Теория вероятности

Храбров Александр Игоревич

18 мая 2023 г.

Содержание

1. Элементарная теория вероятностей

1.1. Основные понятия

***Определение* 1.1.** $\Omega = \{\omega_1, \dots, \omega_n\}$ – пространство элементарных событий (исходов).

1. равновозможные
2. несовместные
3. одно всегда реализуется

***Определение* 1.2.** Событие $A \subset \Omega$

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

Свойства. вероятности

1. $P(\varnothing) = 0,\ P(\Omega) = 1,\ P(A) \in [0, 1]$
2. Если $A \cap B = \varnothing$, то $P(A \cup B) = P(A) + P(B)$
3. $\underbrace{P(A \cup B)}_{=P(A)+P(B\setminus(A\cap B))} = P(A) + P(B) - P(A \cap B)$
4. $P(\overline{A}) = 1 - P(A)$, где $\overline{A} = \Omega \setminus A$
5. $P(A_1 \cup A_2 \dots \cup A_m) = \sum_{i=1}^m P(A_i) - \sum_{i\neq j} P(A_i \cap A_j) + \sum_{i\neq j,\ i\neq k,\ j\neq k} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{m-1} \cdot P(A_1 \cap \dots \cap A_m)$ – формула включений-исключений.

   Доказательство. Индукция по m.

   База $m = 2$.

   Переход $m \to m + 1$:

   $B_j = A_j \cup A_{m+1}$

   $P(\underbrace{A_1 \cup \dots \cup A_m}_{=:B} \cup A_{m+1}) = P(B \cup A_{m+1}) = \underbrace{P(B)}_{\text{это умеем расписывать по инд. предп.}} + P(A_{m+1}) - P(B \cap A_{m+1}) =$

   $= \sum_{j=1}^{m+1} P(A_j) - \sum_{i\neq j}^m P(A_i \cap A_j) + \sum_{i\neq j\neq k}^m P(A_i \cap A_j \cap A_k) - \underbrace{P(A_{m+1} \cap B)}_{=P(B_1\cup B_2 \cdots \cup B_m)}$, где $B_i := A_i \cap A_{m+1}$. □

6. $P(A \cup B) \leq P(A) + P(B)$

   $P(A_1 \cup \dots \cup A_m) \leq \sum_{j=1}^m P(A_j)$

***Определение* 1.3. Условная вероятность.**

$B \neq \varnothing,\ P(B) > 0$.

Знаем, что выполнилось событие B, хотим узнать вероятность наступления A.

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

Свойства.

1. $P(A|A) = 1$, если $B \subset A$, то $P(A|B) = 1$
2. Если $A_1 \cap A_2 = \varnothing$, то $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B)$

   В частности: $P(A|B) + P(\overline{A}|B) = 1$

`Замечание.` $P(A|B) + P(A|\overline{B})$ не обязана быть 1.

Пример: игральный кубик, B – выпало четное число, A – выпало кратное трем.

$P(A|B) = \frac{1}{3},\ P(A|\overline{B}) = \frac{1}{3}$

Теорема 1.1. Формула полной вероятности.

Пусть $\Omega = \bigsqcup_{j=1}^{m} B_j,\ P(B_j) > 0.$

Тогда $P(A) = \sum_{j=1}^{m} P(A|B_j) \cdot P(B_j)$

Доказательство. $\sum_{j=1}^{m} \underbrace{P(A|B_j)}_{=\frac{P(A\cap B_j)}{P(B_j)}} \cdot P(B_j) = \sum_{j=1}^{m} P(A \cap B_j) = P(A \cap \bigsqcup_{j=1}^{m} B_j) = P(A)$ □

Пример. Пусть есть 2 урны с шариками:

I. 3 белых шара, 5 черных шаров

II. 5 белых, 5 черных

2 шара из I положили в II, затем вынули 1 шар из II, P(вынули белый) =?

A – вынули из II белый шар.

$B_0,\ B_1,\ B_2$, где B_j – переложили j белых шаров из I в II.

Тогда $P(A|B_0) = \frac{5}{12},\ P(A|B_1) = \frac{1}{2},\ P(A|B_2) = \frac{7}{12}$.

