Санкт-Петербургский Государственный Университет

Факультет математики и компьютерных наук

Теория вероятностей

Конспект основан на лекциях
Александра Игоревича Храброва

18 июня 2022 г.

Санкт-Петербургский государственный университет

Факультет математики и компьютерных наук СПбГУ

Конспект основан на лекциях по математическому анализу, прочитанных Александром Игоревичем Храбровым студентам Факультета математики и компьютерных наук Санкт-Петербургского государственного университета в весеннем семестре 2019–2020 учебного года.

В конспекте содержится материал 4-го семестра курса теории вероятностей.

Авторы:

Ольга Шиманская
Андрей Кислицын

Консультант:

Михаил Опанасенко

Последняя версия конспекта и исходный код:

https://www.overleaf.com/read/cnrvskqzwhcw

Сайт СПБГУ: https://spbu.ru.
Сайт факультета МКН: https://math-cs.spbu.ru.

Оглавление

§1. Конечное вероятностное пространство. Свойства вероятности. Классическое определение вероятности

Определение. $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ — *множество (пространство) элементарных событий*, если

1. ω_i – равновозможны;
2. ω_i и ω_j не реализуемы одновременно (несовместны);
3. Какая-то ω_i случается;

Примеры 1.1.

1. Монетка — орёл или решка(1/0);
2. Игральный кубик;
3. Колода карт;

Здесь и далее за $\#A$ обозначается мощность множества A.

Определение. *Случайное событие* — это некоторое $A \subset \Omega$. *Вероятность случайного события* — это $P(A) = \frac{\#A}{\#\Omega}$.

Для примеров пространств элементарных событий приведём примеры случайных событий:

Примеры 1.2.

1. Орёл/решка;
2. Чётное число очков/число очков больше трёх;
3. Пики/красные старше валета;

Свойства 1.1 (вероятности).

1. $P(\varnothing) = 0,\ P(\Omega) = 1$;
2. Если $A \cap B = \varnothing$ (говорят, что они *несовместны*), то $P(A \cup B) = P(A) + P(B)$;
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$;
4. $P(A \cup B) \le P(A) + P(B)$;
5. *Формула включений-исключений*:

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_i P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \ldots + (-1)^{n-1} P(A_1 \cap \ldots \cap A_n).$$

6. $P(\overline{A}) = 1 - P(A) = P(\Omega) - P(A)$, где $\overline{A} = \Omega \setminus A$;

Доказательство. 5. Индукция. База — пункт 3. Переход $n \to n+1$. Обозначим $B = A_1 \cup \ldots \cup A_n$. Тогда по 3-му пункту

$$P(B \cup A_{n+1}) = P(B) + P(A_{n+1}) - P(B \cap A_{n+1}).$$

Заметим, что

$$B \cap A_{n+1} = \bigcup_{i=1}^{n} (A_i \cap A_{n+1}).$$

Тогда по индукционному переходу

$$P(B \cap A_{n+1}) = \sum_{i \le n} P(A_i \cap A_{n+1}) - \sum_{i<j \le n} P(A_i \cap A_j \cap A_{n+1}) + \ldots.$$

6. Следствие из первого и третьего. ■

§2. Условная вероятность. Мотивировка, определение и свойства. Пример

Определение. *Условная вероятность*. Пусть $A \neq \varnothing$, $P(A) > 0$. Тогда вероятность B при условии A — это

$$P(B|A) = \frac{\#(A \cap B)}{\#A} = \frac{\#(A \cap B)/\#\Omega}{\#A/\#\Omega} = \frac{P(A \cap B)}{P(A)}.$$

Свойства 2.1 (условной вероятности).

1. $P(A|A) = 1$. Если $A \subset B$, то $P(B|A) = 1$.

2. $P(\varnothing|A) = 0$;

3. Если $B \cap C = \varnothing$, то
$$P(B \cup C|A) = P(B|A) + P(C|A).$$

Доказательство. Докажем пункт 3.

$$\begin{aligned} P(B \cup C|A) &= \frac{P((B \cup C) \cap A)}{P(A)} = \frac{P((B \cap A) \cup (C \cap A))}{P(A)} \\ &= \frac{P(A \cap B) + P(C \cap A)}{P(A)} = P(B|A) + P(C|A). \end{aligned}$$

■

§3. Формула полной вероятности. Формула и теорема Байеса. Примеры

Теорема 3.1 (формула полной вероятности). Пусть $\Omega = \bigsqcup_{k=1}^m A_k$. Тогда

$$P(B) = \sum_{k=1}^m P(B|A_k) \cdot P(A_k).$$

В частности, если $0 < P(A) < 1$, то

$$P(B) = P(B|A) \cdot P(A) + P(B|\overline{A}) \cdot P(\overline{A}).$$

Доказательство.

$$P(B) = \sum_{k=1}^m P(B \cap A_k) = \sum_{k=1}^m \frac{P(B \cap A_k)}{P(A_k)} \cdot P(A_k).$$

■

Пример 3.1. Есть две урны. В первой 3 белых и 5 чёрных шаров, во второй 5 белых и 5 чёрных шаров. Кладём из первой во вторую два шара, и берём шар из второй. Какова вероятность, что он белый (обозначим за B)? Обозначим за A_i событие "взяли i белых шаров из первой". Тогда

$$P(B) = P(B|A_0)P(A_0) + P(B|A_1)P(A_1) + P(B|A_2)P(A_2).$$

$$P(B|A_0) = \frac{5}{12}.$$

$$P(B|A_1) = \frac{1}{2}.$$

$$P(B|A_2) = \frac{7}{12}.$$

$$P(A_0) = \frac{C_5^2}{C_8^2} = \frac{5}{14}.$$

$$P(A_1) = \frac{3 \cdot 5}{C_8^2} = \frac{15}{28}.$$

$$P(A_2) = \frac{C_3^2}{C_8^2} = \frac{3}{28}.$$

$$P(B) = \frac{23}{48}.$$

Теорема 3.2 (формула Байеса). Если $P(A), P(B) > 0$, то

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

Доказательство.

$$\frac{P(B|A)P(A)}{P(B)} = \frac{\frac{P(A\cap B)}{P(A)} \cdot P(A)}{P(B)} = P(A|B).$$

■

Теорема 3.3 (Байеса). Пусть $\Omega = \bigsqcup_{k=1}^{m} A_k$, $P(B) > 0$ и $P(A_k) > 0$. Тогда

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_{i=1}^{m} P(B|A_i)P(A_i)}.$$

Пример 3.2. Турнир по олимпийской системе (проигравший выбывает). 16 участников, из них 2 сестры, известно, что они сыграли друг с другом. Какова вероятность, что этот матч был финальным? Обозначим события B – сёстры сыграли между собой (их матч состоялся), A_i — матч мог состояться в i-м туре.

$$P(A_4|B) = \frac{P(B|A_4)P(A_4)}{\sum P(B|A_i)P(A_i)}.$$
$$P(A_1) = \frac{1}{15}.$$
$$P(A_2) = \frac{2}{15}.$$
$$P(A_3) = \frac{4}{15}.$$
$$P(A_4) = \frac{8}{15}.$$
$$P(B|A_1) = 1.$$
$$P(B|A_2) = \frac{1}{4}.$$
$$P(B|A_3) = \frac{1}{16}.$$
$$P(B|A_4) = \frac{1}{64}.$$
$$P(A_4|B) = \frac{1}{15}.$$

§4. Независимые события. Мотивировка и определение. Примеры. Попарная независимость и независимость в совокупности. Примеры

Определение. Случайные события A и B *независимы*, если $P(A \cap B) = P(A) \cdot P(B)$.

Равносильное определение:

$$P(A \cap B) = P(A)P(B) \iff P(B|A) = \frac{P(A \cap B)}{P(A)} = P(B).$$

Определение. $A_1, \ldots, A_m$ *независимы в совокупности*, если

$$\forall i_1, \ldots, i_k \; P(A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot \ldots \cdot P(A_{i_k}).$$

Лемма 4.1. $A_1, \ldots, A_m$ независимы в совокупности $\implies P(B_1 \cap \ldots \cap B_m) = P(B_1) \cdot \ldots \cdot P(B_m)$, где $B_i = A_i$ или $\overline{A_i}$

Отметим, что независимость в совокупности неравносильна попарной независимости. Приведём пример. Пространство — множество пар чисел при кидании двух кубиков. Обозначим события A — чётное на первом, B — чётная на втором, C — чётная сумма.

$$P(A) = P(B) = P(C) = \frac{1}{2}.$$
$$P(A \cap C) = P(B \cap C) = P(A \cap B) = \frac{1}{4}.$$

Значит, эти события попарно независимы.

$$P(A \cap B \cap C) = P(A \cap B) = \frac{1}{4}.$$
$$P(A)P(B)P(C) = \frac{1}{8},$$

то есть они не независимы в совокупности.

Определение. *B не зависит от совокупности событий* $A_1, \ldots, A_m$, если

$$\forall i_1, \ldots, i_k \; P(B|A_{i_1}, \ldots, A_{i_k}) = P(B) \iff P(B \cap A_{i_1} \cap \ldots \cap A_{i_k}) = P(B) \cdot P(A_{i_1} \cap \ldots \cap A_{i_k}).$$

Теорема 4.2 (Эрдёша-Мозера). В турнире участвует n волейбольных команд. Играют каждая с каждой, без ничей. Пусть k — наибольшее число, для которого всегда найдутся такие команды $a_1, \ldots, a_k$, что a_i выиграла у a_j, если $i < j$. Тогда $k \leq 1 + [2 \log_2 n]$.

Доказательство. Турнир — полный орграф (стрелочки от победителей к проигравшим). Подходящая цепочка — полный ациклический подграф. Пусть событие $A(a_1, \ldots, a_k)$ — подошёл набор $a_1, \ldots a_k$. Тогда $P(A) = 2^{-\binom{k}{2}}$. Способов выбрать набор (выбрать k команд и порядок на них) — $\binom{n}{k} \cdot k!$. Вероятность того, что какой-то набор подойдёт не превосходит

$$2^{-\binom{k}{2}} \cdot \binom{n}{k} \cdot k!.$$

Докажем, что если $k > 1 + [2 \log_2 n]$, то это значение меньше единицы, то есть существует граф, на котором нет такого набора.

$$k > 1 + [2 \log_2 n] \implies \log_2 n < \frac{k-1}{2} \implies n < 2^{\frac{k-1}{2}}.$$
$$2^{-\binom{k}{2}} \cdot \binom{n}{k} \cdot k! = 2^{-\frac{k \cdot (k-1)}{2}} \cdot \frac{n!}{(n-k)!} < 2^{-\frac{k \cdot (k-1)}{2}} \cdot n^k < 1.$$

■

§5. Схема Бернулли. Полиномиальная схема. Теорема Эрдёша–Мозера

Определение. *Схема Бернулли.* Элементарные события в пространстве имеют вид $\omega = (x_1, \ldots, x_n)$, где x_i принимает значение 0 или 1; определим вероятность как

$$P(\omega) = p^{\sum x_i} q^{n - \sum x_i}, \text{ где } p \in [0,1], \ q = 1 - p.$$

Смысл такой — рассмотрим модель, в которой мы делаем n подбрасываний монетки, при которых орёл выпадает с вероятностью p, а решка с вероятность $q = 1 - p$. Тогда элементарные события — это все возможные исходы; нетрудно проверить, что вероятность исхода совпадает с вероятностью из определения.

Подсчитаем вероятность выпадения k орлов. Надо сложить вероятность по всем возможным ω, в которых ровно k орлов(они имеют одинаковые вероятности).

$$P(k \text{ орлов}) = \binom{n}{k} p^k q^{n-k}.$$

Определение. *Полиномиальная схема.* Элементарные события в пространстве имеют вид $\omega = (x_1, \ldots, x_n)$, где x_i принимают целочисленные значения от 1 до m; определим вероятность как

$$P(\omega) = p_1^{\#\{i | x_i = 1\}} \cdot p_2^{\#\{i | x_i = 2\}} \cdot \ldots \cdot p_m^{\#\{i | x_i = m\}}, \text{ где } p_i \geq 0, \ \sum p_i = 1.$$

Рассмотрим модель, аналогичную предыдущей, но x_i принимает значение k с вероятностью p_k.

$$P(k_1 \text{ раз } 1, \ldots, k_m \text{ раз } m) = p_1^{k_1} \cdot \ldots \cdot p_m^{k_m} \cdot \binom{n}{k_1, \ldots, k_m}, \text{ где}$$

$$\binom{n}{k_1, \ldots, k_m} = \frac{n!}{k_1! \ldots k_m!} \text{ — мультиномиальный коэффициент.}$$

§6. Теоремы Пуассона и Прохорова (вторая без доказательства). Пример

Теорема 6.1 (Пуассона). Пусть p_n — вероятность успеха в n-ой схеме Бернулли, $n \cdot p_n \to \lambda$; S_n — количество успехов, тогда при $k = o(\sqrt{n})$, $k = o(\frac{1}{n \cdot p_n - \lambda})$ верно

$$P(S_n = k) \to \frac{\lambda^k}{k!} e^{-\lambda}.$$

Лемма 6.2.

$$(1-a_1)\cdot(1-a_2)\cdot\ldots\cdot(1-a_n) \geq 1-a_1-\ldots-a_n, \quad \text{где} \quad 0 < a_1,\ldots,a_n < 1.$$

Доказательство. Индукция по n. ■

Следствие 6.3.

$$\binom{n}{k} \underset{n\to\infty}{\sim} \frac{n^k}{k!} \quad \text{при} \quad k = o(\sqrt{n}).$$

Доказательство. С одной стороны,

$$\binom{n}{k} = \frac{n\cdot(n-1)\ldots(n-k+1)}{k!} \leq \frac{n^k}{k!}.$$

С другой стороны, по лемме имеем, что

$$\binom{n}{k} = \frac{n^k}{k!}\cdot(1\cdot(1-\frac{1}{n})\ldots(1-\frac{k-1}{n})) \geq \frac{n^k}{k!}\cdot(1-\frac{1}{n}-\ldots-\frac{k-1}{n}) = \frac{n^k}{k!}\cdot(1-\frac{k(k-1)}{2n}) \to \frac{n^k}{k!}.$$

■

Доказательство теоремы 6.1(Пуассона). По следствию из леммы получаем, что

$$P(S_n=k) = \binom{n}{k}p_n^k(1-p_n)^{n-k} \sim \frac{n^k}{k!}p_n^k(1-p_n)^{n-k} \sim \frac{\lambda^k}{k!}(1-p_n)^{n-k}.$$

Осталось показать, что

$$(1-p_n)^{n-k} \overset{?}{\sim} e^{-\lambda},$$

а это равносильно тому, что

$$(n-k)\ln(1-p_n) \overset{?}{\sim} -\lambda.$$

Это верно, поскольку $n-k \sim n$ и $\ln(1-p_n) \sim -p_n$, а по условию $np_n \sim \lambda$. ■

Теорема 6.4 (Прохорова). Пусть вероятность в n-й схеме равна $\frac{\lambda}{n}$. Тогда

$$\sum_{k=0}^{\infty}\left|P(S_n=k) - \frac{\lambda^k}{k!}e^{-\lambda}\right| \leq \frac{2\lambda}{n}\min\{2,\lambda\}.$$

Пример 6.1. Модель — рулетка: есть целые числа от 0 до 36. Игрок играет $n = 111$ раундов, каждый раз ставит одну монетку (считаем, что число монет неограниченно), если угадывает, то получает выигрыш в 37 раз больше, то есть 37 монеток. Понятно, что для того, чтобы “отбить”все потраченные монетки, нужно выиграть 3 раза. Посчитаем вероятность этого. Очевидно, если игрок ставит равновероятно, то вероятность выигрыша в раунде равна $p = \frac{1}{37}$. Посчитаем напрямую:

$$P(S_{111}=3) = \binom{111}{3}\left(\frac{1}{37}\right)^3\left(1-\frac{1}{37}\right)^{111-3} \approx 0,2271\ldots.$$

Если воспользуемся теоремой Пуассона, то получим оценку

$$P(S_{111}=3)\approx 0,224\ldots.$$

Также оценим шанс выигрыша(“выйти в плюс“):

$$\begin{aligned}P(\text{win}) &= 1-P(S_n=0)-P(S_n=1)-P(S_n=2)-P(S_n=3)\\ &\approx 1-e^{-3}-3e^{-3}-\frac{9}{2}e^{-3}-\frac{9}{2}e^{-3}\\ &= 1-13e^{-3}\approx 0,352\ldots.\end{aligned}$$

§7. Локальная теорема Муавра–Лапласа. Пример

Теорема 7.1 (Локальная предельная теорема Муавра-Лапласа). Пусть $0<p<1$, $q=1-p$. T — некоторое число. Обозначим $x=\frac{k-np}{\sqrt{npq}}$, причём k меняется так, что $|x|\le T$ при $n\to+\infty$. Тогда

$$P(S_n=k)\underset{n\to+\infty}{\sim}\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi npq}}\text{ равномерно по } k.$$

Доказательство. Пусть $n\to\infty$. Из условий

$$np+T\sqrt{npq}\ge k=np+x\sqrt{npq}\ge np-T\sqrt{npq}.$$

Тогда $k\to+\infty$ верно из 2-го неравенства и $n-k\to+\infty$ из 1-го (если из n вычесть $np+T\sqrt{npq}$, то это будет стремится к $+\infty$). Обозначим

$$\alpha=\frac{k}{n}=p+x\sqrt{\frac{pq}{n}}\to p\quad\text{ и }\quad\beta=\frac{n-k}{n}=1-\alpha=q-x\sqrt{\frac{pq}{n}}\to q.$$

Тогда

$$\begin{aligned}P(S_n=k)=\binom{n}{k}p^kq^{n-k} &\sim \frac{n^ne^{-n}\sqrt{2\pi n}p^kq^{n-k}}{k^ke^{-k}\sqrt{2\pi k}(n-k)^{n-k}e^{k-n}\sqrt{2\pi(n-k)}}\\ &= \frac{p^kq^{n-k}}{(\frac{k}{n})^k(1-\frac{k}{n})^{n-k}\sqrt{2\pi n}\sqrt{\frac{k}{n}(1-\frac{k}{n})}}\\ &\sim \frac{p^kq^{n-k}}{\alpha^k\beta^{n-k}\sqrt{2\pi npq}}.\end{aligned}$$

Надо доказать, что

$$\frac{p^kq^{n-k}}{\alpha^k\beta^{n-k}}\sim e^{-\frac{x^2}{2}},$$

что равносильно тому, что

$$k \ln\left(\frac{\alpha}{p}\right) + (n-k)\ln\left(\frac{\beta}{q}\right) \sim \frac{x^2}{2}.$$

$$\frac{\alpha}{p} = 1 + x\sqrt{\frac{q}{np}},\ \ln\left(\frac{\alpha}{p}\right) = x\sqrt{\frac{q}{np}} - x^2\frac{q}{2np} + O\left(\frac{1}{n^{\frac{3}{2}}}\right).$$

$$\frac{\beta}{q} = 1 - x\sqrt{\frac{pq}{n}},\ \ln\left(\frac{\beta}{q}\right) = -x\sqrt{\frac{p}{nq}} - x^2\frac{p}{2nq} + O\left(\frac{1}{n^{\frac{3}{2}}}\right).$$

Итого, подставив, получаем

$$\begin{aligned} k \ln\left(\frac{\alpha}{p}\right) + (n-k)\ln\left(\frac{\beta}{q}\right) &= (np + x\sqrt{npq})\left(x\sqrt{\frac{q}{np}} - x^2\frac{q}{2np} + O\left(\frac{1}{n^{\frac{3}{2}}}\right)\right) \\ &\quad + (nq - \sqrt{npq})\left(-x\sqrt{\frac{p}{nq}} - x^2\frac{p}{2nq} + O\left(\frac{1}{n^{\frac{3}{2}}}\right)\right) \\ &= x\sqrt{npq} + x^2 q - x^2\frac{q}{2} + O\left(\frac{1}{\sqrt{n}}\right) \\ &\quad - x\sqrt{npq} + x^2 p - x^2\frac{p}{2} + O\left(\frac{1}{\sqrt{n}}\right) \\ &= \frac{x^2}{2} + O\left(\frac{1}{\sqrt{n}}\right). \end{aligned}$$

■

Пример 7.1. На рулетке 37 секторов — 18 красных и 18 черных и 1 зелёный. Игрок участвует в $n = 222$ раундах. Посчитаем шанс “отбить:

$$P(S_{222} = 111) = \binom{222}{111}(\frac{18}{37})^{111}(\frac{19}{37})^{111} \approx 0,0493228\ldots.$$

Теорема Муавра-Лапласа дает нам оценку

$$P(S_{222} = 111) \approx \frac{e^{-\frac{(k-np)^2}{2npq}}}{\sqrt{2\pi npq}} \approx 0,0493950\ldots.$$

§8. Интегральная теорема Муавра–Лапласа и оценка на скорость сходимости (без доказательства). Неулучшаемость показателя степени в оценке. Задача о театре. Случайное блуждание на прямой

Теорема 8.1 (Интегральная предельная теорема Муавра-Лапласа). Пусть $0 < p < 1$. Тогда

$$P(a < \frac{S_n - np}{\sqrt{npq}} \le b) \underset{n\to\infty}{\to} \frac{1}{\sqrt{2\pi}} \int\limits_a^b e^{-\frac{x^2}{2}}\,\mathrm{d}x \text{ равномерно по } a, b \in \mathbb{R}.$$

Обозначение.

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^x e^{-\frac{t^2}{2}}\,\mathrm{d}t.$$

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{t^2}{2}}\,\mathrm{d}t.$$

Эти функции не выражаются через элементарные. Из матанализа знаем значения в некоторых точках. Например про второй знаем, что $\Phi_0(x) \approx \frac{1}{2}$ при $x > 4$.

Теорема 8.2 (Оценка скорости сходимости. Частный случай теоремы Берри-Эссена).

$$\sup_{x\in\mathbb{R}} |P(\frac{S_n - np}{\sqrt{npq}} \le x) - \Phi(x)| \le \frac{1}{2}\frac{p^2+q^2}{\sqrt{npq}}.$$

Пример 8.1 (Неулучшаемость оценки). Приведем пример, показывающий, что сходимость не может быть быстрее, чем $\frac{C}{\sqrt{n}}$.

