi \in N(a, \sigma^2)$ — случайная величина с нормальным законом распределения ($p_\xi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}$).

$$\begin{aligned} E\xi &= \int_{-\infty}^{+\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx \stackrel{x=\sigma y+a}{=} \int_{-\infty}^{+\infty} (\sigma y + a) \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2}} \sigma dy = \\ &= \underbrace{\sigma \int_{-\infty}^{+\infty} y \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy}_0 + a \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = a \end{aligned}$$

$$\begin{aligned} D\xi &= \int_{-\infty}^{+\infty} (x-a)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx \stackrel{x=\sigma y+a}{=} \int_{-\infty}^{+\infty} \sigma^2 y^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2}} \sigma dy = \\ &= \sigma^2 \int_{-\infty}^{+\infty} y^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \sigma^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y \cdot y e^{-\frac{y^2}{2}} dy = \\ &= \sigma^2 \frac{1}{\sqrt{2\pi}} \left[-y e^{-\frac{y^2}{2}} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy \right] = \sigma^2 \end{aligned}$$

7. ξ — случайная величина с распределением Коши ($p_\xi(x) = \frac{1}{\pi(1+x^2)}$). $\int_{-\infty}^{+\infty} \frac{x}{\pi(1+x^2)} dx$ — этот интеграл расходится, значит, математического ожидания не существует.

16 Моменты случайных величин

Определение 16.1. Пусть k — целое неотрицательное число, ξ — случайная величина. Число $E\xi^k$, если оно существует, называется k -м моментом ξ . Число $E|\xi|^k$ — абсолютный k -й момент ξ . Число $E(\xi - E\xi)^k$ — центральный k -й момент ξ . Число $E|\xi - E\xi|^k$ — абсолютный центральный k -й момент ξ .

Заметим, что все k -е моменты величины ξ существуют или отсутствуют одновременно.

Предложение 16.2 (Свойства моментов).

1. Пусть $n > m$ и существует $E\xi^n$. Тогда существует и $E\xi^m$.
2. (Неравенство Гёльдера) Если $p, q > 0$ и $\frac{1}{p} + \frac{1}{q} = 1$, то $E|\xi\eta| \leq (E|\xi|^p)^{\frac{1}{p}} \cdot (E|\eta|^q)^{\frac{1}{q}}$.
3. (Неравенство Коши-Буняковского-Шварца) $E|\xi\eta| \leq \sqrt{E\xi^2} \cdot \sqrt{E\eta^2}$.
4. (Неравенство Ляпунова) Пусть $0 < p < q$. Тогда $(E|\xi|^p)^{\frac{1}{p}} \leq (E|\xi|^q)^{\frac{1}{q}}$.

Доказательство.

1. Ясно, что $|x|^m \leq |x|^n + 1$ для любого x . То есть, $|\xi|^m \leq |\xi|^n + 1$ и $E|\xi|^m \leq E|\xi|^n + 1 < \infty$.
2. Очевидно из неравенства Гёльдера для интегралов.
3. Это свойство — частный случай предыдущего при $p = q = \frac{1}{2}$.
4. Данное неравенство равносильно $E|\xi|^p \leq (E|\xi|^q)^{\frac{p}{q}}$. С учётом $\frac{p}{q} + \frac{q-p}{q} = 1$, применим неравенство Гёльдера:

$$E(|\xi|^p \cdot 1) \leq (E|\xi|^{p \cdot \frac{q}{q-p}})^{\frac{q-p}{q}} \cdot (E(1^{\frac{q}{q-p}}))^{\frac{q-p}{q}} = (E|\xi|^q)^{\frac{p}{q}}$$

□

Определение 16.3. Пусть $m, n > 0$. Смешанным моментом величин ξ, η порядка m плюс n называется число $E\xi^m\eta^n$, если оно существует. Смешанным центральным моментом ξ, η порядка m плюс n называется число $E(\xi - E\xi)^m(\eta - E\eta)^n$.

Замечание 16.4. Таким образом, $\text{cov}(\xi, \eta)$ является смешанным центральным моментом величин порядка 1 плюс 1.

17 Ковариация. Коэффициент корреляции

Рассматриваем вероятностное пространство (Ω, \mathcal{F}, P) .

Вспомним, что если ξ, η — случайные величины, то их ковариацией $\text{cov}(\xi, \eta)$ называется число $E(\xi - E\xi)(\eta - E\eta)$.

Предложение 17.1 (Свойства ковариации).

1. $\text{cov}(\xi, \eta) = E\xi\eta - E\xi E\eta$
2. Если $\alpha, \beta \in \mathbb{R}$, то $\text{cov}(\alpha\xi + \beta, \eta) = \alpha \text{cov}(\xi, \eta)$
3. Если ξ, η — независимые, то $\text{cov}(\xi, \eta) = 0$. Обратное, вообще говоря, неверно.

Определение 17.2. Пусть существуют $D\xi > 0$ и $D\eta > 0$. Тогда число $\rho(\xi, \eta) = \frac{\text{cov}(\xi, \eta)}{\sqrt{D\xi \cdot D\eta}}$ называется коэффициентом корреляции случайных величин ξ, η .

Заметим:

$$\rho(\xi, \eta) = \frac{E(\xi - E\xi)(\eta - E\eta)}{\sqrt{D\xi} \cdot \sqrt{D\eta}} = E \frac{\xi - E\xi}{\sqrt{D\xi}} \cdot \frac{\eta - E\eta}{\sqrt{D\eta}}$$

Предложение 17.3 (Свойства коэффициента корреляции).

1. $-1 \leq \rho(\xi, \eta) \leq 1$
2. Если ξ, η — независимые, то $\rho(\xi, \eta) = 0$. Обратное, вообще говоря, неверно.
3. $|\rho(\xi, \eta)| = 1 \iff \exists a, b \in \mathbb{R} : P(\xi = a\eta + b) = 1$ и $a\rho(\xi, \eta) > 0$.

Доказательство. Ясно, что $E \frac{\xi - E\xi}{\sqrt{D\xi}} = 0$. Отсюда:

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

Докажем свойства, пользуясь этими фактами.

1.

$$\begin{aligned} |\rho(\xi, \eta)| &= \left| E \frac{\xi - E\xi}{\sqrt{D\xi}} \cdot \frac{\eta - E\eta}{\sqrt{D\eta}} \right| \leq E \left| \frac{\xi - E\xi}{\sqrt{D\xi}} \right| \cdot \left| \frac{\eta - E\eta}{\sqrt{D\eta}} \right| \leq \\ &\leq \sqrt{E \left(\frac{\xi - E\xi}{\sqrt{D\xi}} \right)^2} \sqrt{E \left(\frac{\eta - E\eta}{\sqrt{D\eta}} \right)^2} = 1 \end{aligned}$$

(Мы воспользовались неравенством Коши-Буняковского-Шварца.)

2. Очевидно.

3. Пусть $\rho(\xi, \eta) = 1$. Тогда:

$$\begin{aligned} D \left(\frac{\xi - E\xi}{\sqrt{D\xi}} - \frac{\eta - E\eta}{\sqrt{D\eta}} \right) &= E \left(\frac{\xi - E\xi}{\sqrt{D\xi}} - \frac{\eta - E\eta}{\sqrt{D\eta}} \right)^2 = \\ &= E \left(\frac{\xi - E\xi}{\sqrt{D\xi}} \right)^2 + E \left(\frac{\eta - E\eta}{\sqrt{D\eta}} \right)^2 - 2E \frac{\xi - E\xi}{\sqrt{D\xi}} \cdot \frac{\eta - E\eta}{\sqrt{D\eta}} = 2 - 2\rho(\xi, \eta) = 0 \end{aligned}$$

Это означает, что распределение случайной величины вырожденное, то есть, существует такое $c \in \mathbb{R}$, что:

$$P \left(\frac{\xi - E\xi}{\sqrt{D\xi}} - \frac{\eta - E\eta}{\sqrt{D\eta}} = c \right) = 1$$

Откуда домножением на $\sqrt{D\xi}$ и переносом нужных слагаемых в правую часть получим:

$$P \left(\xi = \frac{\sqrt{D\xi}}{\sqrt{D\eta}} \eta + b \right) = 1$$

Аналогично при $\rho(\xi, \eta) = -1$ будет:

$$D \left(\frac{\xi - E\xi}{\sqrt{D\xi}} + \frac{\eta - E\eta}{\sqrt{D\eta}} \right) = 0$$

И получаем:

$$P \left(\xi = -\frac{\sqrt{D\xi}}{\sqrt{D\eta}} \eta + b \right) = 1$$

□

Примеры 17.4.