$P(B_0) = \frac{C_5^2}{C_8^2} = \frac{5}{14}$

$P(B_1) = \frac{15}{C_8^2} = \frac{15}{28}$

$P(B_2) = \frac{C_3^2}{C_8^2} = \frac{3}{28}$

Подставляем в формулу:

$P(A) = \frac{5}{12} \cdot \frac{5}{14} + \frac{1}{2} \cdot \frac{15}{28} + \frac{7}{12} \cdot \frac{3}{28} = \frac{23}{48}$

Теорема 1.2. Формула Байеса.

Пусть $P(A) > 0,\ P(B) > 0$, тогда $P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$

Доказательство. Расписываем $P(A|B)$, получаем в правой части: $\frac{P(A\cap B)}{P(B)} \cdot P(B) \cdot \frac{1}{P(A)}$. □

Теорема 1.3. Байеса.

Пусть $P(A) > 0,\ P(B_j) > 0,\ \Omega = \bigsqcup_{j=1}^{m} B_j$, тогда

$P(B_j|A) = \frac{P(A|B_j) \cdot P(B_j)}{P(A|B_1)P(B_1) + \cdots + P(A|B_m)P(B_m)}$

Пример. Есть 2 монеты (одна симметричная, вторая P(орла) $= \frac{1}{3}$, P(решка) $= \frac{2}{3}$). Взялу наугад монету, побросили и выпал орел. Какова вероятность, что мы взяли симметричную монету?

A – выпал орел, B – монета симметричная ($\overline{B}$ – монета кривая).

$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\overline{B})P(\overline{B})} = \frac{\frac{1}{2} \cdot \frac{1}{2}}{\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2}} = \frac{3}{5}$

***Определение* 1.4. Независимые события**.

Рассуждения: A не зависит от B, если $P(A) = P(A|B) = \frac{P(A\cap B)}{P(B)}$.

Def: $A,\ B$ независимые события, если $P(A \cap B) = P(A) \cdot P(B)$

***Определение* 1.5.** События $A_1, A_2, \dots, A_m$ – независимы в совокупности, если

$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_k})$ – для любых индексов i_j.

`Замечание.` Независимость в совокупности $\implies$ попарная независимость.

Наоборот неверно.

Пример. Есть два игральных кубика.

A – на первом кубике выпало четное число.

B – на втором выпало четное число.

C – сумма на кубиках четная.

Пространство элементарных исходов это все пары (i, j), где $i, j \in \{1, 2, 3, 4, 5, 6\}$, $\#\Omega = 36$.

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

$A \cap B = A \cap C = B \cap C = A \cap B \cap C$.

$P(A \cap B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(A) \cdot P(B)$, остальные равенства тоже выполняются $\implies$ попарная независимость.

$P(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = P(A) \cdot P(B) \cdot P(C) \implies$ нет независимости в совокупности.

Упражнение. Д-ть, что $A_1, \dots A_m$ независимы в совокупности $\Leftrightarrow$ $P(B_1 \cap B_2 \cap \dots \cap B_m) = P(B_1) \dots P(B_m)$, где $B_j = A_j$ или $\overline{A_j}$ (все 2^m равенств).

`Замечание.` Небольшое обобщение.

$\Omega = \{\omega_1, \dots, \omega_n\}$ – пр-во элементарных исходов.

Также у нас есть $p_1, \dots p_n :\ \sum_{i=1}^{n} p_i = 1,\ \forall i:\ p_i \geq 0$.

$P(A) = \sum_{j:\ \omega_j \in A} p_j$.

Теорема 1.4. Схема Бернулли.

орел = успех = 1.

решка = неудача = 0.

$P(\text{орел}) = p,\ 0 \leq p \leq 1$

$P(\text{решка}) = 1 - p$

Бросаем монету n раз, получаем последовательность исходов:

$\Omega = \{x_1, x_2, \dots, x_n\}:\ x_j = 0$ или 1.

$\omega = (x_1, x_2, \dots, x_n),\ P(\{\omega\}) = p^{\#i:\ x_i=1} \cdot q^{\#i:\ x_i=0} = p^{\sum x_i} \cdot q^{n - \sum x_i}$

Хотим узнать:

$P(\text{выпало ровно } k \text{ орлов}) = C_n^k p^k q^{n-k}$

$P(i\text{-ое подбрасывание }) = P(\text{орел}) = p$ – независимые в совокупности по $i = 1, 2, \dots, n$.