Пусть $p = q = \frac{1}{2}, x = 0$. Точное значение $\Phi(x)$ мы знаем: $\Phi(0) = \frac{1}{2}$, так как это половина всего распределения. Теперь оценим вероятность:

$$P\left(\frac{S_n - \frac{n}{2}}{\sqrt{n\frac{1}{4}}} \le 0\right) = P\left(S_n \le \frac{n}{2}\right) = P(S_{2n} \le n) =$$

Заметим, что $P(S_{2n} \le n) = P(S_{2n} \ge n)$, причем объединение этих двух событий дает все вероятностное пространство за исключением того, что событие $P(S_{2n} = n)$ посчитано 2 раза. Отсюда вытекает равенство:

$$= \frac{1}{2} + \frac{P(S_{2n} = n)}{2} = \frac{1}{2} + \frac{1}{2}\cdot\binom{2n}{n}\cdot\frac{1}{4^n}$$

Осталось вспомнить, что $\binom{2n}{n} \cdot \frac{1}{4^n} \sim \frac{1}{\sqrt{\pi n}}$, поэтому:

$$\left| P\left(\frac{S_n - \frac{n}{2}}{\sqrt{n\frac{1}{4}}} \leq 0 \right) - \Phi(0) \right| \sim \frac{1}{\sqrt{\pi n}}$$

Пример 8.2 (задача о театре). В театре, рассчитанном на $n = 1600$ мест, есть два гардероба. Сколько должно быть мест в каждом гардеробе, чтобы в среднем не чаще раза в месяц какому-то посетителю пришлось идти не к ближайшему из-за отсутствия в нем мест(считаем, что люди заходят в гардеробы равновероятно)? Обозначим число вешалок в каждом из гардеробов за C. Пусть S_n — число людей, сдавших вещи в первый гардероб. Тогда должны выполняться неравенства $S_n \leq C$ и $n - S_n \leq C$. Хотим, чтобы выполнялось

$$P(n - C \leq S_n \leq C) \approx \frac{29}{30}.$$

По теореме 8.1

$$\begin{aligned} P(n - S_n \leq S_n \leq C) &= P(\frac{-C + \frac{n}{2}}{\sqrt{npq}} \leq \frac{S_n - np}{\sqrt{npq}} \leq \frac{C - \frac{n}{2}}{\sqrt{npq}}) \\ &= P(\frac{-C + 800}{20} \leq \frac{S_n - np}{\sqrt{npq}} \leq \frac{C - 800}{20}) \approx \frac{1}{\sqrt{2\pi}} \int\limits_{-\frac{C-800}{20}}^{\frac{C-800}{20}} e^{-\frac{t^2}{2}} \, \mathrm{d}t \\ &= 2\Phi_0(\frac{C - 800}{20}) \approx \frac{29}{30} \implies C \approx 843. \end{aligned}$$

Пример 8.3 (Случайное блуждание на прямой). Ходим по прямой, начиная с 0. Идем на один шаг вправо с вероятностью p, и на один шаг влево с вероятностью $q = (1 - p)$. Заметим, что точка, в которую мы придем, выражается как $a_n = 2S_n - n$. Тогда вероятность придти в конкретную точку после n шагов вычисляется как: $P(a_n = k) = P(S_n = \frac{k+n}{2}) = \binom{n}{\frac{n+k}{2}} p^{\frac{n+k}{2}} q^{\frac{n-k}{2}}$, при условии, что n и k одной четности.

§9. Вероятностное пространство. Условная вероятность. Независимые события

Определение. *Вероятностное пространство* — это $(\Omega, \mathcal{F}, P)$, где Ω — множество элементарных событий; $\mathcal{F}$ — σ-алгебра подмножеств Ω, его элементы — *случайные события*; P — вероятностная мера на $\mathcal{F}$, такая что $P(\Omega) = 1$.

Определение. *Условная вероятность.* Пусть $B \in \mathcal{F}, P(B) > 0$. Тогда $P(A|B) = \frac{P(A \cap B)}{P(B)}$ — вероятность A при условии B.

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

Определение. Множество событий A_i по $i \in I$ является независимым в совокупности, если $\forall i_1, ..., i_k \in I$

$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \ldots \cdot P(A_{i_k})$.

§10. Лемма Бореля–Кантелли. Закон нуля и единицы. Пример

Лемма 10.1 (Бореля-Кантелли). $A_1, A_2, \ldots$ — последовательность случайных событий. Событие B — наступило бесконечное число из событий A.

1. Если $\sum_{n=1}^{\infty} P(A_n) < +\infty$, то $P(B) = 0$.
2. Если $A_1, A_2, \ldots$ независимы в совокупности и $\sum_{n=1}^{\infty} P(A_n) = +\infty$, то $P(B) = 1$.

Доказательство. Поймём, что $B = \bigcap\limits_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$. $\omega \in B \implies \omega \in A_{m_k}$ для некоторой последовательности $m_1, m_2, \ldots \Rightarrow \omega$ лежит в каждом из объединений $\bigcup_{k=n}^{\infty} A_k$, то есть и в их пересечении. В обратную сторону: $\omega \in \bigcap \bigcup \implies \omega \in \bigcup_{k=n}^{\infty} A_k \ \forall n \implies \omega \in A_{m_k}$ для некоторой последовательности $m_1, m_2, \ldots$, а это по определению B означает, что $\omega \in B$.

1. По только что доказанному верно

$$P(B) = P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right) = [\text{из свойств меры}]$$
$$= \lim_{n\to\infty} P\left(\bigcup_{k=n}^{\infty} A_k\right) \le \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0.$$

Последнее верно, так как это предел хвоста сходящегося ряда.

2. По лемме 4.1 $\overline{A_1}, \overline{A_2}, \ldots$ независимы в совокупности.

$$P\left(\bigcap_{k=n}^{\infty} \overline{A_k}\right) \leftarrow P\left(\bigcap_{k=n}^{m} \overline{A_k}\right) = \prod_{k=n}^{m} P(\overline{A_k}) \to \prod_{k=n}^{\infty} P(\overline{A_k}) = \prod_{k=n}^{\infty} (1 - P(A_k)).$$
$$\ln P\left(\bigcap_{k=n}^{\infty} \overline{A_k}\right) = \sum_{k=n}^{\infty} \ln(1 - P(A_k)) \le -\sum_{k=n}^{\infty} P(A_k) = -\infty \implies P\left(\bigcap_{k=n}^{\infty} \overline{A_k}\right) = 0.$$
$$\overline{B} = \bigcup_{n=1}^{\infty} \overline{\left(\bigcup_{k=n}^{\infty} A_k\right)} = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \overline{A_k} \implies P(\overline{B}) \le \sum_{n=0}^{\infty} P\left(\bigcap_{k=n}^{\infty} \overline{A_k}\right) = 0.$$

Значит, $P(B) = 1$. ■

Следствие 10.2 (Закон 0 и 1 Колмогорова). Пусть $A_1, A_2, \ldots$ независимы в совокупности. Тогда вероятность того, что наступило бесконечное количество событий из A равно 0 или 1.

Пример 10.1. Рассмотрим бесконечную орлянку: ОРРРРОРРООРР...

Посчитаем вероятность того, что «ОРРО» встретилось бесконечное число раз. Событие A_k — эта последовательность встретилась, причём начиная с позиции k. Тогда $A_1, A_5, A_9, \ldots$ независимы в совокупности. $P(A_k) = p^2q^2$, $\sum_{n=1}^{\infty} P(A_{4n+1}) = +\infty$. Тогда по лемме «ОРРО» случилось бесконечное число раз с вероятностью 1.

§11. Случайная величина. Распределение случайной величины. Свойства функций распределения

Определение. Пусть $(\Omega, \mathcal{F}, P)$ — вероятностное пространство. *Случайной величиной* называется измеримая функция $\xi\colon \Omega \to \mathbb{R}$.

Определение. *Распределение случайной величины* $P_\xi(A)$ — это мера на борелевских подмножествах, определённая следующим образом:

$$P_\xi(A) = P(\xi \in A) = P(\{\omega \in \Omega \mid \xi(\omega) \in A\}).$$

Докажем корректность определения. Знаем, достаточно определить P на ячейках:

$$P_\xi(a, b] = P(a < \xi \le b) = P(\xi \le b) - P(\xi \le a).$$

Такие слагаемые определены, так как ξ — измеримая. Очевидно, эти функции однозначно задают распределение. Определим их.

Определение. *Функцией распределения* случайной величины называется

$$F_\xi(x) = P(\xi \le x).$$

Определение. ξ, η имеют *одинаковое распределение*, если $P_\xi = P_\eta$.

Отметим, что это равносильно равенству $P(\xi \le b) = P(\eta \le b)$ для всех b.

Свойства 11.1 (функции распределения).

1. F_ξ нестрого монотонно возрастает;
2. $0 \le F_\xi \le 1$;
3. $\lim_{n\to-\infty} F_\xi(x) = 0$ и $\lim_{n\to+\infty} F_\xi(x) = 1$;
4. F_ξ непрерывна справа;
5. $P(\xi < x) = \lim_{y\to x-} F_\xi(y)$;
6. F_ξ непрерывна в точке x_0 равносильно тому, что $P(\xi = x_0) = 0$;
7. $F_{\xi+c}(x) = F_\xi(x - c)$;
8. $F_{c\xi}(x) = F_\xi(\frac{x}{c})$ при $c > 0$;

Доказательство.

3. Пусть $x_n \searrow -\infty$. Тогда множества $\{\xi \leq x_n\}$ вложены, значит можем применить следующее свойство меры:

$$\lim_{n\to-\infty} F_\xi(x) = \lim_{n\to\infty} F_\xi(x_n) = \lim_{n\to\infty} P(\xi \leq x_n) = P\left(\bigcap_{n=1}^{\infty}\{\xi \leq x_n\}\right) = P(\varnothing) = 0.$$

4. Проверим непрерывность в точке x_0, пусть $x_n \searrow x_0$. Тогда

$$\lim_{x\to\ x_0} F_\xi(x) = \lim_{x\to\infty} F_\xi(x_n) = \lim_{x\to\infty} P(\xi \leq x_n) = P\left(\bigcap_{n=1}^{\infty}\{\xi \leq x_n\}\right) = P(\xi \leq x_0).$$

5. Пусть $x_n \nearrow x$, тогда

$$\lim_{y\to x-} F_\xi(y) = \lim_{n\to\infty} F_\xi(x_n) = \lim P(\xi \leq x_n) = P\left(\bigcup_{n=1}^{\infty}\{\xi \leq x_n\}\right) = P(\xi < x).$$

■

Также заметим, что если функция F обладает свойствами 1–4 то это функция распределения для некоторой случайной величины, так как можем определить вероятностную меру как $\mu((a,b]) = F(b) - F(a)$(знаем, что это мера, из теории меры).

§12. Дискретное, непрерывное и абсолютно непрерывное распределения. Свойства

Определение. Случайная величина ξ — *дискретная*(или говорят, что она имеет *дискретное распределение*), если

$$\xi\colon \Omega \to Y,$$

где множество Y не более, чем счётно.

В дискретном вероятностном пространстве распределение устроено следующим образом:

$$P_\xi(A) = \sum_{k: y_k \in A} P(\xi = y_k).$$

Таким образом, распределение полностью определяется вероятностями $P(\xi = y_k)$.

Определение. Случайная величина имеет *непрерывное распределение*, если её функция распределения непрерывна.

Как мы уже знаем, это равносильно тому, что $\forall x \in \mathbb{R}\ P(\xi = x) = 0$.

Определение. Случайная величина ξ имеет *абсолютно непрерывное распределение*, если у её функции распределения есть плотность относительно меры Лебега, то есть $p_\xi(t) \geq 0$ измеримая по Лебегу такая, что

$$P_\xi(A) = \int\limits_A p_\xi(t)\,\mathrm{d}t.$$

Понятно, что для функция распределения такой величины считается как

$$F_\xi(x) = \int\limits_{-\infty}^{x} p_\xi(t)\,\mathrm{d}t.$$

Свойства 12.1 (плотности распределения).

1. $P(\xi \in A) = P_\xi(A) = \int_A p_\xi(t)\,\mathrm{d}t$;
2. $F_\xi(x) = \int_{-\infty}^{x} p_\xi(t)\,\mathrm{d}t$;
3. Если t_0 — точка непрерывности p_ξ, то $p_\xi(t_0) = F'_\xi(t_0)$ (на самом деле равенство есть почти везде, так как монотонно возрастающая функция дифференцируема почти везде);

§13. Примеры вероятностных распределений

Примеры 13.1 (различных распределений).

1. Распределение Бернулли ($\xi \sim \mathrm{Bern}(p)$).

$$0 \leq p \leq 1,\ \xi\colon \Omega \to \{0, 1\},$$
$$P(\xi = 0) = 1 - p,\ P(\xi = 1) = p.$$

2. Биномиальное распределение ($\xi \sim \mathrm{Binom}(p)$).

$$\xi\colon \Omega \to \{0, 1, \ldots, n\},$$
$$P(\xi = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

3. Распределение Пуассона ($\xi \sim \mathrm{Pois}(\lambda)$).

$$\lambda > 0,\ \xi\colon \Omega \to \{0, 1, 2, \ldots\},$$
$$P(\xi = n) = \frac{\lambda^n e^{-\lambda}}{n!}.$$

4. Геометрическое распределение ($\xi \sim \mathrm{Geom}(p)$).

$$0 < p < 1,\ \xi\colon \Omega \to \{1, 2, 3, \ldots\},$$

$$P(\xi = n) = p(1-p)^{n-1}.$$

5. Дискретное равномерное распределение.

$$\xi\colon \Omega \to \{a+1, \dots, b\},$$
$$P(\xi = n) = \frac{1}{b-a}.$$

6. Непрерывное равномерное распределение ($\xi \sim \mathcal{U}[a,b]$).

$$\xi\colon \Omega \to [a,b],$$
$$p_\xi(t) = \frac{1}{b-a}\mathbb{1}_{[a,b]}(t).$$

7. Нормальное распределение ($\xi \sim \mathcal{N}(a,\sigma^2)$).

$$a \in \mathbb{R},\ \sigma > 0.$$
$$p_\xi(t) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(t-a)^2}{2\sigma^2}}.$$

   Для $\xi \sim \mathcal{N}(0,1)$ это называется стандартным нормальным распределением. Нетрудно заметить, что функция распределения для такой величины равна $\Phi(x)$.

8. Экспоненциальное распределение ($\xi \sim \mathrm{Exp}(\lambda)$).

$$\lambda > 0,\ \xi\colon \Omega \to [0,+\infty),$$
$$p_\xi(t) = \lambda e^{-\lambda t}\mathbb{1}_{[0,+\infty)}(t).$$

§14. Совместные распределения. Совместное распределение независимых случайных величин

Определение. Пусть A — борелевское из R^n, $\vec{\xi} = (\xi_1, \dots, \xi_n)\colon \Omega \to \mathbb{R}^n$. Тогда *совместным (многомерным) распределением* этих величин называется

$$P_{\vec{\xi}}(A) = P(\vec{\xi} \in A).$$

Как и в случае с одномерным, многомерное распределение определеяется значением на ячейках:

$$P_{\vec{\xi}}(a,b] = P(\vec{\xi} \in (a,b]) = P(a_1 < \xi_1 \le b_1, \dots, a_n < \xi_n \le b_n).$$

Отметим, что $P_{\vec{\xi}}$ определяет P_{ξ_k}, но не наоборот:

$$A \subset \mathbb{R};\ P_{\xi_1}(A) = P(\xi_1 \in A) = P(\vec{\xi} \in A \times \mathbb{R}^{n-1}) = P_{\vec{\xi}}(A \times \mathbb{R}^{n-1}).$$

В другую сторону: рассмотрим величины ξ_1 и ξ_2 такие, что $P(\xi_i = 0) = P(\xi_i = 1) = \frac{1}{2}$. Если $\xi_1 = \xi_2$, то у нас 2 события с вероятностью $\frac{1}{2}$ — $(0,0)$ и $(1,1)$. Если же они независимы (например, подбрасывание монеток), то 4 события с вероятностью $\frac{1}{4}$. То есть, получили разные совместные распределения.

Определение. Случайные величины $\xi_1, \ldots, \xi_n$ *независимы*, если для любых множеств $A_1, \ldots, A_n \subset \mathbb{R}$ случайные события $\{\xi_1 \in A_1\}, \ldots, \{\xi_n \in A_n\}$ независимы.

Замечание. Если случайные величины $\xi_1, \ldots, \xi_n$ независимы, то

$$P(\xi_1 \in A_1, \ldots, \xi_n \in A_n) = P(\xi_1 \in A_1) \cdot \ldots \cdot P(\xi_n \in A_n),$$
$$P_{\vec{\xi}}(A_1 \times \ldots \times A_n) = P_{\xi_1}(A_1) \cdot \ldots \cdot P_{\xi_n}(A_n).$$

Теорема 14.1. Независимость величин $\xi_1, \ldots, \xi_n$ равносильно тому, что $P_{\vec{\xi}}$ — произведение мер $P_{\xi_1}, \ldots, P_{\xi_n}$.

Доказательство. Достаточно доказать равенство на ячейках:

$$P_{\vec{\xi}}(a, b] = P_{\xi_1}((a_1, b_1]) \cdot \ldots \cdot P_{\xi_n}((a_n, b_n]).$$

Из замечания выше видно, что оно есть. ∎

§15. Совместная функция распределения и совместная плотность. Функции распределения и плотности для независимых случайных величин

Определение. *Совместной функцией распределения* вектора величин $\vec{\xi} \colon \Omega \to \mathbb{R}^n$ называется

$$F_{\vec{\xi}}(\vec{x}) = P(\xi_1 \leq x_1, \ldots, \xi_n \leq x_n).$$

Определение. *Совместная плотность распределения* вектора величин $\vec{\xi}$ — это такая функция $p_{\vec{\xi}}(\vec{t})$, измеримая в $\mathbb{R}^n$ (если она существует), что

$$P_{\vec{\xi}}(A) = \int\limits_A p_{\vec{\xi}}(\vec{t})\, \mathrm{d}\lambda_n(\vec{t}).$$

Как и в случае с одномерной плотностью, нетрудно понять, что

$$F_{\vec{\xi}}(\vec{x}) = \int\limits_{-\infty}^{x_1} \ldots \int\limits_{-\infty}^{x_n} p_{\vec{\xi}}(\vec{t})\, \mathrm{d}t_n \ldots \mathrm{d}t_1.$$

Следствие 15.1. 1. $\xi_1, \ldots, \xi_n$ — независимые $\iff F_{\vec{\xi}}(\vec{x}) = F_{\xi_1}(x_1) \ldots F_{\xi_n}(x_n)$.

2. Пусть $\xi_1, \ldots, \xi_n$ — абсолютно непрерывные случайные величины. Тогда $\xi_1, \ldots, \xi_n$ независимы тогда и только тогда, когда $p_{\vec{\xi}}(\vec{t}) = p_{\xi_1}(t_1) \cdot \ldots \cdot p_{\xi_n}(t_n)$.

Доказательство.

2. $\Longleftarrow$. По определению

$$P_{\vec{\xi}}(A_1 \times \ldots \times A_n) = \int\limits_{A_1 \times \ldots \times A_n} p_{\vec{\xi}}(\vec{t}) \,\mathrm{d}\lambda_n(t) = \int\limits_{A_1 \times \ldots \times A_n} p_{\xi_1}(t_1) \,\mathrm{d}t_1 \cdot \ldots \cdot p_{\xi_n}(t_n) \,\mathrm{d}t_n.$$

По теореме Фубини-Тонелли это равно произведению интегралов:

$$\begin{aligned}\int\limits_{A_1 \times \ldots \times A_n} p_{\xi_1}(t_1) \,\mathrm{d}t_1 \cdot \ldots \cdot p_{\xi_n}(t_n) \,\mathrm{d}t_n &= \int\limits_{A_1} p_{\xi_1}(t_1) \,\mathrm{d}t_1 \cdot \ldots \cdot \int\limits_{A_n} p_{\xi_n}(t_n) \,\mathrm{d}t_n \\ &= P_{\xi_1}(A_1) \cdot \ldots \cdot P_{\xi_n}(A_n).\end{aligned}$$

$\Longrightarrow$. Проверим, что $p_{\xi_1}(t_1) \cdot \ldots \cdot p_{\xi_n}(t_n)$ — совместная плотность. Для этого докажем равенство на ячейках, то есть проверим равенство

$$P_{\vec{\xi}}(a,b] = \int\limits_{(a,b]} p_{\xi_1}(t_1) \cdot \ldots \cdot p_{\xi_n}(t_n) \,\mathrm{d}\lambda_n.$$

С одной стороны,

$$P_{\vec{\xi}}(a,b] = \prod_{k=1}^{n} P_{\xi_k}(a_k, b_k].$$

С другой, по теореме Тонелли,

$$\int\limits_{(a,b]} p_{\xi_1}(t_1) \cdot \ldots \cdot p_{\xi_n}(t_n) \,\mathrm{d}\lambda_n = \prod_{k=1}^{n} \int\limits_{(a_k,b_k]} p_{\xi_k}(t_k) \,\mathrm{d}t_k.$$

Осталось показать, что $\int_{(a_k,b_k]} p_{\xi_k}(t_k) \,\mathrm{d}t_k = P_{\xi_k}(a_k, b_k]$, ну а это является определением плотности. ∎

§16. Свертки мер. Свертки мер, имеющих плотность

В этом паграфе μ и ν — конечные меры на борелевских подмножествах $\mathbb{R}$.

Определение. *Сверткой мер* μ и ν называется

$$\mu * \nu(A) = \int\limits_{\mathbb{R}} \mu(A - x) \,\mathrm{d}\nu(x).$$

Свойства 16.1 (свёртки мер).

1. $\mu * \nu(A) = \int_{\mathbb{R}^2} \mathbb{1}_A(x+y)\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y)$;
2. $\mu * \nu = \nu * \mu$;
3. $(\mu_1 * \mu_2) * \mu_3(A) = \int_{\mathbb{R}^3} \mathbb{1}(x_1 + x_2 + x_3)\,\mathrm{d}\mu_1(x_1)\,\mathrm{d}\mu_2(x_2)\,\mathrm{d}\mu_3(x_3) = \mu_1 * (\mu_2 * \mu_3)(A)$;
4. $\mu_1 * \mu_2 * \ldots * \mu_n(A) = \int_{\mathbb{R}^n} \mathbb{1}(x_1 + x_2 + \ldots + x_n)\,\mathrm{d}\mu_1(x_1)\,\mathrm{d}\mu_2(x_2)\ldots\,\mathrm{d}\mu_n(x_n)$;
5. $(c\mu) * \nu = c \cdot \mu * \nu$;
6. $(\mu_1 + \mu_2) * \nu = \mu_1 * \nu + \mu_2 * \nu$;
7. Пусть δ_a — мера, такая что $\delta_a(\{a\}) = 1$ и $\delta_a(\{\mathbb{R} \setminus \{a\}\}) = 0$. Тогда $\mu * \delta_0 = \mu$ (то есть δ_0 — единица с точки зрения свертки).