1. ξ, η, τ — независимые случайные величины. $\xi \in U_{[-1,0]}$, $\eta \in N(0, 2)$, $\tau \in B(4, \frac{1}{2})$.

$$\begin{aligned} \rho(\xi + \tau, \eta + \tau) &= \frac{E(\xi + \tau)(\eta + \tau) - E(\xi + \tau)E(\eta + \tau)}{\sqrt{D(\xi + \tau)D(\eta + \tau)}} = \\ &= \frac{E\xi E\eta + E\xi E\tau + E\tau E\eta + E\tau^2 - E\xi E\eta - E\xi E\tau - E\tau E\eta - (E\tau)^2}{\sqrt{(D\xi + D\tau)(D\eta + D\tau)}} = \\ &= \frac{D\tau}{\sqrt{(D\xi + D\tau)(D\eta + D\tau)}} = \frac{1}{\sqrt{(\frac{1}{12} + 1)(2 + 1)}} \approx 0.555 \end{aligned}$$

2. Рассмотрим случайный вектор (ξ, η) с законом распределения:

$\eta \backslash \xi$	0	1
−1	0.4	0.1
1	0.3	0.2

Найдём законы распределения каждой величины:

ξ	0	1
p_ξ	0.7	0.3

η	−1	1
p_η	0.5	0.5

$$E\xi\eta = -0.1 + 0.2 = 0.1$$

$$\rho(\xi, \eta) = \frac{E\xi\eta - E\xi E\eta}{\sqrt{D\xi D\eta}} = \frac{0.1 - 0.3 \cdot 0}{\sqrt{0.21 \cdot 1}} = \frac{0.1}{\sqrt{0.21}} \approx 0.218$$

18 Неравенства Маркова и Чебышёва

Пусть ξ — неотрицательная случайная величина.

Лемма 18.1 (Неравенство Маркова). Для любого $t > 0$ выполнено $P(\xi \geq t) \leq \frac{E\xi}{t}$.

Доказательство. Рассмотрим $E\xi$ и распишем его по определению:

$$\begin{aligned} E\xi &= \int_{\Omega} \xi(\omega) dP(\omega) = \int_{\omega: \xi(\omega) \geq t} \xi(\omega) dP(\omega) + \underbrace{\int_{\omega: \xi(\omega) < t} \xi(\omega) dP(\omega)}_{\geq 0} \geq \\ &\geq \int_{\xi(\omega) \geq t} \xi(\omega) dP(\omega) \geq \int_{\xi(\omega) \geq t} t dP(\omega) = t \int_{\xi(\omega) \geq t} dP(\omega) = tP(\xi \geq t) \end{aligned}$$

□

Следствие 18.2 (Неравенство Чебышёва). Пусть существует $D\xi$. Тогда для любых $t > 0$ выполнено $P(|\xi - E\xi| \geq t) \leq \frac{D\xi}{t^2}$.

Доказательство.

$$P(|\xi - E\xi| \geq t) = P((\xi - E\xi)^2 \geq t^2) \leq \frac{E(\xi - E\xi)^2}{t^2} = \frac{D\xi}{t^2}$$

□

19 Виды сходимости случайных величин

Пусть ξ — случайная величина, а ξ_1, ξ_2, \dots — последовательность случайных величин на вероятностном пространстве $(\Omega, \mathfrak{F}, P)$.

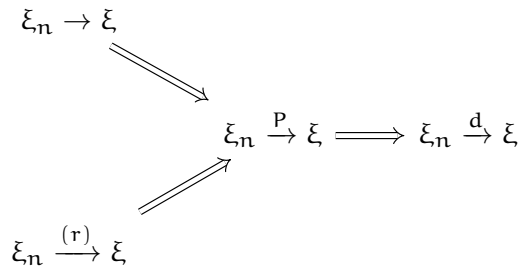
Определение 19.1 (Сходимость по вероятности). Последовательность $\{\xi_i\}_{i=1}^{\infty}$ называется *сходящейся по вероятности* (обозначается $\xi_n \xrightarrow{P} \xi$), если $\forall \varepsilon > 0 \ P(|\xi_n - \xi| \geq \varepsilon) \xrightarrow{n \rightarrow \infty} 0$.

Определение 19.2 (Сходимость почти наверное). Последовательность $\{\xi_i\}_{i=1}^{\infty}$ называется *сходящейся почти наверное* (обозначается $\xi_n \rightarrow \xi$), если $P\left(\lim_{n \rightarrow \infty} \xi_n = \xi\right) = 1$.

Определение 19.3 (Сходимость в среднем). Последовательность $\{\xi_i\}_{i=1}^{\infty}$ называется *сходящейся в среднем порядка $r > 0$* (обозначается $\xi_n \xrightarrow{(r)} \xi$), если $E|\xi_n - \xi|^r \xrightarrow{n \rightarrow \infty} 0$.

Определение 19.4 (Сходимость по распределению). Последовательность $\{\xi_i\}_{i=1}^{\infty}$ называется *сходящейся по распределению, или слабо сходящейся* (обозначается $\xi_n \xrightarrow{d} \xi$), если для всех x — точек непрерывности F_{ξ} выполнено $F_{\xi_n}(x) \rightarrow F_{\xi}(x)$.

Теорема 19.5. Из сходимости почти наверное или в среднем следует сходимость по вероятности; из сходимости по вероятности следует сходимость по распределению.



Доказательство.

1. $(\xi_n \xrightarrow{(r)} \xi \implies \xi_n \xrightarrow{P} \xi)$ Рассмотрим произвольное $\varepsilon > 0$.

$$P(|\xi_n - \xi| \geq \varepsilon) = P(|\xi_n - \xi|^r \geq \varepsilon^r) \leq \frac{E|\xi_n - \xi|^r}{\varepsilon^r} \xrightarrow{n \rightarrow \infty} 0$$

2. $(\xi_n \xrightarrow{P} \xi \implies \xi_n \xrightarrow{d} \xi)$ Пусть τ, η — случайные величины. Заметим, что верно включение:

$$(\eta < x) \subset (|\eta - \tau| \geq \varepsilon) \cup (\tau < x + \varepsilon)$$

Отсюда:

$$F_{\eta}(x) \leq P(|\eta - \tau| \geq \varepsilon) + F_{\tau}(x + \varepsilon)$$

Пусть теперь x — точка непрерывности F_{ξ} .

$$F_{\xi_n}(x) \leq P(|\xi_n - \xi| \geq \varepsilon) + F_{\xi}(x + \varepsilon) \implies \limsup_{n \rightarrow \infty} F_{\xi_n}(x) \leq F_{\xi}(x + \varepsilon)$$

$$F_{\xi}(x - \varepsilon) \leq P(|\xi_n - \xi| \geq \varepsilon) + F_{\xi_n}(x) \implies F_{\xi}(x - \varepsilon) \leq \liminf_{n \rightarrow \infty} F_{\xi_n}(x)$$

То есть:

$$F_{\xi}(x - \varepsilon) \leq \liminf_{n \rightarrow \infty} F_{\xi_n}(x) \leq \limsup_{n \rightarrow \infty} F_{\xi_n}(x) \leq F_{\xi}(x + \varepsilon)$$

Устремив ε к нулю, получим:

$$F_{\xi}(x) \leq \liminf_{n \rightarrow \infty} F_{\xi_n}(x) \leq \limsup_{n \rightarrow \infty} F_{\xi_n}(x) \leq F_{\xi}(x)$$

Значит, существует $\lim_{n \rightarrow \infty} F_{\xi_n}(x) = F_{\xi}(x)$.