Теорема 1.5. Полиномиальная схема.

$p_1, p_2, \dots, p_m : \sum p_i = 1$.

$P(x_i = k) = p_k$, где $x_i \in \{1, 2, \dots, m\}$

$\Omega = \{(x_1, x_2, \dots, x_n)\},\ \omega = (x_1, x_2, \dots, x_n)$

$P(\{\omega\}) = p_1^{\#\{i: x_i = 1\}} \cdot \dots \cdot p_m^{\#\{i: x_i = m\}}$

$k_1 + k_2 + \dots + k_m = n$

$$P(k_1 \text{ раз выпало } 1,\ k_2 \text{ раз выпало } 2,\ ...) = \underbrace{\binom{n}{k_1, k_2, \ldots, k_m}}_{=\frac{n!}{k_1!\ldots k_m!}} \cdot p_1^{k_1} \cdot \cdots \cdot p_m^{k_m}$$

Теорема 1.6. Эрдёша-Мозера

Рассмотрим турнир на n команд. При каком наибольшем k можно всегда выбрать команды $A_1, A_2 \ldots A_k$, так, что A_i выиграла у A_j, если $i < j$? При $k \leqslant 1 + [2\log_2 n]$

Доказательство. Предположим, что $k \geqslant 2 + [2\log_2 n] > 1 + 2\log_2 n$. Хотим показать, что при таких k точно найдётся турнир, в котором нельзя выбрать k команд.

Рассмотрим случайный турнир(Всего встреч $\binom{n}{2}$, тогда $2^{\binom{n}{2}}$ разных турниров. Случайный - берём из этой кучи наугад).

$P(\text{А выиграла у В}) = \frac{1}{2}$.

Рассмотрим $A_1, A_2, \ldots A_k$ команды.

1. $P(A_1, A_2 \ldots A_k \text{ подходят}) = (\frac{1}{2})^{\binom{k}{2}}$.
2. $P(A_1, A_2 \ldots A_k \text{ можно переименовать, так, что они подошли}) \leqslant \frac{k!}{2^{\binom{k}{2}}}$
3. $P(\text{какие-то } k \text{ команд подошли}) \leqslant \binom{n}{k} \cdot \frac{k!}{2^{\binom{k}{2}}}$

Нужно понять, что если $k \geqslant 2 + [2\log_2 n]$, то $\binom{n}{k}\frac{k!}{2^{\binom{k}{2}}} < 1$.

Действительно, $\binom{n}{k}\frac{k!}{2^{\binom{k}{2}}} = \frac{n(n-1)(n-2)\ldots(n-k+1)}{2^{\frac{k(k-1)}{2}}} < \frac{n^k}{(2^{\frac{k-1}{2}})^k} = \left(\frac{n}{2^{\frac{k-1}{2}}}\right)^k$

Мы знаем, что $k > 1 + 2\log_2 n \Leftrightarrow \frac{k-1}{2} > \log_2 n \implies 2^{\frac{k-1}{2}} > n$. И тогда $\left(\frac{n}{2^{\frac{k-1}{2}}}\right)^k < 1$. Это значит, что вероятность, что никакие команды не подходят - положительная, значит есть турнир, в котором k команд выбрать нельзя. $\square$

1.2. Предельные теоремы для схем Бернулли

***Определение* 1.6.** Схема Бернулли с вероятностью успеха $p \in (0, 1)$. S_n - число успехов при n испытаниях. $P(S_n = k) = \binom{n}{k}p^k q^{n-k}$

Что будет больше $P(S_{1000} = 220)$ при $p = \frac{1}{5}$ или $P(S_{2000}) = 360$ при $p = \frac{1}{6}$. Точные вычисления дают 0.008984 и 0.006625 соответственно.

Теорема 1.7. Пуассона

Схема Бернулли с n испытаниями и вероятностью успеха p_n - зависит от n. Если $np_n \to \lambda > 0$. Тогда $P(S_n = k) \to \frac{\lambda^k}{k!}e^{-\lambda}$

`Замечание.` Если $np_n = \lambda$, то теорема верна при $k = o(\sqrt{n})$

Доказательство. $P(S_n = k) = \binom{n}{k}p^k(1-p)^{n-k} = \frac{n(n-1)\ldots(n-k+1)}{k!}p^k(1-p)^{n-k} \sim \frac{n^k}{k!}p^k(1-p)^{n-k} = \frac{(np)^k}{k!}(1-p)^{n-k} \sim \frac{\lambda^k}{k!}(1-p)^{n-k}$.