Доказательство.

1. По определению свёртки

$$\begin{aligned}\mu * \nu(A) &= \int\limits_{\mathbb{R}} \mu(A - x)\,\mathrm{d}\nu(x)\\ &= \int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}} \mathbb{1}_{A-y}(x)\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y)\\ &= \int\limits_{\mathbb{R}^2} \mathbb{1}_A(x+y)\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y).\end{aligned}$$

Пункты 2-6 очевидно следуют из определения и пункта 1.

7. По определению свёртки

$$\mu * \delta_0(A) = \int\limits_{\mathbb{R}} \mu(A - x) d\delta_0(x) = \mu(A - 0) = \mu A.$$

■

Теорема 16.2. Пусть p_μ и p_ν — плотности мер μ и ν относительно меры Лебега λ. Тогда $\mu * \nu$ имеет плотность

$$p(t) = \int\limits_{\mathbb{R}} p_\mu(t - x) p_\nu(x)\,\mathrm{d}x.$$

Это называется *свёрткой функцией*.

Доказательство. Надо проверить, что

$$\mu * \nu(A) = \int\limits_A p(t)\,\mathrm{d}t.$$

По определению $p(t)$

$$\begin{aligned}\int\limits_A p(t)\,\mathrm{d}t &= \int\limits_A \int\limits_{\mathbb{R}} p_\mu(t-x)p_\nu(x)\,\mathrm{d}x\,\mathrm{d}t \\ &= \int\limits_{\mathbb{R}^2} \mathbb{1}_A(t)p_\mu(t-x)p_\nu(x)\,\mathrm{d}x\,\mathrm{d}t \\ &= \int\limits_{\mathbb{R}^2} \mathbb{1}_A(x+y)p_\mu(y)p_\nu(x)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int\limits_{\mathbb{R}^2} \mathbb{1}_A(x+y)\,\mathrm{d}\mu(y)\,\mathrm{d}\nu(x) = \mu * \nu(A).\end{aligned}$$

■

§17. Распределение суммы независимых случайных величин. Примеры

Теорема 17.1. Пусть ξ и η независимые случайные величины. Тогда $P_{\xi+\eta} = P_\xi * P_\eta$.

Доказательство. По определению распределения

$$\begin{aligned}P_{\xi+\eta}(A) &= P(\xi+\eta \in A) = P((\xi,\eta)\in B) = P_{(\xi,\eta)}(B) \\ &= \int\limits_{\mathbb{R}^2} \mathbb{1}_B(x,y)\,\mathrm{d}P_{(\xi,\eta)}(x,y) = \int\limits_{\mathbb{R}^2} \mathbb{1}_B(x,y)\,\mathrm{d}P_\xi(x)\,\mathrm{d}P_\eta(y) \\ &= \int\limits_{\mathbb{R}^2} \mathbb{1}_A(x+y)\,\mathrm{d}P_\xi(x)\,\mathrm{d}P_\eta(y) = P_\xi * P_\eta(A).\end{aligned}$$

■

Примеры 17.1.

1. Свертка с дискретным распределением. Дискретное распределение можно описать как $\nu = \sum p_x\delta_x$ (вес на нагрузку в точке).

$$\mu * \nu = \sum p_x \mu * \delta_x;$$

$$\mu * \delta_a(A) = \int\limits_{\mathbb{R}} \mu(A-x)d\delta_a(x) = \mu(A-a).$$

$$\mu * \nu(A) = \sum p_x \mu * \delta_x(A) = \sum p_x \mu(A-x).$$

2. Если меры μ и ν с нагрузками в $\mathbb{N}\cup\{0\}$, то

$$\nu = \sum q_n\delta_n,\ \mu = \sum p_n\delta_n;$$

$$\mu * \nu(A) = \sum_{n=0}^{\infty} q_n \mu(A-n), \quad A = \{k\},\ k \in \mathbb{Z};$$

$$\mu * \nu(\{k\}) = \sum_{n=0}^{k} q_n p_{k-n}.$$

3. Пусть величины $\xi_1 \sim \mathrm{Pois}(\lambda_1)$ и $\xi_2 \sim \mathrm{Pois}(\lambda_2)$ независимые. Тогда для ξ_1 веса равны $\frac{e^{-\lambda_1}\lambda_1^n}{n!}$, для ξ_2 веса равны $\frac{e^{-\lambda_2}\lambda_2^n}{n!}$.

   Для $\xi_1 + \xi_2$ веса будут равны

$$\sum_{k=0}^{n} \frac{e^{-\lambda_1}\lambda_1^k}{k!} \frac{e^{-\lambda_2}\lambda_2^{n-k}}{(n-k)!} = \frac{e^{-\lambda_1-\lambda_2}}{n!} \sum_{k=0}^{n} \binom{n}{k} \lambda_1^k \lambda_2^{n-k} = \frac{e^{-(\lambda_1+\lambda_2)}}{n!}(\lambda_1+\lambda_2)^n.$$

   Итого получили, что $\xi_1 + \xi_2 \sim \mathrm{Pois}(\lambda_1 + \lambda_2)$.

§18. Математическое ожидание. Свойства (до математического ожидания произведения включительно)

Определение. Пусть $\xi\colon \Omega \to \mathbb{R}$ — случайная величина, являющаяся суммируемой функцией. Тогда её *математическим ожиданием* называется. $\mathbb{E}\xi = \int\limits_{\Omega} \xi \, \mathrm{d}P$.

Свойства 18.1 (матожидания).

1. Линейность : $\mathbb{E}(a\xi + b\eta) = a\mathbb{E}_\xi + b\mathbb{E}_\eta$;
2. Если $\xi \geq 0$ с вероятностью 1, то $\mathbb{E}\xi \geq 0$;
3. Если $\xi \geq \eta$ с вероятностью 1, то $\mathbb{E}\xi \geq \mathbb{E}\eta$;
4. $\mathbb{E}\xi = \int_{\mathbb{R}} x \, \mathrm{d}P_\xi(x)$;
5. Если $f\colon \mathbb{R} \to \mathbb{R}$ такая, что прообразы борелевских множеств — борелевские, то

$$\mathbb{E}f(\xi) = \int\limits_{\mathbb{R}} f(x) \, \mathrm{d}P_\xi(x).$$

6. Если $f : \mathbb{R}^n \to \mathbb{R}$ такая, что прообразы борелевских — борелевские, то

$$\mathbb{E}f(\xi_1, \ldots, \xi_n) = \int\limits_{\mathbb{R}^n} f(x_1, \ldots, x_n) \, \mathrm{d}P_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n).$$

7. Если ξ и η независимы, то

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

Доказательство. Пункты 1-4 очевидно доказываются из определения(по свойствам интеграла).

5. Частный случай 6-го пункта.

6. Воспользуемся стандартной схемой доказательства из теории меры.

   Докажем для простых. Пусть $f = \mathbb{1}_A$. Тогда

$$\begin{aligned}\mathbb{E}f(\xi_1,\dots,\xi_n) &= \int\limits_{\Omega} \mathbb{1}_A(\xi_1(\omega),\dots,\xi_n(\omega))\,\mathrm{d}P(\omega)\\ &= P((\xi_1,\dots,\xi_n)\in A) = P_{\xi_1,\dots,\xi_n}(A)\\ &= \int\limits_{\mathbb{R}} \mathbb{1}_A\,\mathrm{d}P_{\xi_1,\dots,\xi_n}.\end{aligned}$$

   По линейности верно для простых функций. Приблизим произвольную функцию f простыми:

$$\int\limits_{\Omega} f_k(\xi_1,\dots,\xi_n)\,\mathrm{d}P = \int\limits_{\mathbb{R}^n} f_k(x)\,\mathrm{d}P_{\xi_1,\dots,\xi_n}(x), \quad \text{где} \quad f_k \text{ — простые,}$$

   и перейдём к пределу по теореме Беппо Леви.

7. Воспользуемся предыдущим пунктом для $f(x,y) = xy$:

$$\begin{aligned}\mathbb{E}(\xi\eta) &= \int\limits_{\mathbb{R}^2} xy\,\mathrm{d}P_{\xi,\eta}(x,y) = [\text{поскольку они независимы}]\\ &= \int\limits_{\mathbb{R}^2} xy\,\mathrm{d}P_{\xi}(x)\,\mathrm{d}P_{\eta}(y) = \int\limits_{\mathbb{R}} x\,\mathrm{d}P_{\xi}(x)\int\limits_{\mathbb{R}} y\,\mathrm{d}P_{\eta}(y) = \mathbb{E}\xi\mathbb{E}\eta.\end{aligned}$$

■

Если ξ и η не независимы, то может оказаться $\mathbb{E}(\xi\eta) \neq \mathbb{E}\xi\mathbb{E}\eta$. Пример : пусть ξ принимает значение из $\{-1,1\}$ равновероятно, $\eta = \xi$. Тогда $\mathbb{E}(\xi\eta) = \mathbb{E}\xi^2 = 1$, но $\mathbb{E}\xi = \mathbb{E}\eta = 0$.

§19. Математическое ожидание. Свойства (неравенства Гёльдера, Ляпунова и Маркова). Медиана. Примеры

Свойства 19.1 (матожидания).

1. Если $\xi \geq 0$, то

$$\mathbb{E}\xi = \int\limits_0^{+\infty} P(\xi \geq t)\,\mathrm{d}t.$$

2. Неравенства Гёльдера для матожидания. Если $p, q > 1$ и $\frac{1}{p} + \frac{1}{q} = 1$, то
$$\mathbb{E}|\xi\eta| \le (\mathbb{E}|\xi|^p)^{\frac{1}{p}} (\mathbb{E}|\eta|^q)^{\frac{1}{q}}.$$
3. Неравенство Ляпунова. Если $0 < r < s$, то
$$(\mathbb{E}|\xi|^r)^{\frac{1}{r}} \le (\mathbb{E}|\xi|^s)^{\frac{1}{s}}.$$
4. Неравенство Маркова. Пусть $\xi \ge 0$. Тогда
$$P(\xi \ge t) \le \frac{\mathbb{E}\xi^p}{t^p}, \quad \text{где} \quad p, t > 0.$$

Доказательство. Пункты 1-4 очевидно доказываются из определения(по свойствам интеграла).

8. В теории меры была такая теорема: Если $(X, \mathcal{A}, \mu)$ — пространство с σ-конечной мерой и $f \ge 0$ измеримая, то
$$\int\limits_X f \, \mathrm{d}\mu = \int\limits_0^\infty \mu X\{f \ge t\} \, \mathrm{d}t.$$
9. Прямое следствие из неравенства Гёльдера(для интегралов).
10. $$\mathbb{E}|\xi|^r = \mathbb{E}|\xi|^r \cdot 1 \le (\mathbb{E}(|\xi|^r)^{\frac{s}{r}})^{\frac{r}{s}} (\mathbb{E}1^q)^{\frac{1}{q}} = (\mathbb{E}|\xi|^s)^{\frac{r}{s}}.$$
11. Прямое следствие из неравенства Чебышёва(из теории меры). ■

§20. Дисперсия. Свойства дисперсии. Неравенство Чебышёва. Математическое ожидание и дисперсия для равномерного и нормального распределений

Определение. *Дисперсией* случайной величины ξ называется
$$\mathbb{D}\xi = \mathbb{E}(\xi - \mathbb{E}\xi)^2.$$

Свойства 20.1 (дисперсии).

1. $\mathbb{D}\xi \ge 0$ и если $\mathbb{D}\xi = 0$, то $P(\xi = c) = 1$, где c — некоторая константа.
2. $\mathbb{D}\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2$.
3. $\mathbb{D}(c \cdot \xi) = c^2\mathbb{D}\xi$ (в частности, $\mathbb{D}\xi = \mathbb{D}(-\xi)$).

4. Если ξ и η независимы, то $\mathbb{D}(\xi+\eta)=\mathbb{D}\xi+\mathbb{D}\eta$.
5. $\mathbb{E}|\xi-\mathbb{E}\xi|\leq\sqrt{\mathbb{D}\xi}$.
6. Неравенство Чебышёва при $t>0$.

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

Доказательство. 1. Очевидно из определения;

2. По определению

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

3. Очевидно из определения;
4. Воспользуемся свойством матожидания для произведения независимых величин:

$$\begin{aligned}\mathbb{D}(\xi+\eta)&=\mathbb{E}(\xi+\eta)^2-(\mathbb{E}(\xi+\eta))^2\\&=\mathbb{E}\xi^2+2\mathbb{E}\xi\eta+\mathbb{E}\eta^2-(\mathbb{E}\xi+\mathbb{E}\eta)^2\\&=\mathbb{E}\xi^2+\mathbb{E}\eta^2-(\mathbb{E}\xi)^2-(\mathbb{E}\eta)^2.\end{aligned}$$

5. Пусть $\eta=\xi-\mathbb{E}\xi$, тогда по неравенству Ляпунова

$$\mathbb{E}|\eta|\leq(\mathbb{E}|\eta|^2)^{\frac{1}{2}}.$$

6. Из предыдущего свойства

$$P(\eta\geq t)\leq\frac{\mathbb{E}\eta^2}{t^2}.$$

■

Примеры 20.1. 1. Пусть $\xi\sim\mathcal{U}[0,1]$. Тогда

$$\mathbb{E}\xi=\int\limits_{\mathbb{R}}x\,\mathrm{d}P_\xi=\int\limits_0^1 x\,\mathrm{d}x=\left.\frac{x^2}{2}\right|_0^1=\frac{1}{2}.$$

$$\mathbb{E}\xi^2=\int\limits_{\mathbb{R}}x^2\,\mathrm{d}P_\xi(x)=\int\limits_0^1 x^2\,\mathrm{d}x=\left.\frac{x^3}{3}\right|_0^1=\frac{1}{3}.$$

$$\mathbb{D}\xi=\mathbb{E}\xi^2-(\mathbb{E}\xi)^2=\frac{1}{12}.$$

2. Пусть $\xi\sim\mathcal{U}[a,b]$. Нетрудно заметить, что $\xi=(b-a)\eta+a$, где $\eta\sim\mathcal{U}[0,1]$. Тогда

$$\mathbb{E}\xi=\mathbb{E}((b-a)\eta+a)=(b-a)\mathbb{E}\eta+a=\frac{a+b}{2}.$$

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

3. Пусть $\xi \sim \mathcal{N}(0,1)$. Тогда

$$\mathbb{E}\xi = \int\limits_{\mathbb{R}} x \, \mathrm{d}P_\xi(x) = \int\limits_{\mathbb{R}} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, \mathrm{d}x = 0,$$

поскольку функция нечётная.

$$\mathbb{D}\xi = \mathbb{E}\xi^2 = \frac{1}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} x^2 e^{-\frac{x^2}{2}} \, \mathrm{d}x = -\frac{1}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} x \, \mathrm{d}\left(e^{-\frac{x^2}{2}}\right) = \frac{1}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} e^{-\frac{x^2}{2}} \, \mathrm{d}x = 1.$$

4. Пусть $\xi \sim \mathcal{N}(a,\sigma^2)$. Поймём, что $\xi = \sigma\eta + a$, где $\eta \sim \mathcal{N}(0,1)$. Пусть $\xi' = \sigma\eta$. Тогда

$$\begin{aligned} F_{\xi'}(x) &= \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} \, \mathrm{d}t = [t = \sigma s] \\ &= \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{x}{\sigma}} e^{-\frac{s^2}{2}} \, \mathrm{d}s = F_\eta\left(\frac{x}{\sigma}\right), \end{aligned}$$

то есть

$$\xi' = \sigma\eta \sim \mathcal{N}(0,\sigma^2).$$

Аналогично доказывается вторая часть, и тогда

$$\mathbb{E}\xi = \mathbb{E}(\sigma\eta + a) = \sigma\mathbb{E}\eta + a = a.$$
$$\mathbb{D}\xi = \mathbb{D}(\sigma\eta + a) = \sigma^2.$$

§21. Ковариация. Связь с независимостью. Коэффициент корреляции. Моменты случайной величины

Определение. *Ковариацией* случайных величин ξ и η называется

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

Свойства 21.1 (ковариации).

1. $\operatorname{cov}(\xi,\xi) = \mathbb{D}\xi$.
2. $\operatorname{cov}(\xi,\eta) = \mathbb{E}(\xi\eta) - \mathbb{E}\xi\mathbb{E}\eta$.
3. $\mathbb{D}(\xi+\eta) = \mathbb{D}\xi + \mathbb{D}\eta + 2\operatorname{cov}(\xi,\eta)$.
4. $\mathbb{D}(\sum\limits_{k=1}^{n} \xi_k) = \sum\limits_{k=1}^{n} \mathbb{D}\xi_k + \sum\limits_{i\neq k} \operatorname{cov}(\xi_i,\xi_k) = \sum\limits_{k=1}^{n} \mathbb{D}\xi_k + 2\sum\limits_{i<k} \operatorname{cov}(\xi_i,\xi_k)$.

5. Если ξ и η независимые, то $\operatorname{cov}(\xi, \eta) = 0$.

6. $\operatorname{cov}(\xi, \eta) = \operatorname{cov}(\eta, \xi)$.

7. $\operatorname{cov}(c\xi, \eta) = c \cdot \operatorname{cov}(\xi, \eta)$.

8. $\operatorname{cov}(\xi_1 + \xi_2, \eta) = \operatorname{cov}(\xi_1, \eta) + \operatorname{cov}(\xi_2, \eta)$.

Сделаем несколько важных замечаний.

1. Дисперсия и ковариация могут не существовать (например, если не определено матожидание). Для существования ковариации надо, чтобы $\mathbb{E}\xi^2 < +\infty$ и $\mathbb{E}\eta^2 < +\infty$ (а дисперсия — частный случай ковариации).

2. Из $\operatorname{cov}(\xi, \eta) = 0$ не следует независимость этих величин.

   Например, пусть $\Omega = \{0, \frac{\pi}{2}, \pi\}$, каждое значение равновероятно. Возьмём $\xi = \sin\omega$, $\eta = \cos\omega$. Тогда

$$\xi\eta \equiv 0,\ \mathbb{E}(\xi\eta) = 0,\ \mathbb{E}\eta = 0;$$
$$\operatorname{cov}(\xi, \eta) = \mathbb{E}(\xi\eta) - \mathbb{E}\xi\mathbb{E}\eta = 0.$$

   Но они не являются независимыми:

$$P(\xi = 1, \eta = 1) = 0 \neq P(\xi = 1)P(\xi = 1) = \frac{1}{9}.$$

Определение. *Коэффициентом корреляции* случайных величин ξ и η называется

$$\rho(\xi, \eta) = \frac{\operatorname{cov}(\xi, \eta)}{\sqrt{\mathbb{D}\xi}\sqrt{\mathbb{D}\eta}}.$$

Очевидно, это значение лежит на отрезке $[-1, 1]$.

Определение. Случайные величины, для которых $\operatorname{cov}(\xi, \eta) = 0$ называются *некоррелированными*.

Определение. $\mathbb{E}\xi^k = \int\limits_{\mathbb{R}} x^k \, \mathrm{d}P_\xi(x)$ — *k-й момент* случайной величины. $\mathbb{E}|\xi|^k$— *k-й абсолютный момент*. $\mathbb{E}((\xi - \mathbb{E}\xi)^k)$ — *k-й центральный момент*.

Легко заметить, что дисперсия — 2-й центральный момент.

§22. Выбор двудольного подграфа с большим количеством ребер

Пример 22.1. Пусть $G(V, \mathbb{E})$ — граф, в котором $|V| = n$, $|\mathbb{E}| = \frac{nd}{2}$, где $d \geq 1$. Тогда в G можно выбрать $\frac{n}{2d}$ попарно несоединенных друг с другом вершин.

Доказательство. Рассмотрим случайное подмножество вершин S, причём вероятность вхождения вершины в него равно p. Рассмотрим подграф на вершинах S. Если $xy \in \mathbb{E}$, то обозначим за $\xi_{xy} = 1$, если $x, y \in S$, и 0 иначе.