3. $(\xi_n \rightarrow \xi \implies \xi_n \xrightarrow{P} \xi)$ Сходимость почти наверное равносильна тому, что

$$P\left(\underbrace{\bigcap_{r=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \left(|\xi_m - \xi| < \frac{1}{r}\right)}_{\xi_n \rightarrow \xi}\right) = 1$$

Или, что то же самое:

$$P\left(\bigcup_{r=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \left(|\xi_m - \xi| \geq \frac{1}{r}\right)\right) = 0$$

$$\forall r \geq 1 \quad P\left(\underbrace{\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \left(|\xi_m - \xi| \geq \frac{1}{r}\right)}_{B_n}\right) = 0$$

Заметим, что $B_n \supset B_{n+1}$. Поэтому:

$$\forall r \geq 1 \quad \lim_{n \rightarrow \infty} P\left(\underbrace{\bigcup_{m=n}^{\infty} \left(|\xi_m - \xi| \geq \frac{1}{r}\right)}_{\sup_{m \geq n} |\xi_m - \xi| \geq \frac{1}{r}}\right) = 0$$

Если супремум последовательности превосходит некоторое число, то и любой её элемент превосходит это число, именно это означает следующее включение:

$$\left(|\xi_n - \xi| \geq \frac{1}{r}\right) \subset \left(\sup_{m \geq n} |\xi_m - \xi| \geq \frac{1}{r}\right)$$

$$\forall r \geq 1 \quad P\left(|\xi_n - \xi| \geq \frac{1}{r}\right) \xrightarrow{n \rightarrow \infty} 0$$

□

20 Законы больших чисел

Определение 20.1 («Слабый» закон больших чисел). Пусть ξ_1, ξ_2, \dots — последовательность случайных величин, у каждой из которых существует математическое ожидание $E\xi_k = a_k$. Говорят, что для этой последовательности выполняется закон больших чисел, если:

$$\forall \varepsilon > 0 \quad P\left(\left|\frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n a_k\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow \infty} 0$$

Или, что то же самое:

$$\forall \varepsilon > 0 \quad P\left(\left|\frac{1}{n} \sum_{k=1}^n (\xi_k - E\xi_k)\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow \infty} 0$$

Или:

$$\frac{1}{n} \sum_{k=1}^n (\xi_k - E\xi_k) \xrightarrow{P} 0$$

Теорема 20.2 (Закон больших чисел Маркова). Пусть ξ_1, ξ_2, \dots — последовательность случайных величин и пусть для любого $n \geq 1$ существуют $D \sum_{k=1}^n \xi_k, \frac{1}{n^2} D \sum_{k=1}^n \xi_k \xrightarrow{n \rightarrow \infty} 0$. Тогда выполняется закон больших чисел.

Доказательство. Возьмём произвольное $\varepsilon > 0$. Тогда из неравенства Чебышёва (следствие 18.2):

$$P\left(\left|\frac{1}{n} \sum_{k=1}^n \xi_k - \underbrace{\frac{1}{n} \sum_{k=1}^n a_k}_{E \frac{1}{n} \sum_{k=1}^n \xi_k}\right| \geq \varepsilon\right) \leq \frac{D \frac{1}{n} \sum_{k=1}^n \xi_k}{\varepsilon^2} = \frac{D \sum_{k=1}^n \xi_k}{n^2 \varepsilon^2} \xrightarrow{n \rightarrow \infty} 0$$

□

Теорема 20.3 (Закон больших чисел Чебышёва). Пусть ξ_1, ξ_2, \dots — последовательность независимых случайных величин и существует $c > 0$ такое, что $\forall n \ D\xi_n \leq c$. Тогда выполняется закон больших чисел.

Доказательство.

$$P\left(\left|\frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n a_k\right| \geq \varepsilon\right) \leq \frac{D \sum_{k=1}^n \xi_k}{n^2 \varepsilon^2} = \frac{\sum_{k=1}^n D\xi_k}{n^2 \varepsilon^2} \leq \frac{nc}{n^2 \varepsilon^2} \xrightarrow{n \rightarrow \infty} 0$$

□

Замечание 20.4. Теорема 8.8 (Бернулли) является частным случаем теоремы Чебышёва. Пусть η_1, η_2, \dots — последовательность независимых одинаково распределённых случайных величин с законом распределения:

η_k	0	1
p_{η_k}	$1-p$	p

$$E\eta_k = p, \quad D\eta_k = p(1-p) \leq 1$$

$$P\left(\left|\frac{1}{n} \sum_{i=1}^n \eta_i - p\right| \geq \varepsilon\right) \rightarrow 0$$

Теорема 20.5 (Закон больших чисел Хинчина). Пусть ξ_1, ξ_2, \dots — последовательность независимых одинаково распределённых случайных величин, у которых существует математическое ожидание $E\xi_k = a$. Тогда выполняется закон больших чисел.

Мы докажем эту теорему позже.

21 Характеристические функции

Определение 21.1. Пусть ξ, η — случайные величины, определённые на вероятностном пространстве $(\Omega, \mathfrak{F}, P)$. Рассмотрим отображение $\xi + i\eta : \Omega \rightarrow \mathbb{C}$ такое, что $\forall \omega \ (\xi + i\eta)(\omega) = \xi(\omega) + i\eta(\omega)$. $\xi + i\eta$ называется *комплекснозначной случайной величиной*.

Определение 21.2. Математическим ожиданием комплекснозначной величины называется число $E\xi + iE\eta$.

Определение 21.3. Случайные величины $\xi_1 + i\eta_1, \xi_2 + i\eta_2$ называются независимыми, если $\forall C_1, C_2 \subset \mathbb{C} \ (\xi_1 + i\eta_1)^{-1}(C_1)$ и $(\xi_2 + i\eta_2)^{-1}(C_2)$ — независимые события, или, что то же самое, (ξ_1, η_1) и (ξ_2, η_2) — независимые случайные векторы.

Определение 21.4. Рассмотрим функцию $f_\xi : \mathbb{R} \rightarrow \mathbb{C}$ такую, что $\forall t \ f_\xi(t) = Ee^{it\xi}$. Такая функция называется *характеристической функцией* случайной величины ξ .

Рассмотрим случай дискретного распределения с законом:

ξ	a_1	a_2	\dots
p_ξ	p_1	p_2	\dots

Тогда $f_\xi(t) = \sum_k e^{it a_k} p_k$.

В случае абсолютно непрерывного распределения $f_\xi(t) = \int_{-\infty}^{+\infty} e^{itx} p_\xi(x) dx$.

Предложение 21.5 (Свойства характеристической функции).

1. $|f_\xi(t)| \leq 1$.
2. $f_\xi(0) = 1$.
3. $f_{\alpha\xi+\beta}(t) = e^{it\beta} \cdot f_\xi(\alpha \cdot t)$ ($\alpha, \beta \in \mathbb{R}$).
4. $f_\xi(-t) = \overline{f_\xi(t)}$.
5. $f_\xi(t)$ равномерно непрерывна на \mathbb{R} .
6. Если ξ, η — независимые, то $f_{\xi+\eta}(t) = f_\xi(t)f_\eta(t)$.

Доказательство. Свойство 1:

$$|Ee^{it\eta}| \leq E|e^{it\xi}| = E1 = 1$$

Свойство 5:

$$|f_\xi(t+h) - f_\xi(t)| = |Ee^{i(t+h)\xi} - Ee^{it\xi}| = |Ee^{it\xi}(e^{ih\xi} - 1)| \leq E|e^{ih\xi} - 1| \xrightarrow[h \rightarrow 0]{\leq 2} 0$$

Свойство 6:

$$f_{\xi+\eta}(t) = Ee^{it(\xi+\eta)} = Ee^{it\xi}e^{it\eta} = Ee^{it\xi}Ee^{it\eta} = f_\xi(t)f_\eta(t)$$

Здесь мы воспользовались теоремой 14.7. □

Примеры 21.6.