Осталось показать, что $(1-p)^{n-k} \sim e^{-\lambda}$. Прологарифмируем: $\ln(1-p)^{n-k} =$
$= (n-k)\ln(1-p) \sim -np \sim -\lambda$

Доказательство замечания:

Нам нужно показать, что $n(n-1)\dots(n-k+1) \sim n^k$, все остальные переходы будут верны.
$\frac{n(n-1)\dots(n-k+1)}{n^k} = 1 \cdot (1-\frac{1}{n}) \dots \cdot (1-\frac{k-1}{n}) \underbrace{\geqslant}_{(*)} 1 - \frac{1}{n} - \dots - \frac{k-1}{n} = 1 - \frac{k(k-1)}{2n} \to 1$

$(*)$ Неравенство $(1-x_1)\dots(1-x_k) \geqslant 1 - x_1 - x_2 - \dots - x_k$ при $0 \leqslant x_i \leqslant 1$ - индукция. □

Теорема 1.8. Прохорова

Если $\lambda = np$, то $\sum_{k=0}^{+\infty} |P(S_n = k) - \frac{\lambda^k}{k!}e^{-\lambda}| \leqslant \frac{2\lambda}{n} \cdot \min(2, \lambda)$

Пример. Игра в рулетку: 36 чисел и ноль.

$p = \frac{1}{37}, n = 111, np = 3 = \lambda.$

$P(S_{111} = 3) = \binom{111}{3}(\frac{1}{37})^3(1-\frac{1}{37})^{111-3} = 0.227127$

Из Пуассона $\frac{\lambda^3}{3!}e^{-\lambda} = 0.224$

Видим, что приближение хорошее.

P(ставка удачная хотя бы 4 раза) $= 1 - P(S_{111} = 0) - P(S_{111} = 1) - P(S_{111} = 2) - P(S_{111} = 3) = 1 - \frac{\lambda^0}{0!}e^{-\lambda} - \frac{\lambda^1}{1!}e^{-\lambda} - \frac{\lambda^2}{2!}e^{-\lambda} - \frac{\lambda^3}{3!}e^{-\lambda} = 0.352754$

А по формулам 0.352768

Теорема 1.9. Локальная предельная теорема Муавра-Лапласа

Схема Бернулии с вероятностью успеха $p \in (0,1),\ q = 1-p,\ x = \frac{k-np}{\sqrt{npq}}$, где k зависит от n, и n меняется. $|x| \leqslant T$ – при $n \to +\infty$ и любых k. Тогда:

$P(S_n = k) \sim_{n\to+\infty} \frac{1}{\sqrt{2\pi npq}} e^{\frac{-x^2}{2}}$

Доказательство.

1. $k = np + x\sqrt{npq} \geqslant np - T\sqrt{npq} \to +\infty$
2. $n - k = nq - x\sqrt{npq} \geqslant nq - T\sqrt{npq} \to +\infty$

$P(S_n = k) = \binom{n}{k} p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k}$

Напишем формулу Стирлинга ($n! \sim_{n\to\infty} \sqrt{2\pi n} \cdot n^n e^{-n}$):
$\frac{n^n e^{-n} \sqrt{2\pi n} p^k q^{n-k}}{k^k e^{-k} \sqrt{2\pi k} \cdot (n-k)^{n-k} e^{-(n-k)} \sqrt{2\pi(n-k)}} = \frac{p^k q^{n-k}}{(\frac{k}{n})^k \cdot (\frac{n-k}{n})^{n-k} \cdot \sqrt{2\pi \frac{k}{n}(1-\frac{k}{n})n}}.$

Заметим, что $\frac{k}{n} = p + \frac{x\sqrt{pq}}{\sqrt{n}} \to p$
и $\frac{n-k}{n} \to q$

Поэтому остаётся доказать, что $\frac{(\frac{k}{n})^k \cdot (\frac{n-k}{n})^{n-k}}{p^k q^{n-k}} \to e^{\frac{x^2}{2}}$. Прологарифмируем:

Получим: $k \ln \frac{k}{n} + (n-k) \ln \frac{n-k}{n} - k \ln p - (n-k) \ln q \to \frac{x^2}{2}$