Пусть ξ — количество ребер в подграфе на вершинах из S. Тогда $\xi = \sum\limits_{xy \in \mathbb{E}} \xi_{xy}$. Обозначим $\eta = \#S$, очевидно $\mathbb{E}\eta = np$. Пусть $\eta_x = 1$ если $x \in S$ и 0 иначе, тогда

$$\eta = \sum_{x \in V} \eta_x,\ \mathbb{E}\eta = \sum_{x \in V} \mathbb{E}\eta_x = \sum p = np;$$

$$\mathbb{E}\xi = \sum_{xy \in \mathbb{E}} \mathbb{E}\xi_{xy} = \sum_{xy \in \mathbb{E}} P(x, y \in S) = \sum p^2 = \frac{p^2 nd}{2};$$

$$\mathbb{E}(\eta - \xi) = np - \frac{p^2 nd}{2}.$$

Хотим максимизировать это значение, для этого возьмем $p = \frac{1}{d}$ и получим $\frac{n}{2d}$. ■

§23. Теорема Харди–Рамануджана о количестве различных простых делителей числа

Теорема 23.1 (Харди-Рамануджана). Пусть $\nu(k)$ — количество простых в разложении k, тогда если $\omega(n) \to +\infty$, то

$$P(|\nu(k) - \ln\ln n| > \omega(n)\sqrt{\ln\ln n}) \to 0.$$

Доказательство. Обозначим $t = \omega(n)\sqrt{\ln\ln n}$. Рассмотрим

$$\xi_p(k) = \begin{cases} 1, & k \,\vdots\, p \\ 0, & \text{иначе} \end{cases},\ \text{где } p \text{ — простое.}$$

Возьмём $M = n^{\frac{1}{10}}$ и обозначим

$$\xi = \sum_{p \le M} \xi_p,\ 0 \le \nu(k) - \xi(k) \le 10;$$

$$\mathbb{E}_{\xi_p} = \frac{[\frac{n}{p}]}{n} = \frac{1}{p} + O\left(\frac{1}{n}\right);$$

$$\mathbb{E}_{\xi} = \sum \mathbb{E}_{\xi_p} = \sum_{p \le M} \frac{1}{p} + O\left(\frac{M}{n}\right) = \sum_{p \le M} \frac{1}{p} + O(1) = \ln\ln M + O(1) = \ln\ln n + O(1).$$

$$\mathbb{D}\xi = \sum \mathbb{D}\xi_p + \sum \operatorname{cov}(\xi_p, \xi_q);$$

$$\mathbb{D}\xi_p = \mathbb{E}\xi_p^2 - (\mathbb{E}\xi_p)^2 = \mathbb{E}\xi_p - (\mathbb{E}\xi_p)^2 = \frac{1}{p} - \frac{1}{p^2} + O(\frac{1}{n});$$

$$\operatorname{cov}(\xi_p, \xi_q) = \mathbb{E}\xi_p\xi_q - \mathbb{E}\xi_p\mathbb{E}\xi_q = P(k \,\vdots\, pq) - \frac{[\frac{n}{p}]}{n}\frac{[\frac{n}{q}]}{n} = \frac{[\frac{n}{pq}]}{n} = \frac{[\frac{n}{p}] \cdot [\frac{n}{q}]}{n}.$$

С одной стороны,

$$\frac{[\frac{n}{p}]\cdot[\frac{n}{q}]}{n} \geq \frac{\frac{n}{pq}-1}{n} - \frac{\frac{n}{p}\frac{n}{q}}{n^2} = -\frac{1}{n}.$$

С другой

$$\frac{[\frac{n}{p}]\cdot[\frac{n}{q}]}{n} \leq \frac{\frac{n}{pq}}{n} - \frac{(\frac{n}{p}-1)(\frac{n}{q}-1)}{n} = \frac{1}{n}\left(\frac{1}{p}+\frac{1}{q}\right).$$

То есть

$$2\sum \operatorname{cov} \geq -\frac{M^2}{n} = O(1) \quad \text{и} \quad 2\sum \operatorname{cov} \leq \frac{1}{n}\sum\left(\frac{1}{p}+\frac{1}{q}\right) \leq 2\frac{M}{n}\sum\frac{1}{p} = O(1).$$

Таким образом,

$$\mathbb{D}\xi = \sum\left(\frac{1}{p}-\frac{1}{p^2}\right) + O(1) = \sum\frac{1}{p} + O(1) = \ln\ln n + O(1).$$

Теперь применим неравенство Чебышёва:

$$P(|\xi - E\xi| \geq t) \leq \frac{\mathbb{D}\xi}{t^2} = \frac{\ln\ln n + O(1)}{\omega^2(n)\ln\ln n} \to 0.$$

■

§24. Независимость функций от независимых случайных величин

Теорема 24.1. Пусть ξ_i — независимые случайные величины, функции $f_j : \mathbb{R}^{n_j} \to \mathbb{R}$ измеримы относительно борелевской σ-алгебры. Тогда величины $f_1(\xi_1, \ldots, \xi_{n_1})$, $f_2(\xi_{n_1+1}, \ldots, \xi_{n_2})$, ... независимы.

Доказательство. Покажем, что $f(\xi_1, \ldots, \xi_n)$ и $g(\eta_1, \ldots, \eta_m)$ независимы, если величины $\xi_1, \ldots, \xi_n$, $\eta_1, \ldots, \eta_m$ независимы; общий случай делается аналогично. Надо проверить независимость событий $\{f(\xi_1, \ldots, \xi_n) \in A\}$ и $\{g(\eta_1, \ldots, \eta_m) \in B\}$.

$$\{f(\xi_1, \ldots, \xi_n) \in A\} = \{(\xi_1, \ldots, \xi_n) \in \overline{A}\}, \quad \text{где} \quad \overline{A} = f^{-1}(A);$$
$$\{g(\eta_1, \ldots, \eta_m) \in B\} = \{(\eta_1, \ldots, \eta_m) \in \overline{B}\}, \quad \text{где} \quad \overline{B} = f^{-1}(B);$$

Надо доказать, что

$$P((\xi_1, \ldots, \xi_n) \in \overline{A}) \cdot P((\eta_1, \ldots, \eta_m) \in \overline{B}) = P((\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_m) \in \overline{A} \times \overline{B}).$$

То есть надо доказать равенство мер, заданых на борелевских множествах, значит достаточно проверить это равенство на ячейках:

$$P((\xi_1, \ldots, \xi_n) \in (a, b]) = P(\xi_1 \in (a_1, b_1]) \cdot \ldots \cdot P(\xi_n \in (a_n, b_n]).$$

Аналогично расписываем две другие вероятности и получаем равенство. ■

§25. Различные виды сходимости последовательности случайных величин. Связь между схо- димостями

Определение.

- ξ_n сходится к ξ *почти наверное* (*с вероятностью 1*), если

$$P(\{\omega \in \Omega \mid \lim \xi_k(\omega) = \xi(\omega)\}) = 1$$

  (то же самое, что и сходимость почти везде).

- ξ_n сходится к ξ *в среднем порядка* $r > 0$, если

$$\mathbb{E}|\xi_n - \xi|^r \to 0.$$

- ξ_n сходится к ξ *по вероятности*, если

$$\forall \varepsilon > 0 \, P(|\xi_n - \xi| > \varepsilon) \to 0$$

  (то же самое, что и сходимость по мере).

- ξ_n сходится к ξ *по распределению*, если F_{ξ_n} сходится F_ξ во всех точках непрерывности F_ξ.

“Иерархия”сходимостей следующая:

- $1 \implies 3$. Из теории меры по теореме Лебега (в обратную сторону неверно, смотри пример в там же);
- $2 \implies 3$. Применим неравенство Маркова:

$$P(|\xi_n - \xi| > \varepsilon) = P(|\xi_n - \xi|^r > \varepsilon^r) \le \frac{\mathbb{E}|\xi_n - \xi|^r}{\varepsilon^r} \to 0.$$

- $1 \not\Rightarrow 2$ (и, следовательно, $3 \not\Rightarrow 2$). Пример: $\Omega = [0, 1]$.

$$\xi_n = n^{\frac{1}{r}} \mathbb{1}_{[0, \frac{1}{n})} \to \xi \equiv 0,$$

  но $\mathbb{E}{\xi_n}^r = 1$.

- $2 \not\Rightarrow 1$ (и, значит, $3 \not\Rightarrow 1$). Например,

$$\xi_{n,k} = \mathbb{1}_{[\frac{k}{n}, \frac{k+1}{n}]};$$
$$\mathbb{E}\xi_{n,k}^r = \frac{1}{n} \to 0,$$

  но сходимости почти везде нет.

- $3 \implies 4$.

Докажем последнее.

Доказательство.

$$\{\xi_n \le x\} \subset \{\xi \le x+\varepsilon\} \cup \{|\xi_n - \xi| \ge \varepsilon\};$$
$$F\xi_n(x) = P(\xi_n \le x) \le P(\xi \le x+\varepsilon) + P(|\xi_n - \xi| \ge \varepsilon) = F_\xi(x+\varepsilon) + P(|\xi_n - \xi| \ge \varepsilon);$$
$$\varlimsup_{n\to+\infty} F_{\xi_n}(x) = F_\xi(x+\varepsilon) + \lim_{n\to+\infty} P(|\xi_n - \xi| \ge \varepsilon) = F_\xi(x+\varepsilon). \tag{$*$}$$
$$\{\xi_n \le x\} \supset \{\xi \le x-\varepsilon\} \cap \{|\xi_n - \xi| < \varepsilon\};$$
$$\{\xi_n > x\} \subset \{\xi > x-\varepsilon\} \cup \{|\xi_n - \xi| \ge \varepsilon\};$$
$$P(\xi_n > x) \le P(\xi > x-\varepsilon) + P(|\xi_n - \xi| \ge \varepsilon);$$
$$1 - F_{\xi_n}(x) \le 1 - F_\xi(x-\varepsilon) + P(|\xi_n - \xi| \ge \varepsilon);$$
$$F_{\xi_n}(x) \ge F_\xi(x-\varepsilon) + P(|\xi_n - \xi| \ge \varepsilon);$$
$$\varliminf_{n\to+\infty} F_{\xi_n}(x) \ge F_\xi(x-\varepsilon). \tag{$**$}$$

По непрерывности имеем, что

$$\forall\delta > 0\, F_\xi(x) - \delta \le F_\xi(x-\varepsilon),\ F_\xi(x+\varepsilon) < F_\xi(x) + \delta.$$

Итого, из непрерывности, $(*)$ и $(**)$ получаем

$$\forall\delta > 0\, F_\xi(x) - \delta \le F_\xi(x-\varepsilon) \le \varliminf F_{\xi_n}(x) \le \varlimsup F_{\xi_n}(x) \le F_\xi(x+\varepsilon) < F_\xi(x) + \delta.$$

Значит, предел существует и верно

$$\lim F_{\xi_n}(x) = F_\xi(x).$$

■

§26. Закон больших чисел. Следствия

Теорема 26.1 (закон больших чисел). Пусть величины $\xi_1, \ldots$ попарно некоррелированы, для всех i верно $\mathbb{D}\xi_n < M$, $S_n = \xi_1 + \ldots + \xi_n$. Тогда

$$P\left(\left|\frac{S_n}{n} - \mathbb{E}\frac{S_n}{n}\right| \ge t\right) \to 0 \quad \text{при} \quad t > 0.$$

Доказательство. Применим неравенство Чебышёва:

$$P\left(\left|\frac{S_n}{n} - \mathbb{E}\frac{S_n}{n}\right| \ge t\right) \le \frac{\mathbb{D}(\frac{S_n}{n})}{t^2} = \frac{\mathbb{D}(S_n)}{n^2t^2} = [\operatorname{cov}(\xi_i, \xi_j) = 0]$$
$$= \frac{\sum \mathbb{D}\xi_n}{n^2t^2} \le \frac{nM}{n^2t^2} = \frac{M}{nt^2} \to 0.$$

■

Следствие 26.2 (закон больших чисел в форме Чебышёва). Пусть $\xi_1, \ldots$ — независимые одинаково распределенные случайные величины с конечной дисперсией и $a = \mathbb{E}\xi_1$, тогда

$$P\left(\left|\frac{S_n}{n} - a\right| \geq t\right) \to 0 \quad \text{при} \quad t > 0,$$

то есть $\frac{S_n}{n}$ сходится к a по вероятности.

Доказательство. Величины независимы, поэтому некоррелированы, дисперсии ограничены и, поскольку матожидание суммы равно сумме матожиданий (которые равны a, поскольку они одинаково распределены), $\mathbb{E}\frac{S_n}{n} = a$; применим теорему. ■

Следствие 26.3 (закон больших чисел для схем Бернулли). Пусть $\xi_1, \ldots$ независимые бернуллиевские случайные величины с вероятностью p. Тогда

$$P\left(\left|\frac{S_n}{n} - p\right| \geq t\right) \to 0 \quad \text{при} \quad t > 0.$$

Доказательство. Для бернуллиевских величин верно

$$\mathbb{E}\xi_1 = P(\xi_1 = 1) = p;$$
$$\mathbb{D}\xi = \mathbb{E}\xi_1^2 - (\mathbb{E}\xi_1)^2 = p - p^2.$$

Показали ограниченность дисперсий, и можно применить теорему. ■

§27. Усиленный закон больших чисел. Следствие. Метод Монте–Карло

Теорема 27.1 (усиленный закон больших чисел). Пусть $\xi_1, \ldots$ — независимые случайные величины, $\mathbb{E}|\xi_k - \mathbb{E}\xi_k|^4 \leq C$. Тогда $\frac{S_n}{n} - \mathbb{E}\frac{S_n}{n} \to 0$ почти наверное.

Доказательство. Пусть $a_n = \mathbb{E}\xi_n$, $\widetilde{\xi_n} = \xi_n - a_n$; очевидно, $\mathbb{E}\widetilde{\xi_n} = 0$, и тогда нам надо доказать, что при $\mathbb{E}\xi_n^4 \leq c$ верно $\frac{\widetilde{\xi_1} + \ldots + \widetilde{\xi_n}}{n} \to 0$ почти наверное.

Далее считаем, что $a_n = 0$, $A_n = \left\{\left|\frac{S_n}{n}\right| > \varepsilon\right\}$.

Если в какой то точке нет стремления к нулю, то это означает, что она лежит в бесконечном числе A_n. Мы раньше обсуждали, как можно описать все такие множества, они описываются как

$$A = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

Таким образом, надо доказать, что $P(A) = 0$. По лемме 10.1(Борелля-Кантелли) достаточно доказать, что $\sum_{n=1}^{\infty} P(A_n) < +\infty$.

$$P(A_n) = P\left(\left|\frac{S_n}{n}\right|^4 > \varepsilon^4\right) \leq \frac{\mathbb{E}(\frac{S_n}{n})^4}{\varepsilon^4} = \frac{\mathbb{E}S_n^4}{n^4\varepsilon^4}. \qquad (*)$$

Докажем, что $\mathbb{E}s_n^4 \le cn^4$.

$$(\xi_1 + \ldots + \xi_n)^4 = \sum \xi_k^4 + C_1 \sum_{i \neq j} \xi_i^2 \xi_j^2 + C_2 \sum \xi_i^2 \xi_j \xi_k + C_3 \sum \xi_i^3 \xi_k + C_4 \sum \xi_i \xi_j \xi_k \xi_l.$$

Величины независимы, а $\mathbb{E}\xi_k = 0$, так что

$$\begin{aligned}\mathbb{E}S_n^4 &= \sum \mathbb{E}\xi_k^4 + C_1 \sum_{i \neq j} \mathbb{E}(\xi_i^2 \xi_j^2) \\ &\le nC + C_1 \sum_{i \neq j} \mathbb{E}(\xi_i^2)\mathbb{E}(\xi_j^2) \le \text{[по неравенству Ляпунова]} \\ &\le nC + C_1 \sum \sqrt{\mathbb{E}\xi_i^4}\sqrt{\mathbb{E}\xi_j^4} \le nC + C_1 n^2 C \le cn^2.\end{aligned}$$

Применяя к ($*$), получаем, что $P(A_n) \le \frac{c}{n^2 \varepsilon^4}$, значит ряд сходится. ■

Следствие 27.2 (усиленный закон больших чисел схемы Бернулли). Пусть p — вероятность успеха в бернуллиевских величинах ξ_i. Тогда $\frac{S_n}{n} \to p$ почти наверное.

Доказательство. $\mathbb{E}(\xi_1 - p)^4 < +\infty$, так как $(\xi_1 - p)^4$ принимает значение p^4 с вероятностью $1 - p$ и $(1 - p)^4$ с вероятностью p. ■

Пример 27.1 (метод Монте-Карло). Есть фигура на плоскости. Хотим оценить её площадь. Для этого возьмём прямогульник, полностью покрывающий эту фигуру, и будем брать случайные точки внутри него. Пусть случайная величина $\xi_i = 1$, если i-ая случайная точка лежит внутри фигуры, и $\xi_i = 0$, если не лежит. Вероятность того, что точка попадёт? равна $p = \frac{\text{площадь фигуры}}{\text{площадь прямоугольника}}$. По следствию для схемы Бернулли, $\frac{S_n}{n} \to p$ почти наверное, то есть, посчитав большое количество точек, можно оценить площадь фигуры.

Теорема 27.3 (усиленный закон больших чисел в форме Колмогорова). Пусть $\xi_1, \ldots$ независимые одинаково распределенные случайные величины, $S_n = \xi_1 + \ldots + \xi_n$. Тогда $\frac{S_n}{n} \to a$ почти наверное равносильно тому, что $a = \mathbb{E}\xi_1$.

§28. Математическое ожидание функции от последовательности сходящихся по вероятности слу- чайных величин. Доказательство Бернштейна тео- ремы Вейерштрасса

Теорема 28.1. Пусть $a \in \mathbb{R}$, $f : \mathbb{R} \to \mathbb{R}$ — ограниченная M, непрерывная в точке a, $\xi_1, \xi_2, \ldots$ сходятся к a по вероятности. Тогда $\mathbb{E}f(\xi_n) \to f(a)$.

Доказательство. Оценим сверху модуль разности:

$$|\mathbb{E}f(\xi_n) - f(a)| = |\mathbb{E}\left(f(\xi_n) - f(a)\right)| \le \mathbb{E}\left|f(\xi_n) - f(a)\right|$$

$$
\begin{aligned}
&= \mathbb{E}\left(|f(\xi_n) - f(a)| \cdot \mathbb{1}_{\{|\xi_n - a| < \epsilon\}}\right) + \mathbb{E}\left(|f(\xi_n) - f(a)| \cdot \mathbb{1}_{\{|\xi_n - a| \geq \epsilon\}}\right) \\
&\leq \sup_{|x-a|<\epsilon} |f(x) - f(a)| + 2M \cdot P\left(|\xi_n - a| \geq \epsilon\right)
\end{aligned}
$$

Устремим ϵ к 0 и ограничим верхний предел сверху(не знаем про существование обычного, тем не менее нам этого достаточно, так как величина всегда не отрицательна):

$$\overline{\lim}\, |\mathbb{E}f(\xi_n) - f(a)| \leq \sup_{|x-a|<\epsilon} |f(x) - f(a)| + 2M \cdot \overline{\lim}\, P(|\xi_n - a| \geq \epsilon) \to 0$$

Первое слагаемое можем сделать сколь угодно маленьким, так как функция непрерывна. Второе слагаемое стремится к нулю, так как есть сходимость ξ_n к a по вероятности. ∎

Теорема 28.2 (Вейерштрасса). Пусть $f \in C[a, b]$. Тогда существует последовательность многочленов $P_n \in \mathbb{R}[x]$ такая, что $P_n \rightrightarrows f$ на $[a, b]$.

Доказательство. Считаем, что $[a, b] = [0, 1]$, так как можем преобразовать аргументы в обе стороны какими-то линейными преобразованиями и от этого ничего не сломается. Рассмотрим схему бернулли с вероятностью успеха p и введем случайную величину $\xi_n = \frac{S_n}{n}$, где S_n — количество выигрышей среди первых n бросков в нашей схеме.

$$\mathbb{E}f(\xi_n) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) P(S_n = k) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \cdot C_n^k p^k (1-p)^{n-k}$$

Получили какой-то многочлен(многочлен Бернштейна) от p n-ой степени. Тогда оценим разницу такого многочлена и $f(p)$ так же как оценивали разность в прошлой теореме:

$$
\begin{aligned}
|\mathbb{E}f(\xi_n) - f(p)| &\leq \sup_{|x-p|\leq\epsilon} |f(x) - f(p)| + 2M \cdot P(|\xi_n - p| \geq \epsilon) \\
&\leq \omega_f(\epsilon) + P\left(\left|\frac{S_n}{n} - p\right| \geq \epsilon\right) \leq \omega_f(\epsilon) \frac{\mathbb{D}\frac{S_n}{n}}{\epsilon^2} \\
&= \omega_f(\epsilon) + \frac{p(1-p)}{n\epsilon^2} \leq \omega_f(\epsilon) + \frac{1}{4n\epsilon^2}
\end{aligned}
$$

где $\omega_f(\epsilon)$ — это модуль непрерывности функции. Обьяснение предпоследнего перехода — вынесли $\frac{1}{n}$ как коэффициент с квадратом, расписали дисперсию суммы как сумму дисперсий, так как броски монетки независимы.

Таким образом получили какую-то оценку выраженную через n и ϵ. Теперь давайте скажем, что $\epsilon = \frac{1}{\sqrt[3]{n}}$ и получим равномерное стремление к нулю. ∎

§29. Производящие функции для целозначных случайных величин. Примеры

Определение. Пусть $\xi\colon \Omega \to \mathbb{N}_0$ — случайная величина. Тогда её *производящей функцией* называется
$$G_\xi(z) = \sum_{n=0}^{\infty} P(\xi = n) z^n.$$

Свойства 29.1 (производящей функции).

1. $G_\xi(z) = \mathbb{E} z^\xi$;
2. $G_\xi(1) = 1$ и ряд сходится в единичном круге.
3. $G'_\xi(1) = \mathbb{E}\xi$;
4. $\mathbb{E}\xi^2 = G''_\xi(1) + G'_\xi(1)$;
5. $\mathbb{D}\xi = G''_\xi(1) + G'_\xi(1) - (G'_\xi(1))^2$;
6. Если ξ и η независимы, то $G_{\xi+\eta}(z) = G_\xi(z) \cdot G_\eta(z)$.

Доказательство.

1. ξ действует в неотрицательные числа, так что
$$\mathbb{E} z^\xi = \int\limits_{\mathbb{R}} z^x \, \mathrm{d} P_\xi(x) = \sum_{n=0}^{\infty} P(\xi = n) z^n.$$
2. Подставим 1 в определение, получим сумму вероятностей всех возможных событий, что равняется единице. Коэффициенты — некоторые вероятности, — неотрицательны, так что в единичном круге есть сходимость.
3. По определению
$$G'_\xi(z) = \sum_{n=1}^{\infty} n P(\xi = n) z^{n-1};$$
$$G'_\xi(z) = \sum_{n=1}^{\infty} n P(\xi = n),$$
что в точности равно $\mathbb{E}\xi$.
4. По определению
$$G''_\xi(z) = \sum_{n=1}^{\infty} n(n-1) P(\xi = n) z^{n-2}$$

Пользуясь предыдущим пунктом, получим

$$G''_\xi(1) + G'_\xi(1) = \sum_{n=1}^{\infty} n^2 P(\xi = n),$$

что в точности равно $\mathbb{E}\xi^2$.

5. По свойствам дисперсии $\mathbb{D}\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2$. Подставим два предыдущих пункта.