1. $P(\xi = c) = 1$. $f_\xi(t) = e^{itc}$.
2. ξ имеет распределение Бернулли:

ξ	0	1
p_ξ	$1-p$	p

$$f_\xi(t) = pe^{it} + 1 - p.$$

3. $\xi \in B(n, p)$. $f_\xi(t) = \sum_{k=0}^n e^{itk} C_n^k p^k (1-p)^{n-k} = (pe^{it} + 1 - p)^n$.
4. $\xi \in \Pi(\lambda)$. $f_\xi(t) = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}$.
5. Случайная величина ξ имеет закон распределения:

ξ	-1	1
p_ξ	$\frac{1}{2}$	$\frac{1}{2}$

$$f_\xi(t) = \cos t.$$

6. $\xi \in U_{[a,b]}$. $f_\xi(t) = \int_a^b e^{itx} \frac{1}{b-a} dx = \frac{e^{itb} - e^{ita}}{it(b-a)}$. Если $\xi \in U_{[-a,a]}$, то $f_\xi(t) = \frac{\sin at}{at}$.

7. $\xi \in N(0,1)$. $p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

$$f_\xi(t) = \int_{-\infty}^{+\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$f'_\xi(t) = i \int_{-\infty}^{+\infty} x e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \underbrace{-ie^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Big|_{-\infty}^{+\infty}}_0 - t \int_{-\infty}^{+\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = -t f_\xi(t)$$

Отсюда $\ln f_\xi(t) = -\frac{t^2}{2} + C$, $f_\xi(t) = e^{-\frac{t^2}{2}} \cdot e^C$, $C = 0$. Таким образом, $f_\xi(t) = e^{-\frac{t^2}{2}}$

8. $\xi \in N(a, \sigma^2)$. Тогда $\frac{\xi-a}{\sigma} \in N(0,1)$. $f_{\frac{\xi-a}{\sigma}}(t) = e^{-\frac{t^2}{2}}$ и $f_\xi(t) = e^{ita} e^{-\frac{t^2 \sigma^2}{2}}$.

Мы выяснили, что каждой функции распределения соответствует единственная характеристическая. Покажем теперь, что верно и обратное.

22 Формула обращения и теорема единственности

Теорема 22.1 (Формула обращения). Пусть ξ — случайная величина с функцией распределения F_ξ и характеристической функцией $f_\xi(t)$. Пусть также $x \in \mathbb{R}$ — произвольное число, $l > 0$ и $x+l$, $x-l$ — точки непрерывности функции F_ξ . Тогда:

$$F_\xi(x+l) - F_\xi(x-l) = \lim_{\sigma \rightarrow 0} \frac{1}{\pi} \int_{-\infty}^{+\infty} e^{itx} f_\xi(t) \frac{\sin tl}{t} e^{-\frac{t^2 \sigma^2}{2}} dt$$

Теорема 22.2 (Единственности). Разным функциям распределения F и G соответствуют разные характеристические функции f и g .

Доказательство. Докажем теорему от противного. Пусть $F \rightarrow f$, $G \rightarrow f$, то есть двум различным функциям распределения соответствует одна и та же характеристическая функция. Рассмотрим D_F — множество точек непрерывности функции F и D_G — множество точек непрерывности функции G . Заметим, что D_F и D_G — всюду плотные, то есть $\overline{D_F} = \overline{D_G} = \mathbb{R}$. Тогда и $D_G \cap D_F$ — всюду плотное множество на \mathbb{R} . Чтобы воспользоваться формулой обращения, возьмём a, b ($a < b$) — точки непрерывности F . Тогда:

$$F(b) - F(a) = \lim_{\sigma \rightarrow 0} \frac{1}{\pi} \int_{-\infty}^{\infty} e^{-it \frac{a+b}{2}} f(t) \frac{\sin t \frac{b-a}{2}}{t} e^{-\frac{t^2 \sigma^2}{2}} dt$$

Выбрав такие $a, b \in D_G \cap D_F$, получим: $F(b) - F(a) = G(b) - G(a)$. Устремив a к $-\infty$, получим $F(b) = G(b)$, это выполняется для всех точек непрерывности. \square

23 Связь характеристических функций и моментов случайных величин

Теорема 23.1 (О связи моментов и характеристических функций). Пусть ξ — случайная величина, у которой существует момент порядка n . Тогда существует $f_\xi^{(n)}$ и $f_\xi^{(n)}(0) = i^n E\xi^n$.

Доказательство. Будем доказывать по индукции. Пусть $n \geq 1$ и существует момент порядка n . Докажем сначала, что существует первая производная. Рассмотрим функцию $\frac{f_\xi(t+h) - f_\xi(t)}{h}$ и покажем, что у неё существует предел при $h \rightarrow 0$.

$$\begin{aligned}\frac{f_\xi(t+h) - f_\xi(t)}{h} &= E \frac{1}{h} (e^{i(t+h)\xi} - e^{it\xi}) = E \frac{1}{h} (e^{ih\xi} - 1) e^{it\xi} \\ \frac{1}{h} (e^{ih\xi} - 1) &\xrightarrow{h \rightarrow 0} i\xi\end{aligned}$$

Отсюда по теореме Лебега:

$$E \frac{1}{h} (e^{ih\xi} - 1) e^{it\xi} \xrightarrow{h \rightarrow 0} iE\xi e^{it\xi}$$

То есть производная существует и равна:

$$f'_\xi(t) = iE\xi e^{it\xi}$$

Пусть теперь для любого $k \leq n-1$ существует $f_\xi^{(k)}(t) = i^k E\xi^k e^{it\xi}$.

$$\begin{aligned}\frac{f_\xi^{(n-1)}(t+h) - f_\xi^{(n-1)}(t)}{h} &= i^{n-1} E \frac{1}{h} \xi^{n-1} (e^{i(t+h)\xi} - e^{it\xi}) = \\ &= i^{n-1} E \frac{e^{ih\xi} - 1}{h} \xi^{n-1} e^{it\xi} \xrightarrow{h \rightarrow 0} i^n E\xi^n e^{it\xi}\end{aligned}$$

□

Следствие 23.2. Пусть существует момент порядка n случайной величины ξ , который мы обозначим $E\xi^n = m_n$. Тогда:

$$f_\xi(t) = \sum_{k=0}^n \frac{(it)^k}{k!} m_k + o(t^n)$$

24 Непрерывность соответствия между характеристической функцией и функцией распределения

Теорема 24.1. Пусть F, F_1, F_2, \dots — последовательность функций распределения, f, f_1, f_2, \dots — последовательность соответствующих им характеристических функций. Пусть также для любой x — точки непрерывности F выполнено $F_k(x) \xrightarrow{k \rightarrow \infty} F(x)$. Тогда для любого t $f_k(t) \xrightarrow{k \rightarrow \infty} f(t)$. Иными словами, сходимость по распределению некоторой последовательности случайных величин (причём необязательно определённых на одном вероятностном пространстве) влечёт поточечную сходимость соответствующей последовательности характеристических функций.

Теорема 24.2 (Обратная). Пусть для последовательности характеристических функций f_1, f_2, \dots существует $\lim_{n \rightarrow \infty} f_n(t) = f(t)$, непрерывный в точке 0. Тогда:

1. $f(t)$ — характеристическая функция.
2. $F_n(x) \rightarrow F(x)$ для всех x — точек непрерывности F , где F — функция распределения, соответствующая f .

Пользуясь этими утверждениями, докажем ранее сформулированные теоремы.