Введём обозначения: $\alpha = \frac{k}{n} \to p, \beta = \frac{n-k}{n} \to q$. Тогда $k = n\alpha, n - k = n\beta$ и всё перепишется в виде:

$n\alpha \ln \alpha + n\beta \ln \beta - n\alpha \ln p - n\beta \ln q = \underbrace{n\alpha \ln \frac{\alpha}{p} + n\beta \ln \frac{\beta}{q}}_{(*)} \to \frac{x^2}{2}$

Мы знаем, что $\frac{\alpha}{p} = 1 + x\sqrt{\frac{q}{np}}$ и $\frac{\beta}{q} = 1 - x\sqrt{\frac{p}{nq}}$ - из первых двух тождеств в доказательстве.

Напишем Тейлора:

$ln(1+t) = t - \frac{t^2}{2} + o(t^2)$

$\ln\frac{\alpha}{p} = \ln(1 + x\sqrt{\frac{q}{np}}) = x\sqrt{\frac{q}{np}} - \frac{1}{2}x^2\frac{q}{np} + o(\frac{1}{n})$

$\ln\frac{\beta}{q} = \ln(1 - x\sqrt{\frac{p}{nq}}) = -x\sqrt{\frac{p}{nq}} - \frac{1}{2}x^2\frac{p}{nq} + o(\frac{1}{n})$

Тогда $(*) = x\sqrt{pq}\sqrt{n} + x^2q - \frac{1}{2}x^2q + o(\frac{1}{n}) - x\sqrt{pq}\sqrt{n} + x^2p - \frac{1}{2}x^2p + o(\frac{1}{n}) = x^2(\frac{q}{2} + \frac{p}{2}) + o(1) = \frac{x^2}{2} + o(1)$ □

Замечание. Если $\varphi(n) = o(n^{\frac{2}{3}})$ и $|k - np| \leqslant \varphi(n)$, то теорема тоже верна

Пример. Всё та же рулетка. $n = 222, k = 111$. Пытаемся ставить на четное/нечётное(кроме 0). $p = \frac{18}{37}$

$P(S_{222} = 111) \approx \frac{1}{\sqrt{2\pi npq}}e^{\frac{-x^2}{2}} \approx 0.049395\ldots$

Если считать точно, то получим $0.0493228\ldots$

Теорема 1.10. Интегральная теорема Муавра-Лапласа

$0 < p < 1$. $P(a < \frac{S_n - np}{\sqrt{npq}} \leqslant b) \to_{n\to\infty} \frac{1}{\sqrt{2\pi}}\int_a^b e^{\frac{-t^2}{2}}\,dt$

Стремление равномерно по $a,\ b \in \mathbb{R}$.

Теорема 1.11. Берри-Эссеена

Обозначение: $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x e^{-\frac{t^2}{2}}\,dt$, $\Phi_0(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}\,dt$

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

Замечание. Константа лучше, чем $\frac{c}{\sqrt{n}}$ не бывает.

Замечание. $P(a < S_n \leqslant b) = P(\frac{a-np}{\sqrt{npq}} < \frac{S_n-np}{\sqrt{npq}} \leqslant \frac{b-np}{\sqrt{npq}}) \to \Phi(\frac{b-np}{\sqrt{npq}}) - \Phi(\frac{a-np}{\sqrt{npq}})$

Отсюда получили, что лучше всего писать полуцелые a и b.

Замечание. Если p или q очень маленькие, то произведение np маленькое и оценка будет плохой. В таких случаях хорошо использовать Пуассона. Муавра-Лаплас же хорош, когда np большое.

Пример. $p = q = \frac{1}{2}$. Вопрос: $P(S_{2n} = n) = \binom{2n}{n}\frac{1}{2^{2n}} \sim \frac{4^n}{\sqrt{\pi n}}\frac{1}{4^n} = \frac{1}{\sqrt{\pi n}}$.

Но $P(S_{2n} < n) = P(S_{2n} > n)$.