6. Рассмотрим $G(z) = G_\xi(z) \cdot G_\eta(z)$. Это свёртка последовательностей. Коэффициент при n-ой степени равен $c_n = P(\xi = 0)P(\eta = n) + P(\xi = 1)P(\eta = n-1) + \ldots + P(\xi = n)P(\eta = 0) = P(\xi + \eta = n)$, а значит $G(z) = G_{\xi+\eta}(z)$. ■

Примеры 29.1. 1. Равномерное распределение. Пусть ξ равновероятно принимает значения из $\{0, 1, \ldots n-1\}$. Тогда

$$G_\xi = \frac{1 + z + \ldots + z^{n-1}}{n} = \frac{1 - z^n}{n(1-z)}.$$

Чтобы было удобнее работать, сделаем замену $z = 1 + t$:

$$G_\xi(1+t) = \frac{(1+t)^n - 1}{nt} = \frac{1}{n}\binom{n}{1} + \frac{1}{n}\binom{n}{2}t + \frac{1}{n}\binom{n}{3}t^2 \ldots.$$

И тогда можем вычислить

$$G'_\xi(1) = \frac{n-1}{2};$$
$$G''_\xi(1) = \frac{(n-1)(n-2)}{3};$$
$$\mathbb{D}\xi = \frac{n^2-1}{12}.$$

2. Задача Галилея. Бросается три кубика, какова вероятность что сумма очков равна 10?. Пусть ξ_i — количество очков на i-м кубике.

$$G_{\xi_i}(z) = \frac{z + z^2 + \ldots + z^6}{6} = \frac{z(1-z^6)}{6(1-z)};$$
$$G_{\xi_1+\xi_2+\xi_3}(z) = (G_\xi(z))^3 = \frac{z^3(1-z^6)^3}{6^3(1-z)^3} = \frac{1}{6^3} z^3 (1 - 3z^6 + 3z^{12} - z^{18}) \sum_{n=0}^{\infty} \binom{n+2}{n} z^n;$$

По определению производящей функции коэффициент при z_{10} — это $P(\xi_1 + \xi_2 + \xi_3 = 10)$, как раз то, что мы ищем.

$$P(\xi_1 + \xi_2 + \xi_3 = 10) = \frac{1}{6^3}\left(\binom{9}{7} - \binom{3}{1}\right) = \frac{1}{8}.$$

§30. Математическое ожидание для комплекснозначных случайных величин. Свойства. Ковариация

Определим комплекснозначные случайные величины.

Определение. $\xi\colon \Omega \to \mathbb{C}$ — *случайная величина*, если $\operatorname{Re}\xi$ и $\operatorname{Im}\xi$ — вещественнозначные случайные величины; $\xi = \operatorname{Re}\xi + i\operatorname{Im}\xi$, $\mathbb{E}\xi = \mathbb{E}(\operatorname{Re}\xi) + i\mathbb{E}(\operatorname{Im}\xi)$.

Свойства 30.1 (комплекснозначной случайной величины).

1. Комплексная линейность;
2. $|\mathbb{E}\xi| \leq \mathbb{E}|\xi|$;

Доказательство.

1. Докажем, что $\mathbb{E}(i\xi) = i\mathbb{E}\xi$. Пусть $\xi = \zeta + i\eta$, где $\zeta, \eta\colon \Omega \to \mathbb{R}$. Тогда
$$\mathbb{E}(i\xi) = \mathbb{E}(i\zeta - \eta) = i\mathbb{E}\zeta - \mathbb{E}\eta = i(\mathbb{E}\zeta + i\mathbb{E}\eta) = i\mathbb{E}\xi.$$
2. Возьмем $c = \frac{\overline{\mathbb{E}\xi}}{|\mathbb{E}\xi|}$, тогда $|c| = 1$, $\mathbb{E}(c\xi) = |\mathbb{E}\xi|$.
$$\mathbb{E}|\xi| = \mathbb{E}|c\xi| \geq \mathbb{E}|\operatorname{Re}(c\xi)| \geq \mathbb{E}(\operatorname{Re}(c\xi)) = \operatorname{Re}\mathbb{E}(c\xi) = |\mathbb{E}\xi|.$$

■

Определение. *Ковариацией комплекснозначных* случайных величин ξ и η называется
$$\operatorname{cov}(\xi, \eta) = \mathbb{E}((\xi - \mathbb{E}\xi)\overline{(\eta - \mathbb{E}\eta)}).$$

Понятно, что для величин, принимающих вещественное значение определение не изменилось.

Также заметим, что сохранилось равенство
$$\mathbb{D}\xi = \operatorname{cov}(\xi, \xi).$$

§31. Характеристическая функция. Свойства. Характеристическая функция нормального распределения

Определение. Характеристическая функция вещественнозначной случайной величины ξ — это
$$\varphi_\xi(t) = \mathbb{E}e^{i\xi t}.$$

Свойства 31.1 (характеристической функции).

1. $\varphi_\xi(0) = 1$, $|\varphi_\xi(t)| \le 1$;

2. $\varphi_{a\xi+b}(t) = e^{itb}\varphi_\xi(at)$;

3. Если ξ и η независимы, то $\varphi_{\xi+\eta}(t) = \varphi_\xi(t)\varphi_\eta(t)$

4. Если $\xi_1, \ldots, \xi_n$ независимы, то $\varphi_{\xi_1+\ldots+\xi_n}(t) = \varphi_{\xi_1}(t) \cdot \ldots \cdot \varphi_{\xi_n}(t)$

5. $\varphi_\xi(-t) = \overline{\varphi_\xi(t)}$

6. $\varphi_\xi(t)$ — равномерно непрерывна на $\mathbb{R}$

Доказательство.

2. $\varphi_{a\xi+b}(t) = \mathbb{E}(e^{it(a\xi+b)}) = e^{itb}\mathbb{E}e^{ai\xi t} = e^{itb}\varphi_\xi(at)$

3. $\varphi_{\xi+\eta}(t) = \mathbb{E}(e^{it(\xi+\eta)}) = \mathbb{E}(e^{it\xi} \cdot e^{it\eta}) = \mathbb{E}e^{it\xi}\mathbb{E}e^{it\eta} = \varphi_\xi(t)\varphi_\eta(t)$

6. $|\varphi_\xi(t+h) - \varphi_\xi(t)| = |\mathbb{E}(e^{i\xi(t+h))} - e^{it\xi})| = |\mathbb{E}(e^{it\xi}(e^{ih\xi} - 1))| \le \mathbb{E}|e^{ih\xi} - 1|$. Хотим доказать, что $\mathbb{E}|e^{ih\xi} - 1| \to 0$ при $h \to 0$. По определению матожидания
$$\mathbb{E}|e^{ih\xi} - 1| = \int\limits_{\mathbb{R}} |e^{ihx} - 1| \, \mathrm{d}P_\xi(x).$$
$|e^{ihx} - 1| \to 0$. По теореме Лебега подынтегральное выражение всегда ≤ 2, то есть это суммируемая мажоранта и можно менять предел и интеграл местами.
,
∎

Пример 31.1. Пусть $\xi \sim \mathcal{N}(a, \sigma^2)$, найдём его производящую функция $\varphi_\xi(t)$.

Пусть $\eta \sim \mathcal{N}(0,1)$; знаем, что $\xi = \sigma\eta + a$, $\varphi_\xi(t) = e^{ita}\varphi_\eta(\sigma t)$, поэтому достаточно найти характеристическую функциею для η.
$$\varphi_\eta(t) = \frac{1}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} e^{itx} e^{-\frac{x^2}{2}} \, \mathrm{d}x = \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} e^{-\frac{(x-it)^2}{2}} \, \mathrm{d}x.$$

Знаем, что $\int_{\mathbb{R}} f(x)\,\mathrm{d}x = \sqrt{2\pi}$, где $f(z) = e^{-\frac{z^2}{2}}$, хотим посчитать $I = \int\limits_{\mathbb{R}} f(x - it)\,\mathrm{d}x$. Сделаем это с помощью вычетов. Для этого посчитаем интеграл по контуру Γ_R:
$$0 = \int\limits_{\Gamma_R} f(z)\,\mathrm{d}z = \int\limits_{-R}^{R} + \int\limits_{R}^{R-it} + \int\limits_{R-it}^{-R-it} + \int\limits_{-R-it}^{-R}.$$

Знаем, что
$$\int\limits_{-R}^{R} \to \sqrt{2\pi}, \quad \int\limits_{R-it}^{-R-it} \to -I.$$

Оценим второй:

$$\left|\int\limits_R^{R-it} f(z)\,\mathrm{d}z\right| \le \int\limits_0^{-t} |e^{-\frac{(R-iy)^2}{2}}|\,\mathrm{d}y = \int\limits_0^{-t} (e^{-\frac{R^2}{2}+\frac{y^2}{2}})\,\mathrm{d}y \le t\cdot\max(e^{-\frac{R^2}{2}+\frac{y^2}{2}}) = t\cdot e^{\frac{t^2}{2}}\cdot e^{-\frac{R^2}{2}} \to 0.$$

Аналогично с последним интегралом. Итого получаем, что

$$I = \sqrt{2\pi} \implies \varphi_\eta(t) = \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\sqrt{2\pi} = e^{-\frac{t^2}{2}} \implies \varphi_\xi(t) = e^{ita}e^{-\frac{\sigma^2t^2}{2}}.$$

§32. Две теоремы о связи между математическим ожиданием и характеристической функцией

Теорема 32.1. Если $\mathbb{E}|\xi|^n < +\infty$, то при $k \le n$

$$\varphi_\xi^{(k)}(t) = \mathbb{E}((i\xi)^k e^{it\xi}).$$

Доказательство. Индукция. База $k = 0$ — определение характеристической функции.

Переход $k \to k+1$:

$$\begin{aligned}\varphi_\xi(t)^{(k+1)} &= \lim_{h\to 0}\frac{\mathbb{E}((i\xi)^k e^{i(t+h)\xi}) - \mathbb{E}((i\xi)^k e^{it\xi})}{h} = \lim_{h\to 0}\frac{\mathbb{E}((i\xi)^k e^{it\xi}\,\frac{e^{ih\xi}-1}{h})}{h}\\ &= \lim_{h\to 0}\int\limits_{\mathbb{R}} (ix)^k e^{itx}\frac{e^{ihx}-1}{h}\,\mathrm{d}P_\xi(x) = \int\limits_{\mathbb{R}} \lim_{h\to 0}(ix)^k e^{itx}\frac{e^{ihx}-1}{h}\,\mathrm{d}P_\xi(x)\\ &= \int\limits_{\mathbb{R}} (ix)^{k+1}e^{itx}\,\mathrm{d}P_\xi(x) = \mathbb{E}((i\xi)^{k+1}e^{it\xi}).\end{aligned}$$

Нужно пояснить, почему можно менять интеграл и предел местами. Покажем, что есть суммируемая мажоранта. Если $|xh| \ge 1$, то

$$\left|\frac{e^{ihx}-1}{h}\right| \le \frac{2}{|h|} = O(x).$$

Если $|xh| < 1$, то

$$\left|\frac{e^{ihx}-1}{h}\right| = \left|\frac{1+O(ihx)-1}{h}\right| = O(x).$$

Таким образом, в любом случае значение подынтегрального выражения не превосходит $|x|^kO(x) \le C|x|^{k+1}$. Это суммируемая мажоранта, так как интеграл по ней это $(k+1)$-й момент, который конечен по условию. ∎

Следствие 32.2.

$$\mathbb{E}\xi = -i\varphi'_\xi(0);$$

$$\mathbb{D}\xi = -\varphi''_\xi(0) + (\varphi'_\xi(0))^2.$$

Теорема 32.3. Если существует $\varphi''_\xi(0)$, то $\mathbb{E}\xi^2 < +\infty$.

Доказательство.

$$\mathbb{E}\xi^2 = \int\limits_{\mathbb{R}} x^2 \,\mathrm{d}P_\xi(x) = \int\limits_{\mathbb{R}} \lim_{t\to 0}\left(\frac{\sin(tx)}{t}\right)^2 \,\mathrm{d}P_\xi(x). \tag{$\star$}$$

Воспользуемся леммой Фату (интеграл от предела не превосходит предела от интеграла) и продолжим ($\star$) неравенством:

$$\begin{aligned}
\int\limits_{\mathbb{R}} \lim_{t\to 0}(\frac{\sin(tx)}{t})^2 \,\mathrm{d}P_\xi(x) \le \lim \int (\frac{\sin(tx)}{t})^2 \,\mathrm{d}P_\xi(x) = \lim_{t\to 0}\int\limits_{\mathbb{R}} (\frac{e^{itx}-e^{-itx}}{2it})^2 \,\mathrm{d}P_\xi(x) \\
&= \lim_{t\to 0} -\frac{1}{4t^2}\int\limits_{\mathbb{R}} (e^{2itx}+e^{-2itx}-2)\,\mathrm{d}P_\xi(x) \\
&= \lim_{t\to 0} -\frac{1}{4t^2}(\varphi_\xi(2t)+\varphi_\xi(-2t)-2) \\
&= [\varphi_\xi(s) = 1 + \varphi'_\xi(0)\cdot s + \frac{\varphi''_\xi(0)}{2}s^2 + o(s^2) = 1 + as + bs^2 + o(s^2)] \\
&= \lim_{t\to 0} -\frac{1}{4t^2}(1 + 2at + 4t^2 b + o(t^2) + 1 - 2at + 4t^2 b - 2) = -2b \\
&= \varphi''_\xi(0).
\end{aligned}$$

■

§33. Формула обращения

Теорема 33.1 (формула обращения). Пусть $a < b$ такие, что $P(\xi = a) = P(\xi = b) = 0$. Тогда

$$P(a \le \xi \le b) = \lim_{T\to+\infty}\frac{1}{2\pi}\int\limits_{-T}^{T}\frac{e^{-iat}-e^{-ibt}}{it}\varphi_\xi(t)\,\mathrm{d}t. \tag{†}$$

Заметим, что интеграл $\int_{-\infty}^{\infty}$ может расходится, а сходится должен только в смысле главного значения.

Доказательство. Шаг 1. Пусть $\xi = \frac{a+b}{2} + \frac{b-a}{2}\eta$. Тогда $\xi \in [a,b] \iff \eta \in [-1,1]$.

$$\varphi_\xi(t) = e^{i\frac{a+b}{2}t}\varphi_\eta(\frac{b-a}{2}t);$$
$$\varphi_\xi(t) = e^{i\frac{a+b}{2}t}\varphi_\eta(\frac{b-a}{2}t).$$
$$\int\limits_{-T}^{T}\frac{e^{-iat}-e^{-ibt}}{it}\varphi_\xi(t)\,\mathrm{d}t = \int\limits_{-T}^{T}\frac{e^{-iat}-e^{-ibt}}{it}e^{i\frac{a+b}{2}t}\varphi_\eta(\frac{b-a}{2}t)\,\mathrm{d}t$$

$$= \int_{-T}^{T} \frac{e^{-i(\frac{a-b}{2})t} e^{-i(\frac{b-a}{2})t}}{it} \varphi_\eta(\frac{b-a}{2}t)\,\mathrm{d}t = [u = \frac{b-a}{2}t]$$

$$= \int_{-T(\frac{b-a}{2})}^{T(\frac{b-a}{2})} \frac{e^{iu} - e^{-u}}{iu} \varphi_\eta(u)\,\mathrm{d}u. \qquad (*)$$

Если $(*)$ равна $2\pi P(-1 \le \eta \le 1) = 2\pi P(a \le \xi \le b)$, то доказали. Таким образом свели к частному случаю ($a = -1$, $b = 1$).

Шаг 2. Пусть $a = -1$, $b = 1$.

$$\lim_{T\to+\infty} \int_{-T}^{T} \frac{e^{it} - e^{-it}}{it} \int_{\mathbb{R}} e^{itx}\,\mathrm{d}P_\xi(x)\,\mathrm{d}t. \qquad (**)$$

Для подынтегрального выражения верно

$$\left| e^{itx} \frac{e^{it} - e^{-it}}{it} \right| = \left| \frac{e^{it} - e^{-it}}{it} \right| \le M,$$

значит есть суммируемость и можем применить теорему Фубини, поменяв интегралы местами. Продолжим $(**)$ равенством:

$$\lim_{T\to+\infty} \int_{\mathbb{R}} \int_{-T}^{T} \frac{e^{it} - e^{-it}}{it} \int_{\mathbb{R}} e^{itx}\,\mathrm{d}P_\xi(x)\,\mathrm{d}t = \lim_{T\to+\infty} \int_{\mathbb{R}} \Phi_T(x)\,\mathrm{d}P_\xi(x). \qquad (***)$$

Рассмотрим функция под интегралом.

$$\Phi_T(x) = \int_{-T}^{T} e^{itx} \frac{e^{it} - e^{-it}}{it}\,\mathrm{d}t = \int_{-T}^{T} \int_{-1}^{1} e^{itx} e^{iut}\,\mathrm{d}u\,\mathrm{d}t = [\text{по теореме Фубини}]$$

$$= \int_{-1}^{1} \int_{-T}^{T} e^{it(x+u)}\,\mathrm{d}t\,\mathrm{d}u \int_{-1}^{1} \frac{e^{i(x+u)t}}{(x+u)i}\Bigg|_{-T}^{T} \mathrm{d}u = 2\int_{-1}^{1} \frac{\sin((x+u)T)}{(x+u)}\,\mathrm{d}u = [v = (x+u)T]$$

$$= \int_{(x-1)T}^{(x+1)T} \frac{2\sin v}{v}\,\mathrm{d}v = F((x+1)T) - F((x-1)T).$$

Функция $F(y) = \int_0^y \frac{2\sin v}{v}\,\mathrm{d}v$ непрерывна на $\mathbb{R}$ и имеет предел в $\pm\infty$, значит это ограниченная функция, значит есть суммируемая мажоранта(константа) и в $(***)$ можно воспользоваться теоремой Лебега и поменять местами предел и интеграл. Осталось понять, чему равен $\lim_{T\to+\infty} F((x+1)T) - F((x-1)T)$. Если $x > 1$, то $x+1, x-1 > 0$, значит аргументы стремятся к $+\infty$, следовательно $F(\ldots) \to \pi$, итого получаем, что всё выражение стремится к пределу. Аналогично при $x < 1$ будет $x+1, x-1 > 0$, значит аргументы будут стремится к $-\infty$, $F(\ldots) \to \pi$ и снова предел равен нулю. Если

$x \in (-1, 1)$, тогда $\lim = 2\pi$.

Продолжим ($***$) равенством:

$$\lim_{T\to+\infty} \int\limits_{\mathbb{R}} \Phi_T(x)\,\mathrm{d}P_\xi(x) = \int\limits_{\mathbb{R}} \lim_{T\to+\infty} \Phi_T(x)\,\mathrm{d}P_\xi(x)$$
$$= \int\limits_{\mathbb{R}} 2\pi \mathbb{1}_{[-1,1]}(x)\,\mathrm{d}P_\xi(x) = 2\pi P(-1 \le \xi \le 1),$$

что и требовалось доказать. ■

§34. Следствия формулы обращения. Сумма независимых нормальных случайных величин

Следствие 34.1.

1. Если $\varphi_\xi = \varphi_\eta$, то $P_\xi = P_\eta$.
2. Если модуль характеристической функции суммируем — $\int_{\mathbb{R}} |\varphi_\xi(t)|\,\mathrm{d}t < +\infty$, то P_ξ имеет плотность и, более того,

$$p_\xi(x) = \frac{1}{2\pi} \int\limits_{\mathbb{R}} e^{-itx} \varphi_\xi(t)\,\mathrm{d}t.$$

Эта формула называется *преобразованием Фурье*, а обратная ей (определение характеристической функции) *обратным преобразование Фурье*.

Доказательство.

1. Пусть $A = \{x \in \mathbb{R} \mid P(\xi = x) > 0\}$. A — не более чем счётное, так как в точках из A функция распределения F_ξ имеет скачки. Тогда φ_ξ однозначно определяет $P(a \le \xi \le b)$, если $a, b \notin A$ (это из утверждения теоремы). Поймём, что она определяет и всё остальное распределение тоже. $F_\xi(b) = \lim\limits_{n\to+\infty} P(a_n \le \xi \le b)$, где $a_n \searrow -\infty$, $a_n \notin A$ однозначно определяется φ_ξ ($b \notin A$). Пусть $b \in A$, возьмем $b_n \searrow b$, $b_n \notin A$. Тогда, так как функция распределения непрерывна справа, то $F_\xi(b) = \lim\limits_{n\to+\infty} F_\xi(b_n)$ также однозначно определяется φ_ξ. Итого, однозначно определили P_ξ.
2. Поймём что при данных условиях интеграл из формулы обращения (†) сходится абсолютно.

$$\left|\frac{e^{-iat} - e^{-ibt}}{it} \varphi_\xi(t)\right| = \left|\frac{e^{-iat} - e^{-ibt}}{t}\right| |\varphi_\xi(t)| \le M |\varphi_\xi(t)|.$$

Последнее верно, так как $\frac{e^{-iat} - e^{-ibt}}{t}$ — ограниченная функция, поскольку она непрерывная и стремится к нулю при $t \to \pm\infty$.

$$\int\limits_a^b p_\xi(x)\,\mathrm{d}x = \frac{1}{2\pi}\int\limits_a^b \int\limits_{\mathbb{R}} e^{-itx}\varphi_\xi(t)\,\mathrm{d}t\,\mathrm{d}x. \qquad (\spadesuit)$$

Подынтегральное выражение не превосходит по модулю $|\varphi_\xi(t)|$ — суммируемая по условию. Значит можем применить теорему Фубини и в ($\spadesuit$) поменять порядок интегрирования.

$$\frac{1}{2\pi}\int\limits_a^b \int\limits_{\mathbb{R}} e^{-itx}\varphi_\xi(t)\,\mathrm{d}t\,\mathrm{d}x = \frac{1}{2\pi}\int\limits_{\mathbb{R}} \int\limits_a^b e^{-itx}\varphi_\xi(t)\,\mathrm{d}x\,\mathrm{d}t = \frac{1}{2\pi}\int\limits_{\mathbb{R}} \varphi_\xi(t)\frac{e^{-itx}}{-it}\bigg|_a^b \mathrm{d}t$$
$$= \frac{1}{2\pi}\int\limits_{\mathbb{R}} \frac{e^{-iat}-e^{-ibt}}{it}\varphi_\xi(t)\,\mathrm{d}t = P(a \le \xi \le b).$$

Чтобы было верно последнее равенство, то есть формула обращения, нужно условие $P(\xi = a) = P(\xi = b) = 0$. Поймем, что при условии следствия не может быть $P(\xi = c) > 0$; пусть $a_n \nearrow c$, $b_n \searrow c$ $(a_n, b_n \in A)$, тогда

$$P(\xi = c) = \lim_{n\to+\infty} P(a_n \le \xi \le b_n) = \lim_{n\to+\infty}\int\limits_{a_n}^{b_n} p_\xi(x)\,\mathrm{d}x \le (b_n - a_n)C \to 0.$$

■

Теорема 34.2. Пусть $\xi_k \sim \mathcal{N}(a_k, \sigma_k^2)$ — независимые случайные величины и $\eta = a_0 + \sum\limits_{k=1}^n c_k\xi_k$, причем хотя бы одна $c_k \ne 0$. Тогда

$$\eta \sim \mathcal{N}(a_0 + \sum_{k=1}^n c_k a_k, \sum_{k=1}^n c_k^2\sigma_k^2).$$

Доказательство. Рассмотрим характеристическую функцию ν.

$$\varphi_\eta(t) = e^{ita_0}\varphi_{\xi_1}(c_1 t)\cdot\ldots\cdot\varphi_{\xi_n}(ct_n) = e^{ita_0}e^{ic_1ta_1 - \frac{c_1^2t_1^2\sigma_1^2}{2}}\cdot\ldots$$
$$= e^{it(a_0+\ldots+a_nc_n)}e^{\frac{-t^2(c_1^2\sigma_1^2+\ldots+c_n^2\sigma_n^2)}{2}} = e^{iat-\frac{\sigma^2t^2}{2}}.$$

Это характеристическая функция $\mathcal{N}(a,\sigma^2)$, где

$$a = a_0 + \ldots + a_nc_n,\ \sigma^2 = c_1^2\sigma_1^2 + \ldots + c_n^2\sigma_n^2.$$

■

§35. Теорема о сходимости по распределению (все, кроме 7 $\Rightarrow$ 1)

Теорема 35.1. Пусть $\xi_1, \xi_2, \ldots$ последовательность случайных величин, $F_1, F_2, \ldots$ — их функции распределения, $\varphi_1, \varphi_2, \ldots$ — их характеристические функции. Тогда следующие условия равносильны:

1. $\xi_n \to \xi$ по распределению;
2. Для любого открытого $U \subset \mathbb{R}$ верно
$$\varliminf_{n\to+\infty} P(\xi_n \in U) \geq P(\xi \in U).$$
3. Для любого замкнутого $A \subset \mathbb{R}$ верно
$$\varlimsup_{n\to+\infty} P(\xi_n \in A) \leq P(\xi \in A).$$
4. Для любого борелевского регулярного множества B (то есть $P(\xi \in \operatorname{Cl} B \setminus \operatorname{Int} B) = 0$) верно
$$\lim_{n\to+\infty} P(\xi_n \in B) = P(\xi \in B).$$
5. Для любого борелевского регулярного множества B верно
$$\lim_{n\to+\infty} \mathbb{E}\mathbb{1}_B(\xi_n) = \mathbb{E}\mathbb{1}_B(\xi).$$
6. Для любой ограниченной непрерывной на $\mathbb{R}$ функции f верно
$$\lim_{n\to+\infty} \mathbb{E}f(\xi_n) = \mathbb{E}f(\xi).$$
7. $\varphi_n \to \varphi$ поточечно;

Доказательство.