Доказательство теоремы 20.5. Рассмотрим $E(\xi_k - a) = 0$. Тогда:

$$f_{\xi_k - a}(t) = 1 + o(t)$$

Рассмотрим характеристическую функцию:

$$f_{\frac{1}{n} \sum_{k=1}^n (\xi_k - a)}(t) = f_{\sum_{k=1}^n (\xi_k - a)}\left(\frac{t}{n}\right) = \left(f_{\xi_1 - a}\left(\frac{t}{n}\right)\right)^n = \left(1 + o\left(\frac{t}{n}\right)\right)^n \xrightarrow{n \rightarrow \infty} 1$$

Таким образом, последовательность характеристических функций сходится к функции, являющейся характеристической для вырожденного в нуле распределения. Отсюда по теореме 24.2:

$$F_{\frac{1}{n} \sum_{k=1}^n (\xi_k - a)}(x) \rightarrow \begin{cases} 0, & x < 0 \\ 1, & x > 0 \end{cases}$$

Мы показали сходимость по распределению, а нам требуется сходимость по вероятности. Возьмём $\varepsilon > 0$.

$$P\left(\left|\frac{1}{n} \sum_{k=1}^n (\xi_k - a)\right| \geq \varepsilon\right) \leq \underbrace{F_{\frac{1}{n} \sum_{k=1}^n (\xi_k - a)}\left(-\frac{\varepsilon}{2}\right)}_{\xrightarrow{n \rightarrow \infty} 0} + 1 - \underbrace{F_{\frac{1}{n} \sum_{k=1}^n (\xi_k - a)}(\varepsilon)}_{\xrightarrow{n \rightarrow \infty} 0}$$

Здесь мы использовали $-\frac{\varepsilon}{2}$, чтобы сохранить нестрогость неравенства. □

25 Центральная предельная теорема

Определение 25.1. Пусть ξ_1, ξ_2, \dots — последовательность случайных величин, $A_n = E \sum_{i=1}^n \xi_i$, $B_n^2 = D \sum_{i=1}^n \xi_i$, причём $B_n^2 \xrightarrow{n \rightarrow \infty} \infty$. Говорят, что для этой последовательности выполняется центральная предельная теорема, если для всех $x \in \mathbb{R}$:

$$P\left(\frac{\sum_{i=1}^n \xi_i - A_n}{B_n} < x\right) \xrightarrow{n \rightarrow \infty} \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$$

Или, что то же самое:

$$\frac{\sum_{i=1}^n \xi_i - A_n}{B_n} \xrightarrow{d} N(0, 1)$$

Теорема 25.2 (Центральная предельная теорема Леви). Пусть ξ_1, ξ_2, \dots — последовательность независимых одинаково распределённых случайных величин, у которых существует математическое ожидание $E\xi_k = a$ и дисперсия $D\xi_k = \sigma^2 < \infty$. Тогда для этой последовательности выполняется центральная предельная теорема.

Доказательство. Нам нужно доказать, что $\frac{\sum_{k=1}^n \xi_k - na}{\sqrt{n\sigma^2}} \xrightarrow{d} N(0, 1)$. Мы знаем, что это равносильно сходимости характеристических функций:

$$f_{\frac{\sum_{k=1}^n \xi_k - na}{\sqrt{n\sigma^2}}}(t) \xrightarrow{n \rightarrow \infty} e^{-\frac{t^2}{2}}$$

Рассмотрим характеристическую функцию случайной величины $\frac{\xi_k - a}{\sigma}$. Заметим, что $E \frac{\xi_k - a}{\sigma} = 0$, $D \frac{\xi_k - a}{\sigma} = 1 = E \left(\frac{\xi_k - a}{\sigma} \right)^2$. Тогда характеристическую функцию можно разложить по формуле Тейлора:

$$f_{\frac{\xi_k - a}{\sigma}}(t) = 1 + \frac{(it)^2}{2} + o(t^2) = 1 - \frac{t^2}{2} + o(1)$$

Отсюда:

$$\begin{aligned} f_{\frac{\sum_{k=1}^n \xi_k - na}{\sqrt{n\sigma^2}}}(t) &= f_{\frac{1}{\sqrt{n}} \sum_{k=1}^n \frac{\xi_k - a}{\sigma}}(t) = f_{\sum_{k=1}^n \frac{\xi_k - a}{\sigma}}\left(\frac{t}{\sqrt{n}}\right) = \prod_{k=1}^n f_{\frac{\xi_k - a}{\sigma}}\left(\frac{t}{\sqrt{n}}\right) = \left(f_{\frac{\xi_k - a}{\sigma}}\left(\frac{t}{\sqrt{n}}\right)\right)^n = \\ &= \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n \xrightarrow{n \rightarrow \infty} e^{-\frac{t^2}{2}} \end{aligned}$$

□

Докажем, что интегральная теорема Муарва-Лапласа (8.4) является частным случаем теоремы Леви.

Пусть ξ_1, ξ_2, \dots — независимые одинаково распределённые случайные величины с законом распределения:

ξ_k	0	1
p_k	$1 - p$	p

$$P\left(\sum_{k=1}^n \xi_k < b\right) = P\left(\frac{\sum_{k=1}^n \xi_k - np}{\sqrt{np(1-p)}} < \frac{b - np}{\sqrt{np(1-p)}}\right) \sim \Phi\left(\frac{b - np}{\sqrt{np(1-p)}}\right)$$

Теорема 25.3 (Центральная предельная теорема Ляпунова). Пусть ξ_1, ξ_2, \dots — последовательность независимых случайных величин, у которых существуют математические ожидания $E\xi_k = a_k$, дисперсии $D\xi_k = \sigma_k^2 < \infty$ и третьи центральные абсолютные моменты $c_k^3 = E|\xi_k - a_k|^3 < \infty$. Возьмём $A_n = \sum_{k=1}^n a_k$, $B_n^2 = \sum_{k=1}^n \sigma_k^2$, $C_n^3 = \sum_{k=1}^n c_k^3$. Пусть также $L_n = \frac{C_n}{B_n^3} \xrightarrow{n \rightarrow \infty} 0$. Тогда для последовательности выполняется центральная предельная теорема.

26 Математическая статистика

Пример 26.1. Допустим, у нас имеется ящик с M занумерованными шариками. Мы хотим за конечное число экспериментов («доставаний») определить число M . Пусть X_i — результат i -го эксперимента. Мы можем рассматривать X_1, X_2, \dots, X_n как независимые одинаково распределённые случайные величины с некоторой функцией распределения $F(x)$.

В математической статистике набор X_1, X_2, \dots, X_n называется *случайной выборкой*, а функция распределения $F(x)$ — *генеральной функцией распределения*.

Другой тип задач заключается в нахождении некоторого параметра α , от которого может зависеть функция распределения $F(x, \alpha)$.

Функция $G(X_1, X_2, \dots, X_n) = g_n$, не зависящая от неизвестных параметров распределения, называется *статистикой*, или *оценкой*.

Рассмотрим функцию:

$$\tilde{F}_n(t) = \frac{\text{число тех } X_i, \text{ которые меньше } t}{n} = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{(-\infty, t)}(X_i)$$

Здесь $\mathbb{I}_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$. Такая функция называется *выборочной (эмпирической) функцией распределения*.

1. $\frac{1}{n} \sum_{i=1}^n X_i^k$ называется k -м начальным выборочным моментом.
2. $\frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$ — выборочное математическое ожидание.
3. $\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k$ — k -й центральный выборочный момент.
4. $s^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2$ — выборочная дисперсия.
5. $s_0^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ — выборочная исправленная дисперсия.

Рассмотрим $X_{(1)} = \min(X_1, \dots, X_n)$, $X_{(2)}$ — следующее по величине значение, и так далее, $X_{(n)} = \max(X_1, \dots, X_n)$. Последовательность $X_{(1)}, X_{(2)}, \dots$ называется *вариационным рядом*, а её элемент $X_{(k)}$ — k -й порядковой статистикой.

Функция распределения k -й порядковой статистики имеет вид:

$$G_k(x) = P(X_{(k)} < x) = \sum_{i=k}^n C_n^i (F(x))^i (1 - F(x))^{n-i}$$

Определим *медиану* выборки как $m_{\frac{1}{2}} = \begin{cases} X_{([\frac{n}{2}] + 1)}, & \text{если } n \text{ — нечётное} \\ \frac{1}{2}(X_{(\frac{n}{2})} + X_{(\frac{n}{2} + 1)}), & \text{если } n \text{ — чётное} \end{cases}$

Выборочная α -квантиль: $\gamma_\alpha = X_{([\alpha n] + 1)}$. Здесь $0 < \alpha < 1$ и $\alpha \neq \frac{1}{2}$.