$3 \implies 2$. Возьмём $U = \mathbb{R} \setminus A$, тогда $P(\xi_n \in U) = 1 - P(\xi_n \in A)$ и
$$\varliminf_{n\to+\infty} P(\xi_n \in U) = 1 - \varlimsup_{n\to+\infty} P(\xi_n \in A) \geq 1 - P(\xi \in A) = P(\xi \in U).$$

$2 \implies 3$. Аналогично предыдущему.

$4 \implies 5$. Понятно, что $P(\xi_n \in B) = \mathbb{E}\mathbb{1}_B(\xi_n)$ и $P(\xi \in B) = \mathbb{E}\mathbb{1}_B(\xi)$, так что
$$\lim_{n\to+\infty} P(\xi_n \in B) = \lim_{n\to+\infty} \mathbb{E}\mathbb{1}_B(\xi_n) = \mathbb{E}\mathbb{1}_B(\xi) = P(\xi \in B).$$

$5 \implies 4$. Аналогично предыдущему.

$6 \implies 7$.
$$\varphi_\xi(t) = \mathbb{E}e^{it\xi} = \mathbb{E}(\cos(\xi t)) + i\mathbb{E}(\sin(\xi t)).$$

$2+3 \implies 4$. Пусть $A = \mathrm{Cl}\, B$, $U = \mathrm{Int}\, B$ $P(\xi \in A \setminus U) = 0$. Тогда

$$\overline{\lim} P(\xi_n \in B) \le \overline{\lim} P(\xi_n \in A) \le P(\xi \in A) = P(\xi \in B) \le \underline{\lim} P(\xi_n \in A) \le \underline{\lim} P(\xi_n \in B).$$

Значит, предел существует и он равен $P(\xi \in B)$.

$1 \implies 2$. Про открытое знаем, что

$$U = \bigsqcup_{k=1}^{+\infty} (a_k, b_k], \qquad \text{где} \qquad P(\xi = a_k) = P(\xi = b_k) = 0.$$

$$P(\xi_n \in (a_k, b_k]) = F_n(b_k) - F_n(a_k) \to F(b_k) - F(a_k) = P(\xi \in (a_k, b_k]);$$

Получаем, что

$$P(\xi_n \in U) \ge P(\xi_n \in \bigsqcup_{k=1}^{m} (a_k, b_k]) = \sum_{k=1}^{m} P(\xi_n \in (a_k, b_k])$$
$$\underset{m\to\infty}{\to} \sum_{k=1}^{m} P(\xi \in (a_k, b_k]) = P(\xi \in \bigsqcup_{k=1}^{m} (a_k, b_k]);$$

Возьмём нижний предел:

$$\underline{\lim} P(\xi_n \in U) \ge \underline{\lim} P(\xi_n \in \bigsqcup_{k=1}^{m} (a_k, b_k]) = \lim \ldots = P(\xi \in \bigsqcup_{k=1}^{m} (a_k, b_k]) \underset{m\to\infty}{\to} P(\xi \in U).$$

Значит, $\underline{\lim} P(\xi_n \in U) \ge P(\xi \in U)$.

$5 \implies 6$. Обозначим $D = \{x \in \mathbb{R} \mid P(f(\xi) = x) > 0\} = \{x \mid P_\xi(f^{-1}) > 0\}$ —не более, чем счётное (такие точки — точки разрыва функции распределения). По условию $f \in C(\mathbb{R})$ и $|f| \le M$. Нарежем $[-M, M]$ на маленькие кусочки: $-M = t_0 < t_1 < \ldots < t_n = M$ такие, что $\forall i\ t_i \notin D$ (такое возможно, так как D не более чем счетно). Введем следующие множества A_j, B_j, U_j:

$$A_j = \{x \mid t_{j-1} \le f(x) \le t_j\} \supset B_j = \{x \mid t_{j-1} \le f(x) < t_j\} \supset \{x \mid t_{j-1} < f(x) < t_j\} = U_j.$$

A_j — замкнуто, U_j — открыто и $P(f(\xi) \in A_j \setminus U_j) = 0$, потому что это означает, что $f(\xi) \in t_{j-1}, t_j$, а эти точки не лежат в D, поэтому у них вероятность нулевая. Итого, B_j — регулярное Из этих B_j теперь соорудим ступенчатую функцию

$$g(x) = \sum_{j=1}^{m} t_j \cdot \mathbb{1}_{B_j}(x).$$

По 5-му пункту $\lim \mathbb{E} g(\xi_n) = \mathbb{E} g(\xi)$. Заметим, что

$$|f(x) - g(x)| = \max\{|t_j - t_{j-1}|\},$$

тогда это верно и для матожидания:

$$|\mathbb{E}f(\xi) - \mathbb{E}g(\xi)| \le \mathbb{E}|f(\xi) - g(\xi)| \le \max|t_j - t_{j-1}|.$$

$$\begin{aligned} |\mathbb{E}f(\xi_n) - \mathbb{E}f(\xi)| &\le |\mathbb{E}f(\xi_n) - \mathbb{E}g(\xi_n)| + |\mathbb{E}g(\xi_n) - \mathbb{E}g(\xi)| + |\mathbb{E}g(\xi) - \mathbb{E}f(\xi)| \\ &< 2\varepsilon + |\mathbb{E}g(\xi_n) - \mathbb{E}g(\xi)| < 3\varepsilon. \end{aligned}$$

Последнее верно при достаточно большом n. ■

§36. Теорема о сходимости по распределению (7 ⇒ 1)

Теорема 36.1. $7 \implies 1$ ($\varphi_n \to \varphi$ поточечно $\implies$ $\xi_n \to \xi$ по распределению

Доказательство. Возьмем $\eta_\sigma \sim \mathcal{N}(0, \sigma^2)$, не зависящую от $\xi_1, \xi_2, \dots$

Для независимых величин хар. функция - произведение хар. функций.

$$\varphi_{\xi_n + \eta_\sigma}(t) = \varphi_n(t)\varphi_{\eta_\sigma}(t) = \varphi_n(t)e^{\frac{-\sigma^2 t^2}{2}}$$

φ для нормального распределения мы уже считали, поэтому смогли написать.

$|\varphi_n(t)| \le 1$, поэтому можно написать:

$$|\varphi_{\xi_n + \eta_\sigma}(t)| \le e^{\frac{-\sigma^2 t^2}{2}}$$

В частности, интеграл от этой хар. функции сходится.

Пусть $G_n(x) := F_{\xi_n + \eta_\sigma}$. Это непрерывная функция: мы доказывали, что если к произвольному распределению прибавить непрерывное, то результатом будет непрерывное.

Напишем формулу обращения (из-за непрерывности можно выбирать любые a, b):

$$G_n(b) - G_n(a) = \frac{1}{2\pi} \int\limits_{\mathbb{R}} \frac{e^{-iat} - e^{-ibt}}{2i} \varphi_n(t) e^{\frac{-\sigma^2 t^2}{2}} \, dt$$

Написали сразу интеграл а не предел, потому что интеграл от $e^{\frac{-\sigma^2 t^2}{2}}$ сходится, а остальной множитель ограничен. Значит весь интеграл сходится. В частности, есть суммируемая мажоранта, поэтому можно перейти к пределу под знаком интеграла (теорема Лебега):

$$\frac{1}{2\pi} \int\limits_{\mathbb{R}} \frac{e^{-iat} - e^{-ibt}}{2i} \varphi_n(t) e^{\frac{-\sigma^2 t^2}{2}} \, dt \to \frac{1}{2\pi} \int\limits_{\mathbb{R}} \frac{e^{-iat} - e^{-ibt}}{2i} \varphi(t) e^{\frac{-\sigma^2 t^2}{2}} \, dt = G(b) - G(a)$$

Получили $G_n(b)-G_n(a) \to G(b)-G(a)$ для всех $a,\ b$. Следовательно, $G_n(x) \to G(x)$ поточечно. (Также обсуждали, что сходимость разностей равносильна сходимости поточечной)

Осталось доказать сходимость F_n. Пусть x - точка непрерывности F, и $\epsilon > 0$. Возьмем из непрерывности $\delta > 0$, такое что $|F(x \pm 2\delta) - F(x)| < \epsilon$.

$F_n(x) = P(\xi_n \leq x)$.

$$\{\xi_n \leq x\} \subset \{\xi_n + \eta_\sigma \leq x + \delta\} \cup \{|\eta_\sigma| \geq \delta\}$$

$P(|\eta_\sigma| \geq \delta) \leq \frac{\mathbb{D}\eta_\sigma}{\delta^2} = \frac{\sigma^2}{\delta^2}$ - неравенство Чебышева.

$G_n \to G$ поточечно, поэтому $G_n(x+\delta) < G(x+\delta) + \epsilon$

Теперь можно написать

$$F_n(x) = P(\xi_n \leq x) \leq P(\xi_n+\eta_\sigma \leq x+\delta)+P(|\eta_\sigma| \geq \delta) \leq G_n(x+\delta)+\frac{\delta^2}{\sigma^2} \leq G(x+\delta)+\epsilon+\frac{\delta^2}{\sigma^2} \leq$$

Аналогично оценим G:

$$\{\xi + \eta_\sigma \leq x + \delta\} \subset \{\xi \leq x + 2\delta\} \cup \{|\eta_\sigma| \geq \delta\}$$

Продолжим неравенство

$$< F(x+2\delta) + 2\frac{\delta^2}{\sigma^2} + \epsilon < F(x) + 2\epsilon + 2\frac{\delta^2}{\sigma^2}$$

Оценка снизу получается аналогично, надо рассмотреть

$\{\xi_n \leq n\} \supset \{\xi_n + \eta_\sigma \leq x - \delta\} \setminus \{|\eta_\sigma| \geq \delta\}$

Получим такую оценку:

$F_n(x) > F(x) - 2\epsilon - 2\frac{\delta^2}{\sigma^2}$

Порядок выбора маленьких величин: сначала ϵ, по нему δ, по нему σ. После этого выбора все G_n зафиксированы, как и рассматриваемая точка, и можно переходить к пределу по n и выбрать такое N, что $|G_n(x+\delta) - G(x+\delta)| < \epsilon$, аналогично для $x - \delta$.

Теперь $2\epsilon + \frac{2\sigma^2}{\delta^2} < 4\epsilon$, что и требовалось доказать.

■

§37. Равномерная сходимость к непрерывной функции распределения. Центральная предельная теорема в форме Леви

Теорема 37.1. Пусть функции $F_n, F : \mathbb{R} \to [0,1]$ монотонно возрастают и F непрерывна на $\mathbb{R}$. Тогда, если $F_n \to F$ поточечно, то $F_n \rightrightarrows F$ равномерно.

Доказательство. Найдем t_k, такие что $F(t_k) = \frac{k}{m}$, $t_0 < \ldots < t_m$. По условию

$$|F_n(t_k) - F(t_k)| < \varepsilon = \frac{1}{m},$$

начиная с некоторого номера. Рассмотрим $t_k \le t < t_{k+1}$:

$$F_n(t_k) \le F_n(t) \le F_n(t_{k+1}) < F(t_{k+1}) + \varepsilon = \frac{k+1}{m} + \frac{1}{m} = \frac{k}{m} + 2\varepsilon = F(t_k) + 2\varepsilon \le F(t) + 2\varepsilon.$$

$$F_n(t_k) > F(t_k) - \varepsilon = \frac{k}{m} - \frac{1}{m} = \frac{k+1}{m} - 2\varepsilon = F(t_{k+1}) - 2\varepsilon \ge F(t) - 2\varepsilon.$$

Итого

$$F(t) - 2\varepsilon < F_n(t) < F(t) + 2\varepsilon.$$

■

Теорема 37.2 (Центральная предельная теорема в форме Поля Леви[1]). Пусть $\xi_1, \ldots$ независимые одинаково распределенные случайные величины. $a = \mathbb{E}\xi_1$, $\sigma^2 = \mathbb{D}\xi_1$, $S_n = \xi_1 + \ldots + \xi_n$. Тогда

$$P\left(\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}} \le x\right) = P\left(\frac{S_n - na}{\sqrt{n}\sigma} \le x\right) \rightrightarrows \Phi = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{u^2}{2}}\,\mathrm{d}u.$$

Доказательство. Обозначим

$$\varphi(t) = \varphi_{\xi_1 - a}(t) = 1 - \frac{\sigma^2 t^2}{2} + o(t^2).$$

Равенство верно, поскольку у нас есть дисперсия, то характеристическая функция дважды дифференцируема, поэтому есть формула через матожидание и дисперсию для неё. Пусть

$$\varphi_n(t) = \varphi_{\frac{S_n - na}{\sqrt{n}\sigma}}(t) = \varphi_{S_n - an}(\frac{t}{\sqrt{n}\sigma}).$$

Тогда

$$S_n - an = (\xi_1 - a) + \ldots + (\xi_n - a),$$

и, так как ξ_i независимы, то

$$\varphi_{S_n - an}\left(\frac{t}{\sqrt{n}\sigma}\right) = \left(\varphi\left(\frac{t}{\sqrt{n}\sigma}\right)\right)^n = \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n = \left(1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n \to e^{-\frac{t^2}{2}}.$$

Значит, $\frac{S_n - na}{\sqrt{n}\sigma} \to \mathcal{N}(0,1)$ по распределению, следовательно, поскольку предельная функция непрерывна, то сходимость будет равномерной. ■

[1] Леви́

§38. Интегральная теорема Муавра–Лапласа. Теорема Пуассона

Следствие 38.1 (теорема Муавра-Лапласа). Пусть $\xi_1, \ldots$ — независимые испытания Бернулли с вероятностью успеха $p \in (0,1)$, $S_n = \sum\limits_{i=1}^{n} \xi_i$. Тогда

$$P\left(\frac{S_n - np}{\sqrt{npq}} \le x\right) \rightrightarrows \Phi(x).$$

Доказательство. $\mathbb{E}\xi_1 = p, \mathbb{D}\xi_1 = pq$, подставим в теорему. ■

Теорема 38.2 (Пуассона). Пусть

$$P(\xi_{nk} = 1) = p_{nk},\ P(\xi_{nk} = 0) = 1 - p_{nk} = q_{nk};$$
$$a_n = \max_{1 \le k \le n} p_{nk} \underset{n\to\infty}{\to} 0,\ p_{n1} + \ldots + p_{nn} \to \lambda > 0;$$

события ξ_{nk} независимы при фиксированном n. Тогда

$$P(S_n = m) \underset{n\to\infty}{\to} \frac{e^{-\lambda}\lambda^m}{m!},$$

где $S_n = \xi_{n1} + \ldots + \xi_{nn}$.

Доказательство. Докажем методом характеристических функций.

$$\varphi_{\xi_{nk}}(t) = \mathbb{E}e^{it\xi_{nk}} = p_{nk}e^{it} + 1 - p_{nk}.$$

Надо доказать, что характеристическая функция $\varphi_{S_n}(t)$ стремится к $e^{\lambda(e^{it}-1)}$ (так называемая характеристическая функция Пуассона), и тогда равенство из теоремы очевидно верно.

$$\varphi_{S_n}(t) = \prod_{k=1}^{n} \varphi_{\xi_{nk}}(t) = \prod_{k=1}^{n}(1 + p_{nk}(e^{it} - 1)),$$

Прологарифмируем, тогда надо показать, что

$$\sum \ln(1 + p_{nk}(e^{it} - 1)) \to \lambda(e^{it} - 1).$$

Распишем левую часть:

$$\sum \ln(1 + p_{nk}(e^{it} - 1)) = \sum((p_{nk}(e^{it} - 1)) + O(p_{nk}^2)) \to \lambda(e^{it} - 1) + \sum_{k=1}^{n} O(p_{nk}^2).$$

Оценим второе слагаемое:

$$\sum_{k=1}^{n} O(p_{nk}^2) \le \sum_{k=1}^{n} O(a_n p_{nk}) = a_n O\left(\sum p_{nk}\right) \le Ca_n \to 0.$$

■

§39. Центральная предельная теорема в форме Линденберга (без доказательства). Центральная предельная теорема в форме Ляпунова. Оценки на скорость сходимости

Теорема 39.1 (Центральная предельная теорема в форме Линденберга). Пусть $\xi_1, \xi_2, \ldots$ — независимые случайные величины. $a_k = \mathbb{E}\xi_k, \sigma_k^2 = \mathbb{D}\xi_k > 0, S_n = \xi_1 + \ldots + \xi_n$. Обозначим

$$\mathrm{Lind}(\varepsilon, n) = \frac{1}{\mathbb{D}_n^2} \sum_{k=1}^{n} \mathbb{E} f(\xi_k - a_k),$$

где

$$\mathbb{D}_n^2 = \sum_{k=1}^{n} \sigma_k^2, \; f(x) = x^2 \mathbb{1}_{\{|x| \geq \varepsilon \mathbb{D}_n\}}(x).$$

Тогда, если $\mathrm{Lind}(\varepsilon, n) \to 0$ при $n \to \infty$ при всех $\varepsilon > 0$, то

$$P(\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}} \leq x) \rightrightarrows \Phi.$$

Упражнение. Проверьте, что для независимых одинаково распределенных случайных величин с конечной дисперсией выполняется условие Линденберга

Теорема 39.2 (Центральная предельная теорема в форме Ляпунова). Пусть $\xi_1, \xi_2, \ldots$ — независимые случайные величины. $a_k = \mathbb{E}\xi_k$, $\sigma_k^2 := \mathbb{D}\xi_k > 0$, $S_n := \xi_1 + \ldots + \xi_n$. Обозначим

$$L(\delta, n) = \frac{1}{\mathbb{D}_n^{2+\delta}} \sum_{k=1}^{n} \mathbb{E}|\xi_k - a_k|^{2+\delta}, \quad \text{где} \quad \mathbb{D}_n^2 = \sum_{k=1}^{n} \sigma_k^2.$$

Тогда, если $L(\delta, n) \to 0$ при $n \to \infty$ при некотором $\delta > 0$, то

$$P(\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}} \leq x) \rightrightarrows \Phi.$$

Доказательство. Докажем, что из теоремы в форме Линденберга следует теорема в форме Ляпунова, то есть надо показать, что из условия Ляпунова следует условие Линденберга.

$$\mathrm{Lind}(\varepsilon, n) = \frac{1}{\mathbb{D}_n^2} \sum_{k=1}^{n} \mathbb{E}((\xi_k - a_k)^2 \mathbb{1}_{\{|\xi_k - a_k| \geq \varepsilon \mathbb{D}_n\}}(\xi_k - a_k))$$

$$\leq \frac{1}{\mathbb{D}_n^2}\sum_{k=1}^{n}\mathbb{E}((\xi_k - a_k)^2(\frac{|\xi_k - a_k|}{\varepsilon\mathbb{D}_n})^\delta)$$
$$= \frac{1}{\varepsilon^\delta}\frac{1}{\mathbb{D}_n^{2+\delta}}\sum_{k=1}^{n}\mathbb{E}|\xi_k - a_k|^{2+\delta}$$
$$= \frac{L(\delta, n)}{\varepsilon^\delta} \underset{n\to\infty}{\longrightarrow} 0.$$

■

Теорема 39.3. Пусть $0 < \delta \leq 1$. Тогда в условии центральной предельной теоремы в форме Ляпунова.

$$\sup_{\sup x\in\mathbb{R}}\left|(P(\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}} \leq x) - \Phi\right| \leq C_\delta L(\delta, n).$$

Замечание. Пусть случайные величины ξ_i независимы и одинаково распределены, $a = \mathbb{E}\xi_1$, $\sigma^2 = \mathbb{D}\xi_1$, $\mathbb{D}_n^2 = n\sigma^2$. Тогда

$$L(\delta, n) = \frac{1}{(\sqrt{n}\sigma)^{2+\delta}}n\mathbb{E}|\xi_1 - a|^{2+\delta} = \frac{1}{n^{\frac{\delta}{2}}}\cdot\frac{1}{\sigma^{2+\delta}}\mathbb{E}|\xi_1 - a|^{2+\delta}.$$

Теорема 39.4 (Берри-Эссеена). Пусть $\xi_1, \xi_2, \ldots$ — независимые и одинаково распределенные случайные величины;$\mathbb{E}\xi_1 = a$. Тогда

$$\left|P(\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}} \leq x) - \Phi\right| \leq \frac{C\mathbb{E}|\xi_1 - a_1|^3}{\sqrt{n}\sigma^3}.$$

Теорема 39.5 (Хартмана — Витнера, "закон повторного логарифма"). Пусть $\xi_1, \xi_2, \ldots$ — независимые и одинаково распределенные случайные велЬчины, $\mathbb{E}\xi_1 = 0$, $\mathbb{D}\xi_1 = \sigma^2 > 0$. Тогда

$$\overline{\lim_{n\to\infty}}\frac{S_n}{\sqrt{2n\ln\ln n}} = \sigma;$$
$$\underline{\lim_{n\to\infty}}\frac{S_n}{\sqrt{2n\ln\ln n}} = -\sigma.$$

Теорема 39.6 (Штрассена). Любая точка из $[-\sigma, \sigma]$ — предельная точка последовательности $\frac{S_n}{\sqrt{2n\ln\ln n}}$.

§40. Оценки Чернова для больших уклонений. Примеры функций уклонения

Теорема 40.1 (закон больших чисел в форме Хинчина). Пусть $\xi_1, \xi_2, \ldots$ — независимые одинаково распределенные случайные величины, $a = \mathbb{E}\xi_1$, $S_n = \xi_1 + \xi_2 + \ldots + \xi_n$. Тогда для всех $r > a$

$$P\left(\frac{S_n}{n} \geq r\right) \to 0.$$

Доказательство. По неравенству Чебышёва

$$P\left(\frac{S_n}{n} \geq r\right) \leq \frac{\mathbb{D}\left(\frac{S_n}{n}\right)}{(r-a)^2} = \frac{\mathbb{D}\xi_1}{n(r-a)^2} \underset{n\to+\infty}{\to} 0.$$

■

Определение. Случайная величина ξ удовлетворяет *условию Крамера*, если для всех λ в некоторой окрестности нуля выполняется $\mathbb{E}e^{\lambda\xi} < +\infty$.

Оценка Чернова. Пусть $\lambda \geq 0$. Тогда

$$P\left(\frac{S_n}{n} \geq r\right) = P(S_n\lambda \geq \lambda nr) = P(e^{\lambda S_n} \geq e^{\lambda nr}) \leq \frac{\mathbb{E}e^{\lambda S_n}}{e^{\lambda nr}} = \frac{\mathbb{E}\prod_{k=1}^n e^{\lambda\xi_k}}{e^{\lambda nr}} = \left(\frac{\mathbb{E}e^{\lambda\xi_1}}{e^{\lambda r}}\right)^n.$$

Пусть

$$\psi(\lambda) = \ln\mathbb{E}e^{\lambda\xi_1},$$

тогда

$$P\left(\frac{S_n}{n} \geq r\right) \leq \exp(n(\psi(\lambda) - \lambda r)).$$

Введём обозначение

$$I(r) = \sup_{\lambda>0}(\lambda r - \psi(\lambda)).$$

Это функция называется функцией отклонения. Тогда

$$P\left(\frac{S_n}{n} \geq r\right) \leq \exp(-nI(r)).$$

Примеры 40.1 (оценок Чернова).

1. Пусть $\xi_k \sim \mathcal{N}(0,1)$. Тогда

$$\psi(\lambda) = \ln\mathbb{E}e^{\lambda\xi} = \ln\left(\frac{1}{\sqrt{2\pi}}\int\limits_{\mathbb{R}} e^{\lambda t}e^{-\frac{t^2}{2}}\right)\mathrm{d}t = \left[\lambda t - \frac{t^2}{2} = \frac{-(t-\lambda)^2}{2} + \frac{\lambda^2}{2}\right] = \frac{\lambda^2}{2}.$$

   Найдём функцию отклонения. Максимум выражения $\lambda r - \psi(\lambda) = \lambda r - \frac{\lambda^2}{2}$ достигается при $\lambda = r$, значит $I(r) = \frac{r^2}{2}$ и

$$P\left(\frac{S_n}{n} \geq r\right) \leq e^{\frac{-nr^2}{2}}.$$

2. Пусть $\xi_k \sim \mathrm{Exp}(1)$. Тогда

$$\psi(\lambda) = \ln\mathbb{E}^{\lambda\xi} = \ln\left(\int\limits_0^{+\infty} e^{\lambda t}e^{-t}\,\mathrm{d}t\right) = \ln\left(\frac{1}{1-\lambda}\right)$$

при $\lambda < 1$. Максимум выражения $\lambda r - \psi(\lambda) = \lambda r + \ln(1-\lambda)$ достигается при $\lambda = \frac{r-1}{r}$, значит $I(r) = r - 1 - \ln r$ и

$$P\left(\frac{S_n}{n} \geq r\right) \leq e^{-n(r-1-\ln r)}.$$

Упражнение. Пусть $\xi_k \sim \mathrm{Bern}(p_k)$. $\mu = p_1 + p_2 + \ldots + p_k$. Тогда

$$\forall \delta > 0 \;\; P(S_n \geq (1+\delta)\mu) < \exp\left(\frac{-\delta^2\mu}{\delta+2}\right).$$

§41. Условные математические ожидания относительно событий и относительно разбиений. Примеры

Определение. Пусть $(\Omega, \mathcal{F}, P)$ — вероятностное пространство. $\mathcal{A} \subset \mathcal{F}$ — σ-алгебра; ξ — случайная величина, $\mathbb{E}|\xi| < +\infty$. Тогда случайная величина $\eta = \mathbb{E}(\xi|\mathcal{A})$ — *математическое ожидание при условии* $\mathcal{A}$, если

1. η измерима относительно $\mathcal{A}$;
2. $\forall A \in \mathcal{A} \quad \mathbb{E}(\xi \mathbb{1}_A) = \mathbb{E}(\eta \mathbb{1}_A)$.

Пример 41.1. $\mathcal{A} = \{\varnothing, \Omega\}$, $\eta = \mathrm{const}$, $\mathbb{E}\xi = \mathbb{E}(\mathrm{const}) = \mathrm{const}$.

Теорема 41.1. Условное матожидание $\mathbb{E}(\xi|\mathcal{A})$ существует и единственно в следующем смысле: если η_1 и η_2 — условные матожидания, то они равны почти наверное.

Доказательство. Докажем единственность. Пусть $A = \{\eta_1 > \eta_2\} \in \mathcal{A}$. Тогда

$$\begin{aligned}\mathbb{E}(\eta_1 \mathbb{1}_A) = \mathbb{E}(\xi \mathbb{1}_A) = \mathbb{E}(\eta_2 \mathbb{1}_A) \implies \mathbb{E}((\eta_1 - \eta_2)\mathbb{1}_A) = 0 \\ \implies P((\eta_1 - \eta_2)\mathbb{1}_A = 0) = 1 \implies P(\eta_1 > \eta_2) = 0.\end{aligned}$$

Аналогичное доказательство для $\{\eta_1 < \eta_2\}$, значит $P(\eta_1 = \eta_2) = 1$.

Докажем существование. Пусть $\mu_\pm A = \mathbb{E}(\xi_\pm \mathbb{1}_A) \geq 0$ — меры на $\mathcal{A}$ (это меры, так как $\mathbb{E}$ счетно аддитивно). Если $P(A) = 0$, то $\mu_\pm A = 0$, то есть эти меры абсолютно непрерывна относительно P. По теореме Радона-Никодима существует функция $\omega_\pm \geq 0$, измеримая относительно $\mathcal{A}$ и суммируемая, такая что

$$\mu_\pm A = \int\limits_A \omega_\pm \, \mathrm{d}P = \int\limits_\Omega \omega_\pm \mathbb{1}_A \, \mathrm{d}P.$$

Тогда

$$\mathbb{E}(\xi_\pm \mathbb{1}_A) = \mathbb{E}(\omega_\pm \mathbb{1}_A) \implies \mathbb{E}(\xi \mathbb{1}_A) = \mathbb{E}((\omega_+ - \omega_-)\mathbb{1}_A),$$

значит функция $\omega_+ - \omega_-$ подходит. ∎

Свойства 41.2 (условных матожиданий).

1. $\mathbb{E}(c|\mathcal{A}) = c$;
2. $\mathbb{E}(\xi|\mathcal{A})$ линейно;
3. Если $\mathcal{A}_1 \subset \mathcal{A}_2$, то $\mathbb{E}(\xi|\mathcal{A}_1) = \mathbb{E}(\mathbb{E}(\xi|\mathcal{A}_2)|\mathcal{A}_1)$ почти наверное.
4. $\mathbb{E}\xi = \mathbb{E}(\mathbb{E}(\xi|\mathcal{A}))$;
5. Если ξ измеримо относительно $\mathcal{A}$, то $\mathbb{E}(\xi|\mathcal{A}) = \xi$;
6. Если $\xi \leq \eta$, то $\mathbb{E}(\xi|\mathcal{A}) \leq \mathbb{E}(\eta|\mathcal{A})$;

Доказательство.

1. Очевидно из определения;
2. Очевидно из определения;
3. Надо показать, что $\eta == \mathbb{E}(\mathbb{E}(\xi|\mathcal{A}_2)|\mathcal{A}_1)$ подходит под условие условного матожидания, то есть она измерима относительно $\mathcal{A}_1$ (что сразу следует из определения η) и $\mathbb{E}(\eta\mathbb{1}_A) = \mathbb{E}(\xi\mathbb{1}_A) \quad \forall A \in \mathcal{A}$. Докажем второе; по определению

   $$\mathbb{E}(\eta\mathbb{1}_A) = \mathbb{E}(\mathbb{E}(\xi|\mathcal{A}_2)\mathbb{1}_A) = \mathbb{E}(\xi\mathbb{1}_A).$$

   Последнее равенство верно так как $A \in \mathcal{A}_2$.
4. Подставим в 3-й пункт $\mathcal{A}_1 = \{\varnothing, \Omega\}$ и $\mathcal{A}_2 = \mathcal{A}$).
5. Нужно проверить, что $\forall A \in \mathcal{A}$ $\mathbb{E}(\xi\mathbb{1}_A) = \mathbb{E}(\mathbb{E}(\xi|\mathcal{A})\mathbb{1}_A)$, ну а в случае $\mathbb{E}(\xi|\mathcal{A}) = \xi$ и проверять нечего.
6. Если $\xi \geq 0$, то $\mathbb{E}(\xi|\mathcal{A}) \geq 0$. Вспомним, как мы находили условное матожидание в теореме о его существовании(41.1), а именно как сумму неотрицательных величин. ■

Приведём важный пример условного матожидания.

Пример 41.2 (Условное матожидание относительно разбиения). Пусть $\Omega = \bigsqcup A_k$, а $\mathcal{A}$ — натянутая на A_k σ-алгебра. Тогда условное матожидание $\mathbb{E}(\xi|\mathcal{A})$ —- случайная величина и из определения $\mathcal{A}$ понятно, что они должны быть константными на A_k, значит $\mathbb{E}(\xi|\mathcal{A}) = \sum c_k\mathbb{1}_{A_k}$. Тогда

$$\mathbb{E}(\xi\mathbb{1}_A) = \mathbb{E}(\sum c_k\mathbb{1}_{A_k}\mathbb{1}_A) \quad \forall A \in \mathcal{A}.$$

Подставим $A = A_n$:

$$\mathbb{E}(\xi\mathbb{1}_{A_n}) = \mathbb{E}(\sum c_k\mathbb{1}_{A_k}\mathbb{1}_{A_n}) = c_nP(A_n) \implies c_n = \frac{\mathbb{E}(\xi\mathbb{1}_{A_n})}{P(A_n)}.$$

Итого

$$\mathbb{E}(\xi|\mathcal{A}) = \sum \frac{\mathbb{E}(\xi\mathbb{1}_{A_k})}{P(A_k)}\mathbb{1}_{A_k}.$$

Определение. *Условная вероятность относительно* $\mathcal{A}$ — это

$$P(B|\mathcal{A}) = \mathbb{E}(\mathbb{1}_B|\mathcal{A}).$$

Определение. Пусть η — случайная величина; $\sigma(\eta)$ — σ-алгебра, натянутая на множества $\{\eta \le c\}$. Тогда *условным матождианием относительно случайной величины* η называется

$$\mathbb{E}(\xi|\eta) = \mathbb{E}(\xi|\sigma(\eta)).$$

Пример 41.3. Пусть η — дискретная случайная величина, $\{\eta = y_k\}$ — измеримы, где $\{y_k\}$ — значения η. Тогда

$$\mathbb{E}(\xi|\eta) = \sum \frac{\mathbb{E}(\xi \mathbb{1}_{\{\eta=y_k\}})}{P(\{\eta = y_k\})} \mathbb{1}_{\{\eta=y_k\}} = \sum \mathbb{E}(\xi|\eta = y_k) \mathbb{1}_{\{\eta=y_k\}}.$$

§42. Математическое ожидание и производящая функция суммы случайного количества случайных величин

Пример 42.1. Пусть $N, \xi_1, \xi_2, \dots$ — независимые случайные величины, $\xi_1, \xi_2, \dots$ одинаково распределены, $\mathbb{E}\xi_1 = a$; обозначим $S = \xi_1 + \ldots + \xi_N$. Найдем $\mathbb{E}S$:

$$\begin{aligned}\mathbb{E}S &= \mathbb{E}(\mathbb{E}(S|N)) = \mathbb{E}(\sum \mathbb{E}(S|N=n)\mathbb{1}_{\{N=n\}}) = \sum \mathbb{E}(S|N=n)P(N=n) \\ &= \sum \mathbb{E}(\xi_1 + \ldots + \xi_n)P(N=n) = \sum naP(N=n) = a\sum nP(N=n) = a\mathbb{E}N.\end{aligned}$$

Упражнение. Найдите $\mathbb{D}S$, если известно $\mathbb{D}\xi_1, \mathbb{D}N, \mathbb{E}\xi_1, \mathbb{E}N$.

Пример 42.2. Пусть случайные величины $N, \xi_1, \xi_2, \dots$ принимают натуральные значения. G — производящая функция N, F — производящая функция ξ_1. Найдем производящую функцию для S:

$$\begin{aligned}\mathbb{E}z^S &= \mathbb{E}(\mathbb{E}(z^S|N)) = \sum \mathbb{E}(z^S|N=n)P(N=n) = \sum \mathbb{E}(z^{\xi_1+\ldots+\xi_n})P(N=n) \\ &= \sum \mathbb{E}z^{\xi_1} \cdot \ldots \cdot \mathbb{E}z^{\xi_n} P(N=n) = \sum (F(z))^n P(N=n) = G(F(z)).\end{aligned}$$

Замечание (геометрическая интерпретация). Рассмотрим ξ, такие что $\mathbb{E}\xi^2 < +\infty$. Они образуют пространство $L^2(\Omega, \mathcal{F}, P)$; возьмём σ-алгебру $\mathcal{A} \subset \mathcal{F}$. Условные матожидания живут в пространстве $L^2(\Omega, \mathcal{A}, P) \subset L^2(\Omega, \mathcal{F}, P)$. Тогда условное матожидание $\mathbb{E}(\xi|\mathcal{A})$ — проекция ξ на $L^2(\Omega, \mathcal{A}, P)$.

Доказательство. Нужно доказать, что

$$\xi - \mathbb{E}(\xi|\mathcal{A}) \perp L^2(\Omega, \mathcal{A}, P).$$

Достаточно проверить на $\mathbb{1}_A$ для $A \in \mathcal{A}$, то есть проверить, что

$$\mathbb{E}((\xi - \mathbb{E}(\xi|\mathcal{A}))\mathbb{1}_A) = 0,$$

что равносильно тому, что

$$\mathbb{E}(\xi\mathbb{1}_A) = \mathbb{E}(\mathbb{E}(\xi|\mathcal{A})\mathbb{1}_A,$$

а что в свою очередь является определением. ■

Теорема 42.1.

1. Если ξ, η независимы, то $\mathbb{E}(\xi|\eta) = \mathbb{E}\xi$
2. Если η измерима относительно $\mathcal{A}$, то $\mathbb{E}(\xi\eta|\mathcal{A}) = \eta\mathbb{E}(\xi|\mathcal{A})$

Доказательство.

1. Проверим, что $\mathbb{E}\xi$ подходит, то есть то, что

$$\mathbb{E}(\mathbb{E}\xi\mathbb{1}_A) = \mathbb{E}(\xi\mathbb{1}_A) \quad \forall A \in \sigma(\eta).$$

   $\sigma(\eta)$ натянута на $\{\eta \leq c\}$, достаточно проверить для $A = \{\eta \leq c\}$. Надо показать, что

$$\mathbb{E}(\mathbb{E}\xi\mathbb{1}_A) = \mathbb{E}\xi\mathbb{E}\mathbb{1}_A = \mathbb{E}(\xi\mathbb{1}_A)$$

   Первое верно, так как $\mathbb{1}_A$ — константа, а второе верно, если ξ и $\mathbb{1}_A$ независимы, то есть когда события $\{\xi \leq a\}$ и $\{\mathbb{1}_A \leq b\}$ независимы (заметим, что это верно, поскольку $\{\mathbb{1}_A \leq c\} = \varnothing$ или $A = \{\eta \leq c\}$).

2. Докажем по стандартной схеме(проверяем для индикаторной функции, по линейности верно для простых, приближаем произвольную простыми и переход к пределу по теореме Леви). Проверяем для $\eta = \mathbb{1}_B$ при $B \in \mathcal{A}$. Надо проверить, что

$$\mathbb{E}(\xi\mathbb{1}_B|\mathcal{A}) = \mathbb{1}_B\mathbb{E}(\xi|\mathcal{A}),$$

   то есть то, что $\mathbb{1}_B\mathbb{E}(\xi|\mathcal{A})$ подходит, что равносильно тому, что для любого $A \in \mathcal{A}$ у нас есть равенство

$$\mathbb{E}(\mathbb{1}_B\mathbb{E}(\xi|\mathcal{A})\mathbb{1}_A) = \mathbb{E}(\xi\mathbb{1}_B\mathbb{1}_A).$$

   Оно есть, так как

$$\mathbb{E}(\mathbb{1}_B\mathbb{E}(\xi|\mathcal{A})\mathbb{1}_A) = \mathbb{E}(\mathbb{E}(\xi|\mathcal{A})\mathbb{1}_{A\cap B}) = \mathbb{E}(\xi\mathbb{1}_{A\cap B}) = \mathbb{E}(\xi\mathbb{1}_B\mathbb{1}_A).$$

■

§43. Ветвящиеся процессы. Вероятность вырождения

Модель следующая. В начальный момент времени есть одна одна частица. Далее на каждом шаге каждая частица с вероятностью f_k делится на k частиц, причём

$\sum_{k=0}^{+\infty} f_k = 1$. Обозначим за η_i количество частиц на i-м шаге, а $\xi_j^(i)$ — количество потомков j-ой частицы (на каждом шаге своя нумерация) на i-м шаге. Тогда

$$\eta_0 = 1,$$
$$\eta_1 = \xi_1^{(1)},$$
$$\eta_2 = \xi_1^{(2)} + \ldots + \xi_{\eta_1}^{(2)}, \ldots,$$

причем $P(\xi_j^{(n)} = k) = f_k$.

Пусть $G_n(z)$ — производящая функция для η_n, $G(z)$ — производящая функция для $\xi_i^{(k)}$. Из написанного выше $G_1 = G$.

$$G_n(z) = \mathbb{E}z^{\eta_n} = \sum_{k=0}^{+\infty} P(\eta_{n-1} = k)\mathbb{E}z^{\xi_1^{(n)}+\ldots+\xi_k^{(n)}} = \sum_{k=0}^{+\infty} P(\eta_{n-1} = k)G^k(z) = G_{n-1}(G(z)).$$

Поэтому $G_n(z) = \underbrace{G(G(\ldots(G}_{n \text{ раз}}(z))))$.

Таким образом, поняли как устроена производящая функция, теперь посчитаем матожидание числа частиц:

$$\mathbb{E}\eta_n = G_n'(1) = G_{n-1}'(G(1)) \cdot G'(1) = G_{n-1}'(1) \cdot G'(1) = [\text{ по индукции }] = (\mathbb{E}\eta_1)^n.$$

Теорема 43.1. Вероятность вырождения процесса — наименьший неотрицательный корень уравнения $G(x) = x$.

Заметим, что

$$G(x) = \sum_{k=0}^{+\infty} f_k x^k;$$
$$G'(x) = \sum_{k=1}^{+\infty} k f_k x^{k-1} \geq 0 \quad \text{при} \quad x \in [0,1];$$
$$G''(x) = \sum_{k=2}^{+\infty} k(k-1) f_k x^{k-2} \geq 0 \quad \text{при} \quad x \in [0,1].$$

Таким образом, функция монотонна и выпукла.

Доказательство. Обозначим $A_n = \{\eta_n = 0\}$, очевидно $A_n \subset A_{n+1}$. Также понятно, что $P(A_n) = G_n(0) \leq 1$. Раз события вложены, то вероятности неубывают и эти вероятности само собой ограничены, а значит есть предел, обозначим его за $q = \lim P(A_n) = \lim G_n(0)$.

С одной стороны,

$$G_{n+1}(0) \to q.$$

С другой, как мы выяснили выше,

$$G_{n+1}(0) = G(G_n(0)) \to G(q).$$

Получается $q = G(q)$, то есть найденный предел — корень уравнения $G(x) = x$.