Рассмотрим независимые одинаково распределённые двумерные векторы $(X_1, Y_1), \dots, (X_n, Y_n)$ с генеральной функцией распределения $F(x, y)$. Мы можем рассмотреть две одномерные выборки: $X_1, \dots, X_n \sim G(x)$ и $Y_1, \dots, Y_n \sim H(y)$. Тогда величина $\overline{\text{cov}}(X, Y) = \frac{1}{n} \sum_{i=1}^n X_i Y_i - \bar{X}\bar{Y} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$ называется ковариацией выборки, а $R(X, Y) = \frac{\overline{\text{cov}}(X, Y)}{\sqrt{s_X^2 s_Y^2}}$ — выборочным коэффициентом корреляции.

Пусть g_n — оценка, α — некоторый неизвестный параметр выборки.

1. Если $Eg_n = \alpha$, то g_n называется *несмещённой оценкой*. Если $Eg_n \xrightarrow{n \rightarrow \infty} \alpha$, то g_n называется *асимптотически несмещённой оценкой*.
2. Если $g_n \xrightarrow{P} \alpha$, то g_n — *состоятельная оценка*.
3. Если $\sqrt{n} \frac{g_n - \alpha}{\Delta} \xrightarrow{d} N(0, 1)$, то g_n — *асимптотически нормальная оценка* с дисперсией Δ^2 .

Теорема 26.2 (Гливенко). Выборочная (эмпирическая) функция распределения $\bar{F}(x)$ является несмещённой, состоятельной и асимптотически нормальной оценкой для генеральной функции распределения $F(x)$.

Доказательство. Вспомним, что выборочная функция распределения может быть записана как $\bar{F}(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{(-\infty, x)}(X_i)$. Заметим, что $\mathbb{I}_{(-\infty, x)}(X_1), \mathbb{I}_{(-\infty, x)}(X_2), \dots, \mathbb{I}_{(-\infty, x)}(X_n)$ — независимые одинаковы распределённые случайные величины.

$$\mathbb{I}_{(-\infty, x)}(X_i) = \begin{cases} 0, & X_i \geq x \\ 1, & X_i < x \end{cases}$$

$$P(\mathbb{I}_{(-\infty, x)}(X_i) = 1) = P(X_i < x) = F(x)$$

$$E\bar{F}(x) = E \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{(-\infty, x)}(X_i) = \frac{1}{n} \sum_{i=1}^n E\mathbb{I}_{(-\infty, x)}(X_i) = F(x)$$

Это означает несмещённость. По закону больших чисел Хинчина (теорема 20.5):

$$\frac{1}{n} \sum_{i=1}^n \mathbb{I}_{(-\infty, x)}(X_i) \xrightarrow{P} F(x)$$

Это означает состоятельность. По центральной предельной теореме Леви (25.2):

$$\frac{\sum_{i=1}^n \mathbb{I}_{(-\infty, x)}(X_i) - nF(x)}{\sqrt{(1-F(x))F(x)n}} \xrightarrow{d} N(0, 1)$$

$$\frac{\sqrt{n}(\bar{F}(x) - F(x))}{\sqrt{F(x)(1-F(x))}} \xrightarrow{d} N(0, 1)$$

Это означает асимптотическую нормальность. □

Пусть $a = EX_i$ для некоторой выборки X_1, \dots, X_n с генеральной функцией распределения $F(x)$. Мы хотим оценить a .

Теорема 26.3 (Свойства выборочного среднего).

1. Выборочное среднее \bar{X} — несмещённая оценка для математического ожидания a .
2. \bar{X} — состоятельная оценка для a .
3. Если существует второй момент EX_i^2 , то \bar{X} — асимптотически нормальная оценка для a .

Доказательство.

$$1. E\bar{X} = E \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{n} \sum_{i=1}^n EX_i = a$$

$$2. \text{ По закону больших чисел Хинчина } \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} a$$

$$3. \text{ По центральной предельной теореме Леви } \frac{\sum_{i=1}^n X_i - na}{\sqrt{n(EX_i^2 - (EX_i)^2)}} \xrightarrow{d} N(0, 1)$$

□

Попробуем оценить дисперсию $\sigma^2 = DX_i$. Вспомним, что $s^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2$

Лемма 26.4. Пусть $\Delta_n = \xi_n + \eta_n$, $\xi_n \xrightarrow{d} \xi$, $\eta_n \xrightarrow{P} 0$. Тогда $\Delta_n \xrightarrow{d} \xi$.

Доказательство.

$$|f_{\Delta_n}(t) - f_{\xi}(t)| = |Ee^{it(\xi_n + \eta_n)} - Ee^{it\xi}| \leq |Ee^{it(\xi_n + \eta_n)} - Ee^{it\xi_n}| + \underbrace{|Ee^{it\xi_n} - Ee^{it\xi}|}_{\rightarrow 0}$$

$$|Ee^{it(\xi_n + \eta_n)} - Ee^{it\xi_n}| = |Ee^{it\xi_n}(e^{it\eta_n} - 1)| \leq E|e^{it\eta_n} - 1| \leq 2$$

Из сходимости η_n по вероятности к нулю следует сходимость по распределению, что влечёт сходимость характеристических функций $e^{it\eta_n} \xrightarrow[n \rightarrow \infty]{} 1$. По теореме Лебега о предельном переходе под знаком интеграла $E|e^{it\eta_n} - 1| \xrightarrow[n \rightarrow \infty]{} 0$. □

Теорема 26.5 (Свойства выборочной дисперсии).

1. s^2 — инвариантная относительно сдвига выборки оценка.
2. s^2 — асимптотически несмещённая оценка для σ^2 (которая смещённая).
3. s^2 — состоятельная оценка для σ^2 .
4. Если существует EX_i^4 , то s^2 — асимптотически нормальная оценка с дисперсией $EX_i^4 - (EX_i^2)^2$ для σ^2 .

Доказательство.

1. Рассмотрим сдвиг $X_1 + c, \dots, X_n + c$.

$$s_c^2 = \frac{1}{n} \sum_{i=1}^n (X_i + c - (\bar{X} + c))^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = s^2$$

Поэтому далее разумно предположить, что $EX_i = 0$

- 2.

$$Es^2 = E\left(\frac{1}{n} \sum_{i=1}^n X_i^2 - E(\bar{X})^2\right) = \frac{1}{n} \sum_{i=1}^n \underbrace{EX_i^2}_{\sigma^2} - E\left(\frac{1}{n} \sum_{i=1}^n X_i\right)^2 =$$

$$= \sigma^2 - \frac{1}{n^2} \left(E \sum_{i \neq j} X_i X_j + E \sum_{i=1}^n X_i^2 \right) = \sigma^2 - \frac{1}{n} \sigma^2 = \frac{n-1}{n} \sigma^2 \xrightarrow[n \rightarrow \infty]{} \sigma^2$$

$= \sum_{i \neq j} EX_i EX_j = 0$

3. Возьмём $\varepsilon > 0$.

$$P(|s^2 - \sigma^2| \geq \varepsilon) = P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2 - \sigma^2\right| \geq \varepsilon\right) \leq P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i^2 - \sigma^2\right| \geq \frac{\varepsilon}{2}\right) + P\left(\bar{X}^2 \geq \frac{\varepsilon}{2}\right)$$

По закону больших чисел Хинчина первое слагаемое в правой части неравенства стремится к нулю при $n \rightarrow \infty$.

$$P\left(\bar{X}^2 \geq \frac{\varepsilon}{2}\right) \leq \frac{E\bar{X}^2}{\frac{\varepsilon}{2}} = \frac{2E\left(\frac{1}{n} \sum_{i=1}^n X_i\right)^2}{\varepsilon} = \frac{2n\sigma^2}{\varepsilon n^2} \xrightarrow[n \rightarrow \infty]{} 0$$

Состоятельность доказана.

4.

$$\sqrt{n}(s^2 - \sigma^2) = \sqrt{n}\left(\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2 - \sigma^2\right) = \frac{\sum_{i=1}^n X_i^2 - n\sigma^2}{\sqrt{n}} - \sqrt{n} \cdot \bar{X}^2$$

Первое слагаемое по центральной предельной теореме Леви стремится к нормальному закону, а второе стремится к нулю по вероятности:

$$P(\sqrt{n} \cdot \bar{X}^2 > \varepsilon) \leq \frac{n^{\frac{3}{2}} \sigma^2}{n^2 \varepsilon} \xrightarrow{n \rightarrow \infty} 0$$

Осталось только воспользоваться леммой.

□

Предположим, что у нас есть выборка X_1, \dots, X_n с генеральной функцией распределения $F(x, \theta)$, где θ — какой-то параметр, который мы хотим оценить.

Рассмотрим понятие *доверительного интервала*.

Определение 26.6. Пусть g_n, h_n — статистики, причём $P(g_n < h_n) = 1$. Пусть $0 < \alpha \ll 1$. Если $P(g_n < \theta < h_n) \geq 1 - \alpha$, то (g_n, h_n) называется доверительным интервалом для параметра θ с доверительной вероятностью α .

Геометрический смысл доверительного интервала заключается в том, что мы покрываем с заданной вероятностью α точку θ этим интервалом.

Пусть $\xi_0, \xi_1, \dots, \xi_n$ — независимые одинаково распределённые по стандартному нормальному закону случайные величины. Пусть $\chi_n^2 = \sum_{i=1}^n \xi_i^2$. Такую случайную величину называют *распределённой по закону хи-квадрат с n степенями свободы*.

Рассмотрим $T_n = \frac{\xi_0}{\sqrt{\frac{1}{n} \sum_{i=1}^n \xi_i^2}}$. Говорят, что эта случайная величина распределена по закону

Стьюдента с n степенями свободы. При $n = 1$ мы получим распределение Коши, а при $n \rightarrow \infty$ T_n будет стремиться к стандартному нормальному закону.

Лемма 26.7 (Фишера). Пусть $X_1, \dots, X_n \sim N(a, \sigma^2)$ — случайная выборка. Тогда:

1. $\sqrt{n} \frac{\bar{X} - a}{\sigma} \in N(0, 1)$
2. \bar{X}, s^2 — независимые случайные величины.
3. $\frac{ns^2}{\sigma^2}$ распределена по закону хи-квадрат с $n - 1$ степенями свободы.
4. $\sqrt{n-1} \frac{\bar{X} - a}{s}$ имеет распределение Стьюдента с $n - 1$ степенями свободы

Примеры 26.8.

1. Построение доверительного интервала для a , когда σ^2 известно. Воспользуемся первым утверждением вышеприведённой леммы. Возьмём число α : $0 < \alpha \ll 1$. С помощью таблиц стандартного нормального закона возьмём $z_{1-\frac{\alpha}{2}}$ такое, что функция распределения $\Phi(z_{1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$.

$$P\left(-z_{1-\frac{\alpha}{2}} < \sqrt{n} \frac{\bar{X} - a}{\sigma} < z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$P\left(\bar{X} - \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{n}} < a < \bar{X} + \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right) = 1 - \alpha$$

То есть, доверительный интервал равен $\left(\bar{X} - \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{n}}, \bar{X} + \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right)$.

2. Построение доверительного интервала для α , когда σ^2 неизвестно. С помощью таблиц распределения Стьюдента с $n - 1$ степенями свободы найдём $z_{1-\frac{\alpha}{2}}$ такое, что функция распределения $S_{n-1}(z_{1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$.

$$P\left(-z_{1-\frac{\alpha}{2}} < \sqrt{n-1} \frac{\bar{X} - \alpha}{s} < z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$P\left(\bar{X} - \frac{sz_{1-\frac{\alpha}{2}}}{\sqrt{n-1}} < \alpha < \bar{X} + \frac{sz_{1-\frac{\alpha}{2}}}{\sqrt{n-1}}\right) = 1 - \alpha$$

Доверительный интервал: $\left(\bar{X} - \frac{sz_{1-\frac{\alpha}{2}}}{\sqrt{n-1}}, \bar{X} + \frac{sz_{1-\frac{\alpha}{2}}}{\sqrt{n-1}}\right)$. Заметим, что длина этого интервала $\frac{2sz_{1-\frac{\alpha}{2}}}{\sqrt{n-1}} \xrightarrow{P} 0$ при $n \rightarrow \infty$.

27 Методы построения оценок для неизвестных параметров

Мы рассматриваем статистики X_1, \dots, X_n с генеральной функцией распределения $F(x, \theta)$, причём θ может быть вектором. Мы хотим оценить θ .

Метод моментов Обозначим через $a_i = \frac{1}{n} \sum_{k=1}^n X_k^i$ i -й выборочный момент, через $\alpha_i(\theta) = E_\theta X_1^i$. Метод моментов заключается в решении следующей системы:

$$\begin{cases} \alpha_1(\theta) = a_1 \\ \dots \\ \alpha_k(\theta) = a_k \end{cases}$$

Примеры 27.1.

1. Пусть $X_1, \dots, X_n \sim N(\alpha, \sigma^2)$. Имеем систему:

$$\begin{cases} \alpha = \bar{X} \\ \sigma^2 + \alpha^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 \end{cases}$$

Решив систему, получим оценку (\bar{X}, s^2) .

2. Пусть $X_1, \dots, X_n \sim U_{[0, \alpha]}$. $\frac{\alpha}{2} = \bar{X}$, откуда $\alpha = 2\bar{X}$.

Метод максимального правдоподобия В случае абсолютно непрерывного распределения и $X_1, \dots, X_n \sim F(x, \theta)$, $p(x, \theta)$ — плотности, будем рассматривать *функцию правдоподобия*:

$$L(X_1, \dots, X_n, \theta) = \prod_{i=1}^n p(X_i, \theta)$$

Возьмём $\tilde{\theta} : L(X_1, \dots, X_n, \tilde{\theta}) \geq L(X_1, \dots, X_n, \theta) \quad \forall \theta$. $\tilde{\theta}$ будет оценкой для θ , полученной методом максимального правдоподобия.

$$\frac{dL(X_1, \dots, X_n, \theta)}{d\theta} = 0$$

В случае дискретного распределения функция правдоподобия:

$$L(X_1, \dots, X_n, \theta) = \prod_{i=1}^n P_\theta(X = X_i)$$

И будем брать $\tilde{\theta}$ точкой максимума $L(X_1, \dots, X_n, \theta)$.

Пример 27.2. Предположим $X_i \sim \mathcal{U}_{[0,a]}$ и оценим a .

$$p(x, a) = \frac{1}{a} \mathbb{I}_{[0,a]}(x)$$

$$L(X_1, \dots, X_n, a) = \frac{1}{a^n} \prod_{i=1}^n \mathbb{I}_{[0,a]}(X_i)$$

Построим вариационный ряд, то есть упорядочим X_i : $X_{(1)}, \dots, X_{(n)}$. Тогда:

$$L(X_1, \dots, X_n, a) = \frac{1}{a^n} \prod_{i=1}^n \mathbb{I}_{[0,a]}(X_{(i)})$$

Таким образом, $L = 0$ при $a < X_{(n)}$, в точке $a = X_{(n)}$ будет $L = \frac{1}{a^n}$, после чего с увеличением a L убывает, отсюда $\tilde{a} = X_{(n)}$.

Заметим, что полученная оценка отличается от той, что мы получили методом моментов.

Пусть \tilde{g}_1, \tilde{g}_2 — оценки параметра θ . Пусть мы получили:

$$E(\tilde{g}_1 - \theta)^2 \geq E(\tilde{g}_2 - \theta)^2 \quad \forall \theta$$

Это означает, что оценка \tilde{g}_2 более эффективна.

Упражнение 27.3. Предположим $X_1, \dots, X_n \sim p(x, \alpha) = \frac{1}{2} e^{-|x-\alpha|}$. Оценить α методами моментов и максимального правдоподобия. Отметим, что сложность последнего в данном случае заключается в том, что приравнивание производной к нулю не работает, так как функция не дифференцируема в этой точке.