Осталось доказать, что q — наименьший корень. Пусть y — наименьший неотрицательный корень. Тогда докажем, что $G_n(0) \le y$ всегда, тогда в пределе получим, что $q \le y$, доказав то, что нужно. По условию $0 \le y$, значит

$$G(0) \le G(y) = y \implies G_2(0) \le G(y) = y$$

и так до n, получаем $G_n(0) \le y$. ∎

§44. Скорость вырождения ветвящегося процесса в критическом случае

Теорема 44.1. Если $m = \mathbb{E}\eta_1 = 1$, $0 < b = \mathbb{D}\eta_1$, q_n — вероятность вырождения к n-му шагу, $\gamma_n = q_{n+1} - q_n$ — вероятность вырождения точно на n-м шаге. Тогда

$$\gamma_n \sim \frac{2}{bn^2} \quad \text{и} \quad 1 - q_n \sim \frac{2}{bn}.$$

Доказательство. Заведем вспомогательную функцию $g(x) = 1 - G(1-x)$. $g(0) = 0$ (так как $G'(1) = 1$)

$$g'(x) = G'(1-x);$$
$$g'(0) = m = 1;$$
$$g''(x) = -G''(1-x);$$
$$g''(0) = -G''(1) = -b,$$

поскольку

$$b = \mathbb{D}\eta_1 = G''(1) + G'(1) - (G'(1))^2 = G''(1).$$
$$g(x) = x - \frac{bx^2}{2} + o(x^2);$$
$$p_n = 1 - q_n;$$
$$g(p_n) = p_{n+1};$$
$$\gamma_n = q_{n+1} - q_n = p_n - p_{n+1};$$

. Пусть $a_n = \frac{1}{p_n}$. Тогда

$$a_{n+1} - a_n = \frac{1}{p_{n+1}} - \frac{1}{p_n} = \frac{p_n - p_{n+1}}{p_{n+1}p_n} = \frac{p_n - g(p_n)}{p_n g(p_n)} = \frac{\frac{b}{2}p_n^2 + o(p_n^2)}{p_n(p_n - \frac{bp_n^2}{2} + o(p_n^2))} \sim \frac{b}{2}.$$

И тогда по теореме Штольца.

$$a_n \sim \frac{bn}{2} \implies p_n \sim \frac{2}{bn}.$$

$$\gamma_n = p_n - p_{n+1} = p_n p_{n+1}(a_{n+1} - a_n) \sim p_n^2 \frac{b}{2} \sim \left(\frac{2}{bn}\right)^2 \frac{b}{2} = \frac{2}{bn^2}.$$

■

§45. Марковские цепи. Примеры. Вероятность фиксированной траектории. Теорема существования (без доказательства)

Определение. Пусть Y — конечное или счетное множество (так называемое *фазовое пространство*, *пространство состояний*). $(\Omega, \mathcal{F}, P)$ — вероятностное пространство, $\xi_0, \xi_1, \ldots : \Omega \implies Y$ — случайные величины. Для любого n выполнялось

$$P(\xi_n = a_n | \xi_{n-1} = a_{n-1}, \ldots, \xi_0 = a_0) = P(\xi_n = a_n | \xi_{n-1} = a_{n-1})$$

при $P(\xi_{n-1} = a_{n-1}, \ldots, \xi_0 = a_0) > 0$. Тогда последовательность случайных величин $\xi_0, \xi_1, \ldots$ — *цепь Маркова*.

Примеры 45.1.

1. Случайное блуждание на $\mathbb{Z}$. Пусть η_k — независимые случайные величины, $\eta_k = 1$ с вероятностью p и -1 с вероятностью $1-p$. $\xi_n = \eta_1 + \ldots + \eta_n$. Очевидно, что если мы стоим в какой-то позиции ξ_{n-1}, то значение ξ_n будет зависеть только от ξ_{n-1}, поэтому это цепь Маркова. Итого $\xi_n = \xi_{n-1} + \eta_n$

   Отметим, что не всякая последовательность реализуется (например, нельзя за четное число шагов попасть в нечетную позицию и наоборот).

2. Есть прибор, у которого 2 состояния — сломан и работает. Если он исправен, то через фиксированный квант времени с вероятностью p он ломается, а с вероятностью $1 - p$ остаётся исправным. Если же сломан, то с вероятностью q он становится исправным и с вероятностью $1 - q$ не меняет своего состояния. Это тоже цепь Маркова, так как состояние зависит только от предыдущего шага, а то, что было до этого, не важно.

Определение. Функция $\pi\colon Y \to [0,1]$ — распределение на Y, если $\sum_{y \in Y} \pi(y) = 1$.

Заметим, что цепь Маркова определяется двумя величинами — начальным распределением ($\pi_0 = P_{\xi_0}$ — вероятность на Y) и функцией перехода ($p_n(a,b) = P(\xi_n = b | \xi_{n-1} = a)$).

Определение. Назовем цепь Маркова *однородной*, если $p_n(a,b) = p_{ab}$, то есть не зависит от n.

Нетрудно заметить, что для случайного блуждания по $\mathbb{Z}$ цепь однородна, $p_{k,k+1} = p$ и $p_{k,k-1} = 1 - p$.

Теорема 45.1.

$$P(\xi_0 = a_0, \ldots, \xi_n = a_n) = \pi_0(a_0) \cdot p_{a_0,a_1} \cdot \ldots \cdot p_{a_{n-1},a_n}.$$

(эта последовательность называется *траекторией*).

Доказательство. Индукция по n. База $n = 0$ — определение. Переход $n - 1 \implies n$:

$$\begin{aligned} P(\xi_0 = a_0, \ldots, \xi_n = a_n) &= P(\xi_0 = a_0, \ldots, \xi_{n-1} = a_{n-1}) \cdot P(\xi_n = a_n | \xi_0 = a_0, \ldots, \xi_{n-1} = a_{n-1}) \\ &= P(\xi_0 = a_0, \ldots, \xi_{n-1} = a_{n-1}) \cdot P(\xi_n = a_n | \xi_{n-1} = a_{n-1}) \\ &= \pi_0(a_0) \cdot p_{a_0,a_1} \cdot \ldots \cdot p_{a_{n-1},a_n}. \end{aligned}$$

■

Теорема 45.2. Если заданы $\pi_0 \colon Y \implies [0,1]$ и $p \colon Y \times Y \implies [0,1]$, такие что $\sum_{y \in Y} \pi_0(y) = 1$ и $\sum_{y \in Y} p_{xy} = 1$, то существует такое вероятностное пространство $(\Omega, \mathcal{F}, P)$ и цепь Маркова с начальным распределением π_0 и вероятностные переходы p.

§46. Распределение положений на n-м шаге. Стационарное распределение. Пример

Теорема 46.1. Пусть $\pi_n = P_{\xi_n}$, то есть распределение после n шагов. Его можно представить как вектор длины $|Y|$. P — матрица переходов, то есть матрица $|Y| \times |Y|$, элемент с координатами (a, b) которой равен p_{ab}. Тогда $\pi_n = \pi_0 P^n$.

Доказательство. База $n = 0$ очевидна, матрица в нулевой степени — единичная, получаем $\pi_0 = \pi_0$. Переход $n - 1 \implies n$: надо проверить, что $\pi_n = \pi_{n-1} P$. Рассмотри элемент этого вектора, соответствующий значению $a \in Y$:

$$P(\xi_n = a) = \sum_{y \in Y} P(\xi_{n-1} = y) P(\xi_n = a | \xi_{n-1} = y) = \sum_{y \in Y} \pi_{n-1}(y) p_{ya}.$$

Последнее — это произведение π_{n-1} и столбца P, соответствующего a. ■

Определение. Распределение π называется *стационарным*, если $\pi P = \pi$.

Из предыдущей теоремы становится ясно, что стационарное распределение — то, которое не меняется со временем.

Обозначение. Вероятность перехода из a в b за n шагов

$$p_{ab}(n) = P(\xi_n = b | \xi_0 = a).$$

Пример 46.1. Рассмотрим случайное блуждание на $\mathbb{Z}$. Выбираем сторону с вероятностью $\frac{1}{2}$. Пусть $\pi(y)$ — стационарное распределение, тогда

$$\frac{1}{2}\pi(y-1)+\frac{1}{2}\pi(y+1) = \pi(y)\alpha\pi(y+1)-\pi(y) = \pi(y)-\pi(y-1) \implies \pi(y+1)-\pi(y) = \mathrm{const}.$$

Эта константа не может не равняться нулю, так как иначе через некоторое количество шагов вероятность станет больше 1 или меньше 0, а такого быть не может. Значит, $\pi(y) = \pi(y+1)$, то есть вероятность оказаться в точке для каждой точки одинакова, но такого тоже не бывает, поскольку мы знаем, что $\sum_{y\in Y}\pi(y) = 1$ и поэтому $\pi(y)$ не может не равняться нулю. Итого поняли, что случайное блуждание не имеет стационарного распределения.

Теорема 46.2 (Маркова). Пусть пространство состояний Y конечно и $p_{ab} > 0$ для всех $a, b \in Y$, тогда существует единственное π — стационарное распределение, причём $\pi(b) = \lim_{n\to+\infty} p_{ab}(n)$. Более того, существуют $c > 0$ и $\lambda \in (0,1)$ такие, что

$$|\pi(b) - p_{ab}(n)| \le c\lambda^n \quad \forall a, b \in Y.$$

Стоит отметить, что условие не зависит ни от начального распределения, ни начальной позиции.

Доказательство. Вспомним про теорему Банаха о сжатии: если есть (X, ρ) – полное метрическое пространство и $T : X \to X$ – сжимающее отображение с коэффициентом $\lambda \in (0,1)$ (т.е. это такое отображение, что $\rho(T(x), T(y)) \le \lambda \cdot \rho(x,y)$), тогда существует единственная неподвижная точка (т.е. такой x', что $T(x') = x'$). Сведем нашу теорему к этой.

В качестве полного метрического пространства возьмем $\mathbb{R}^d$, где d– количество элементов в Y, а в качетве нормы: $\| x \| := |x_1| + \cdots + |x_d|$. Рассмотрим $X = \{x \in \mathbb{R}^d : \| x \| = 1, x_i \ge 0\}$ – это соответствует всем распределениям. В качестве отображения логично взять умножение на матрицу перехода $T(x) := xP$. Проверим, что оно сжимающее.

Пусть $z := y - x$, тогда нам надо проверить, что:

$$\| T(y) - T(x) \| \le \lambda \| y - x \|$$
$$\| T(z) \| \le \lambda \| x \|$$

Важно, что у z сумма координат 0. Пусть $\delta := \min\limits_{a,b,\in Y} p_{ab} > 0$. Оценим $\| T(z) \|$:

$$\begin{aligned}
\| T(z) \| &= \sum_{k=1}^{d} |(T(z))_k| \\
&= \sum_{k=1}^{d} |\sum_{j=1}^{d} z_j p_{jk}| (\text{вспомним что } z_1 + \cdots + z_d = 0) \\
&= \sum_{k=1}^{d} |\sum_{j=1}^{d} z_j (p_{jk} - \delta)|
\end{aligned}$$

$$\begin{aligned}
&\le \sum_{k=1}^{d}\sum_{j=1}^{d}|z_j|(p_{jk}-\delta)\\
&= \sum_{j=1}^{d}|z_j|\sum_{k=1}^{d}(p_{jk}-\delta)(\text{вспомним что } p_{1k}+\dots+p_{kk}=1)\\
&= (1-d\delta)\sum_{j=1}^{d}|z_j| = (1-d\delta)\parallel z\parallel
\end{aligned}$$

Осталось сказать, что $(1-d\delta)$ и есть λ из теоремы Банаха о сжатии. Скорость сходимости тоже следует из теоремы Банаха. ■

Следствие 46.3. Пусть Y — конечное множество, и для некоторого n выполняется

$$p_{ab}(n) > 0 \quad \forall a, b \in Y.$$

Тогда существует единственное стационарное распределение π, такое что

$$\lim_{m\to+\infty} p_{ab}(m) = \pi(b) \quad \forall a, b \in Y.$$

§47. Эргодическая теорема Маркова

§48. Классификация состояний цепи Маркова. Критерий возвратности. Теорема солидарности

Определение. Состояние b *достижимо* из состояния a, если $p_{ab}(n) > 0$ для некоторого n.

Два состояния называются *сообщающимися*, если a достижимо из b, а b достижимо из a.

Состояние a — *существенное*, если для любого b, достижимого из a, следует, что a достижимо из b.

Упражнение. Докажите, что в любой конечной цепи существует хотя бы одно существенное состояние.

Обозначим

$$f_a(n) = P(\xi_n = a, \xi_{n-1}\ne a, \dots, \xi_1 \ne a | \xi_0 = a),$$

то есть вероятность того, что мы впервые вернёмся в a за n шагов.

Определение. Если $\sum_{n=1}^{\infty} f_a(n) = 1$, то a — *возвратное* состояние.

Если $\sum_{n=1}^{\infty} f_a(n) < 1$, то a — *невозвратное* состояние.

Определение. Если $p_{aa}(n) \to 0$ при $n\to\infty$, то a — *нулевое* состояние.

Введём обозначение

$$F_a = \sum_{n=1}^{\infty} f_a(n).$$

Теорема 48.1 (критерий возвратности). Состояние a — возвратно тогда и только тогда, когда $P_a = +\infty$, где
$$P_a = \sum_{n=1}^{\infty} p_{aa}(n),$$
а если a невозвратно, то
$$F_a = \frac{P_a}{1 + P_a}.$$

Доказательство. Считаем, что $f_a(0) = 0$ и $p_{aa}(0) = 1$. Заведём производящая функции

$$\mathcal{P}(z) = \sum_{n=0}^{\infty} p_{aa}(n) z^n;$$
$$\mathcal{F}(z) = \sum_{n=0}^{\infty} f_a(n) z^n.$$

Тогда нетрудно понять, что $F_a = \mathcal{F}(1)$ и $P_a = \mathcal{P}(1) - 1$.Также поймём, что

$$p_{aa}(n) = \sum_{k=0}^{n} f_a(k) p_{aa}(n-k).$$

Тогда верно

$$\begin{aligned} \mathcal{P}(z) - 1 = \sum_{n=1}^{\infty} p_{aa}(n) z^n &= \sum_{n=1}^{\infty} \sum_{k=1}^{n} f_a(k) p_{aa}(n-k) z^n \\ &= \sum_{k=0}^{\infty} f_a(k) z^k \sum_{m=0}^{\infty} p_{aa}(m) z^m = \mathcal{F}(z)\mathcal{P}(z). \end{aligned}$$

Таким образом,
$$\mathcal{P}(z) = 1 + \mathcal{F}(z)\mathcal{P}(z),$$
и, следовательно,
$$\mathcal{F}(z) = \frac{\mathcal{P}(z) - 1}{\mathcal{P}(z)}. \tag{$*$}$$

Если ряд $\sum p_{aa}$ сходится, то по теореме Абеля

$$\lim_{\to 1-} \mathcal{P}(z) = P_a.$$

Если ряд $\sum p_{aa} = +\infty$, то
$$\lim_{z \to 1-} \mathcal{P}(z) = P_a = +\infty$$
(упражнение). Таким образом, равенство есть всегда, и можем перейти к пределу в ($*$):
$$F_a = \lim_{z \to 1-} \mathcal{F}(z) = \lim_{z \to 1-} \frac{\mathcal{P}(z) - 1}{\mathcal{P}(z)} = \frac{P_a}{P_a + 1} \begin{cases} < 1, & P_a < +\infty; \\ = 1, & \text{иначе}. \end{cases}$$

■

Следствие 48.2. Невозвратное состояние является нулевым.

Доказательство. Если a невозвратно, то ряд $\sum p_{aa}$ сходится, а значит $p_{aa}(h) \to 0$ при $h \to \infty$, то есть a — нулевое. ■

Теорема 48.3 (солидарности). Сообщающиеся состояния возвратны/невозвратны (нулевые/ненулевые) одновременно.

Доказательство. Пусть a, b — сообщающиеся состояния, то есть по определению $p_{ab}(i) > 0$, $p_{ba}(j) > 0$ для некоторых i, j. Тогда

$$p_{bb}(i+j+k) \geq p_{ba}(j)p_{aa}(k)p_{ab}(i).$$

Если b — нулевое, то

$$p_{bb}(i+j+k) \underset{k\to+\infty}{\to} 0 \implies p_{aa}(k) \implies 0,$$

то есть a — тоже нулевое.

Если a — возвратное, то

$$\sum_{k=1}^{+\infty} p_{aa}(k) = +\infty \implies +\infty = \sum_{k=1}^{+\infty} p_{ba}(j)p_{aa}(k)p_{ab}(i) \leq \sum_{k=1}^{+\infty} p_{bb}(i+j+k),$$

то есть b — возвратное по критерию. ■

Пример 48.1 (управление запасами). Максимальное количество товаров на складе равно s. Если на складе $\leq s$, то заказываем до максимума. Спрос в n-й момент времени η_n — независимо одинаково распределённые случайные величины. Тогда последовательность величин

$$\xi_{n+1} = \begin{cases} \xi_n - \eta_{n+1}, \xi_n > s; \\ S - \eta_{n+1}, \text{ иначе.} \end{cases}$$

будет цепью Маркова.

§49. Случайные блуждания на Z (в соседние точки и произвольное симметричное)

Теорема 49.1. Случайное блуждание на $\mathbb{Z}$ возвратно тогда и только тогда, когда $p = \frac{1}{2}$ (то есть оно симметрично).

Доказательство. Воспользуемся критерием возвратности.

$$\sum_{n=1}^{+\infty} p_{00}(n) = \sum_{n=1}^{+\infty} p_{00}(2n) = \sum_{n=1}^{+\infty} \binom{2n}{n} p^n(1-p)^n.$$

Знаем формулу(например, как следствие формулы Стирлинга)

$$\binom{2n}{n} p^n (1-p)^n \sim \frac{(4p(1-p))^n}{\sqrt{\pi n}}.$$

Тогда если $p \neq \frac{1}{2}$ $4p(1-p) < 1$, то ряд сходящийся, и по критерию блуждание невозвратное. Если $p = \frac{1}{2}$ $\binom{2n}{n} p^n (1-p)^n \sim \frac{1}{\sqrt{\pi n}}$, то ряд расходится и блуждание возвратное. ■

Теперь опишем произвольное симметричное случайное блуждание (то есть можем перейти не только на соседние клетки) на $\mathbb{Z}$ следующим образом. $\xi_1, \xi_2, \ldots$ — независимые одинаково распределенные симметричные целочисленные случайные величины. Обозначим $S_n = \xi_1 + \ldots + \xi_n$.

Теорема 49.2. Если ξ_k симметричные и имеют матожидание, то случайное блуждание возвратно.

Теперь рассмотрим случайные блуждания в $\mathbb{Z}^d$, вероятность перехода в каждую сторону равна $\frac{1}{2d}$.

§50. Теорема Пойя о возвращении

Теорема 50.1 (Пойя о возвращении). Такое случайное блуждание на $\mathbb{Z}^d$ возвратно если и только если $d = 1$ или 2.

Доказательство. Для $d = 1$ и $p = \frac{1}{2d} = \frac{1}{2}$ доказали в теореме 49.1.

Докажем для $d = 2$. Пусть $\overrightarrow{\xi_n}$ — случайное блуждание вдоль в $\mathbb{Z}^2$, а η_n и $\widetilde{\eta_n}$ — блуждания вдоль прямых $y = x$ и $y = -x$, они независимы. Тогда

$$P(\overrightarrow{\xi_{2n}} = 0) = P(\eta_{2n} = 0, \widetilde{\eta_{2n}} = 0) = P(\eta_{2n} = 0)P(\widetilde{\eta_{2n}} = 0) = \left(\binom{2n}{n}\frac{1}{4^n}\right)^2 \sim \frac{1}{\pi n},$$

поэтому ряд $\sum P(\overrightarrow{\xi_{2n}} = 0)$ расходится и по критерию блуждание возвратно.

Пусть $d = 3$. Тогда

$$\begin{aligned} p_{00}(2n) &= \sum_{k+j\le n} \binom{2n}{k,k,j,j,n-j-k,n-j-k} \frac{1}{6^{2n}} \\ &= \binom{2n}{n}\frac{1}{6^{2n}} \sum_{k+j\le n}\left(\binom{n}{k,j,n-j-k}\right)^2 \le \left[\sum_{k+j\le n}\binom{n}{k,j,n-j-k} = 3^n\right] \\ &\le \binom{2n}{n}\frac{1}{6^{2n}} 3^n \cdot \max\binom{n}{k,j,n-j-k} \sim \left[\max \sim 3^n \frac{3\sqrt{3}}{2\pi n}\right] \\ &\sim \frac{4^n}{\sqrt{\pi n}} \frac{1}{4^n 3^n 3^n} 3^n 3^n \frac{3\sqrt{3}}{2\pi n} \end{aligned}$$

$$= \frac{3\sqrt{3}}{2}\frac{1}{(\pi n)^{\frac{3}{2}}}$$

Итого ряд сходящийся и возвратности нет.

Пусть $d \geq 4$. Делаем проекцию с $\mathbb{Z}^d$ на $\mathbb{Z}^3$ и она будет невозвратной, но тогда и исходная очевидно тоже, иначе противоречие с невозвратностью для проекции. ■

§51. Задача о разорении

Пример 51.1 (задача о разорении). Два игрока, у которых A и B монет соответственно играют в орлянку, с вероятностью $q = 1 - p$ первый платит второму и с вероятностью p — второй первому. Найдем вероятность разорения, обозначим за $\beta_k(x)$ вероятность на k-ом шаге оказаться в B, если на нулевом шаге мы находимся в точке x, $-A < x < B$.

$$\beta_k(x) = p\beta_{k-1}(x+1) + q\beta_{k-1}(x-1);$$
$$\beta_{k-1}(x) \leq \beta_k(x) \leq 1.$$

Тогда существует $\lim_{k\to\infty} \beta_k(x)$ обозначим его за $\beta(x) \leq 1$. Получили

$$\beta(x) = p\beta(x-1) + q\beta(x+1);$$
$$\beta(-A) = 0,\ \beta(B) = 1.$$

Нужно решить это соотношение. Если $p \neq q$, то $pt^2 - t + q = 0$, $t_1 = 1, t_2 = \frac{q}{p}$ — его корни.

Тогда $at_1^n + bt_2^n$ — решение соотношения, потому что

$$pt_1^n = pt_1^2 \cdot t_1^{n-2} = (t_1 - q)t_1^{n-2} = t_1^{n-1} - qt_1^{n-2}.$$

Таким образом, $\beta(x) = a + b(\frac{q}{p})^x$, осталось подобрать a, b так, чтобы совпали значения в $-A, B$. Итого

$$\beta(x) = \frac{(\frac{q}{p})^x - (\frac{q}{p})^{-A}}{(\frac{q}{p})^B - (\frac{q}{p})^{-A}}.$$