28 Проверка статистических гипотез

Как обычно, рассматриваем случайную выборку X_1, \dots, X_n с генеральной функцией распределения F . Пусть \mathfrak{F} — некоторое множество функций распределения, представимое в виде объединения двух непересекающихся подмножеств $\mathfrak{F}_0, \mathfrak{F}_1$: $\mathfrak{F}_0 \cap \mathfrak{F}_1 = \mathfrak{F}$, $\mathfrak{F}_0 \cap \mathfrak{F}_1 = \emptyset$. Может возникнуть две гипотезы: $H_0 = \{F \in \mathfrak{F}_0\}$ или $H_1 = \{F \in \mathfrak{F}_1\}$. Например, H_0 («нулевая гипотеза») могут быть такого вида:

- $F \in \mathcal{N}$;
- $F \in \mathcal{N}(0, 4)$;
- F таково, что математическое ожидание положительно.

Методология проверки гипотез такова: задаётся некоторое близкое к нулю положительное α . На практике это может быть $\alpha = 0.1, \alpha = 0.05, \alpha = 0.01, \dots$

	Верна гипотеза H_0	Верна гипотеза H_1
Принята гипотеза H_0	Верное решение	Ошибка II рода
Принята гипотеза H_1	Ошибка I рода	Верное решение

В качестве основной («нулевой») гипотезы обычно берут ту, для которой вероятность ошибки первого рода не превосходит α , но при этом вероятность ошибки второго рода минимизируется.

Рассмотрим основные типы задач на проверку гипотез.

1. Проверка гипотезы согласия. $X_1, \dots, X_n \sim F$ и F_0 — некоторая известная функция. Верётся $H_0 = \{F = F_0\}$, $H_1 = \{F \neq F_0\}$
2. Проверка однородности. $X_1, \dots, X_n \sim F$, $Y_1, \dots, Y_n \sim G$. Верётся $H_0 = \{F = G\}$, $H_1 = \{F \neq G\}$.
3. Проверка независимости. $(X_1, Y_1), \dots, (X_n, Y_n) \sim F(x, y)$, где $X_1, \dots, X_n \sim G(x)$, $Y_1, \dots, Y_n \sim H(y)$. Верётся $H_0 = \{F(x, y) = G(x)H(y)\}$, $H_1 = \{\exists x_0, y_0 : F(x_0, y_0) \neq G(x_0)H(y_0)\}$.

Изучим некоторые инструменты, помогающие проверить гипотезу согласия.

Критерий Колмогорова

Рассмотрим выборочную функцию распределения:

$$\bar{F}_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{(-\infty, t)}(X_i)$$

И статистику Колмогорова:

$$D_n = \sup_{t \in \mathbb{R}} |\bar{F}_n(t) - F_0(t)|$$

Теорема 28.1 (Колмогорова). Если H_0 верна, то $P(\sqrt{n}D_n < x) \xrightarrow{n \rightarrow \infty} K(x)$. $K(x)$ — функция распределения Колмогорова.

Алгоритм:

1. Вычисляем D_n .
2. По таблицам $K(x)$ находим $z_{1-\alpha}$ такое, что $K(z_{1-\alpha}) = 1 - \alpha$.
3. Сравниваем: если $\sqrt{n}D_n \geq z_{1-\alpha}$, то H_0 отвергаем; если $\sqrt{n}D_n < z_{1-\alpha}$, то принимаем.

Критерий хи-квадрат для проверки гипотезы согласия Разбиваем вещественную прямую на r зон A_1, \dots, A_r ($5 \leq r \leq 15$). Считаем вероятности $p_k = P_{H_0}(X \in A_k)$. Считаем $v_k = \sum_{i=1}^n \mathbb{I}_{A_k}(X_i)$ — число элементов выборки, попавших в зону A_k . Построим статистику $\chi_n^2 = \sum_{i=1}^r \frac{(v_i - np_i)^2}{np_i}$. Справедлива следующая теорема.

Теорема 28.2 (Пирсона). Если H_0 верна, то $P(\chi_n^2 < x) \xrightarrow{n \rightarrow \infty} K_{r-1}(x)$. $K_{r-1}(x)$ — функция закона хи-квадрат с $r - 1$ степенью свободы.

Алгоритм:

1. Вычисляем χ_n^2 .
2. По таблицам $K_{r-1}(x)$ находим $z_{1-\alpha}$ такое, что $K(z_{1-\alpha}) = 1 - \alpha$.
3. Сравниваем: если $\chi_n^2 \geq z_{1-\alpha}$, то H_0 отвергаем; если $\chi_n^2 < z_{1-\alpha}$, то принимаем.

Замечание 28.3. В случае, когда несколько критериев дают принятие гипотезы согласия, но один не даёт, мы эту гипотезу отвергаем.

Критерий хи-квадрат для проверки независимости

Проведём разбиение на зоны, аналогичное тому, что мы использовали для проверки гипотезы согласия. Построим таблицу, обозначив за n_{ij} количество элементов, попавших в соответствующую зону:

$Y \setminus X$	A_1	A_2	\dots	A_r	сумма по строке
B_1	n_{11}	n_{21}	\dots	n_{r1}	$n_{\bullet 1}$
B_2	n_{12}	n_{22}	\dots	n_{r2}	$n_{\bullet 2}$
\dots	\dots	\dots	\dots	\dots	\dots
B_s	n_{1s}	n_{2s}	\dots	n_{rs}	$n_{\bullet s}$
сумма по столбцу	$n_{1\bullet}$	$n_{2\bullet}$	\dots	$n_{r\bullet}$	n

$$\chi_n^2 = \sum_{i,j} \frac{\left(n_{ij} - \frac{n_{i\bullet} n_{\bullet j}}{n}\right)^2}{n_{i\bullet} n_{\bullet j}}$$

Теорема 28.4. Если H_0 верна, то $P(\chi_n^2 < \chi) \xrightarrow{n \rightarrow \infty} K_{(r-1)(s-1)}(\chi)$.

Пусть имеется случайная выборка X_1, \dots, X_n и вариационный ряд $X_{(1)}, \dots, X_{(n)}$. Пусть R_i — место, которое занимает величина X_i в вариационном ряду, называемое *рангом*. Итак по имеющейся выборке мы можем построить вариационный ряд и вектор рангов (R_1, \dots, R_n) . Очевидно, что мы также можем восстановить выборку по её вариационному ряду и вектору рангов. Оказывается, для решения многих задач достаточно знать только вектор рангов.

Критерий Вилкоксона для проверки однородности $X_1, \dots, X_n \sim F, Y_1, \dots, Y_m \sim G$. Будем проверять гипотезы $H_0 = \{F = G\}, H_1 = \{F \neq G\}$. Объединим выборки X_i и Y_j и построим ранги R_i только для X_i . Возьмём $w_{n,m} = \sum_{i=1}^n R_i$.

Теорема 28.5. Если H_0 верна, то:

1. $E w_{n,m} = \frac{n(n+m+1)}{2}$
2. $D w_{n,m} = \frac{n(n+m+1)}{12}$
3. $P\left(\frac{w_{n,m} - E w_{n,m}}{\sqrt{D w_{n,m}}} < \chi\right) \xrightarrow{n \rightarrow \infty, m \rightarrow \infty} \Phi(\chi)$

Алгоритм:

1. По α находим $z_{1-\frac{\alpha}{2}}$ такое, что $\Phi(z_{1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$.
2. Если $\frac{w_{n,m} - E w_{n,m}}{\sqrt{D w_{n,m}}} \geq z_{1-\frac{\alpha}{2}}$, то отвергаем H_0 . Иначе — оставляем.

Критерий Спирмена для проверки независимости Пусть есть выборка $(X_1, Y_1), \dots, (X_n, Y_n)$, причём R_i — ранг X_i , S_i — ранг Y_i .

$$d_n^2 = \sum_{i=1}^n (R_i - S_i)^2$$

$$r_n = 1 - \frac{6d_n^2}{n(n-1)(n+1)}$$

r_n называется *коэффициентом корреляции Спирмена*. Он действительно является коэффициентом корреляции.

Теорема 28.6. Если H_0 верна, то $P(\sqrt{n-1}r_n < x) \xrightarrow{n \rightarrow \infty} \Phi(x)$;

Алгоритм:

1. По α находим $z_{1-\frac{\alpha}{2}}$ такое, что $\Phi(z_{1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$.
2. Если $|\sqrt{n-1}r_n| \geq z_{1-\frac{\alpha}{2}}$, то отвергаем H_0 . Иначе — оставляем.