Тогда $P(S_{2n} \leqslant n) = \frac{1 + P(S_{2n}=n)}{2} = \frac{1}{2} + \frac{1}{2\sqrt{\pi n}} + o(\frac{1}{\sqrt{n}})$

Муавра-Лаплас нам говорит, что $P(S_{2n} \leqslant n) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^0 e^{-\frac{t^2}{2}}\,dt = \frac{1}{2}$

Но $P(S_{2n} \leqslant n) = \frac{1}{2} + \frac{1}{2\sqrt{\pi n}} + o(\frac{1}{\sqrt{n}})$

Пример. Задача о театре

Есть театр и 2 входа. У каждого входа расположен гардероб. В театре $n = 1600$ мест. Хотим сделать размер гардероба как можно меньше, но чтобы переполнения случались как можно реже (не чаще, чем раз в месяц).

Пусть c мест в итоге в гардеробе.

За успех считаем ситуацию, когда человек вошел в театр и пошел в ближайший к нему гардероб (т.е. в ближайшем гардеробе было место, и человек не пошел в дальний гардероб). Пусть S_n – кол-во успешных испытаний.

Так как в каждый гардероб мы допускаем c мест, то кол-во успехов $S_n \leq c$.

$p = q = \frac{1}{2}$. Нужно, чтобы $n - c \leqslant S_n \leqslant c$. И $P(n - c \leqslant S_n \leqslant c) > \frac{29}{30}$ – т.е. хотя бы в 29 днях из 30 ближайший к каждому входу гардероб не переполняется.

$P(n-c \leqslant S_n \leqslant c) = P\left(\frac{n-c-\frac{n}{2}}{\sqrt{n\cdot\frac{1}{4}}} \leqslant \frac{S_n-\frac{n}{2}}{\sqrt{n\cdot\frac{1}{4}}} \leqslant \frac{c-\frac{n}{2}}{\sqrt{n\cdot\frac{1}{4}}}\right) =$

$P\left(\frac{800-c}{20} \leqslant \frac{S_n-800}{20} \leqslant \frac{c-800}{20}\right) \to \Phi(\frac{800-c}{20}) - \Phi(\frac{c-800}{20}) = \frac{1}{\sqrt{2\pi}}\int_{\frac{800-c}{20}}^{\frac{c-800}{20}} e^{-\frac{t^2}{2}}\,dt = 2\cdot\Phi_0(\frac{c-800}{20}) > \frac{29}{30}$

$\Phi_0(\frac{c-800}{20}) > \frac{29}{60} \implies c = 843.$

Пример. Случайное блуждание на прямой

Есть прямая, будем считать, что у нас блуждания исключительно по целым точкам.

В каждой точке подбрасываем монетку. С вероятностью p идём вперёд, q - идём назад.

$a_{n+1} = a_n + 1$ с вероятностью p

$a_{n+1} = a_n - 1$ с вероятностью q

$a_n \equiv n \mod 2$

Если представить, что шаг влево – это 0, а шаг в право – это 1, то сдвиг из точки a_n будет определяться как $2\cdot x - 1$, где $x = 0,\ 1$ в зависимости от того, в какую сторону идем (т.е. $a_{n+1} = a_n + (2\cdot x - 1)$).

Тогда пусть S_n – кол-во единичек, тогда $a_n = 2\cdot S_n - n$.

Пусть мы всегда стартуем с $a_0 = 0$, тогда определим, чему равна вероятность попасть за n шагов в точку k ($a_n = k$):

$$P(a_n = k) = P(S_n = \tfrac{n+k}{2}) = \begin{cases} 0, \text{ если } n \not\equiv k \mod 2 \\ \binom{n}{\frac{n+k}{2}} p^{\frac{n+k}{2}} q^{\frac{n-k}{2}}, \text{ иначе} \end{cases}$$

Теорема 1.12. ван дер Вардена

Рассмотрим числа $1, 2\ldots k$ и покрасим их в 2 цвета.

Тогда существует k_n, такое, что, если $k > k_n$, то при любой раскраске найдётся одноцветная n-членная арифметическая прогрессия.

Теорема 1.13. Эрдеша-Радо

$k_{n+1} \geqslant \sqrt{n\cdot 2^{n+1}}$

Доказательство. $A_1, A_2\ldots A_m$ - все арифметические прогрессии длины $n+1$ из чисел $1, 2\ldots k$.

С разностью $1 : k - n$ прогрессий.

С разностью $2 : k - 2n$ прогрессий.

...

С разностью $[\frac{k}{n}] : k - [\frac{k}{n}]\cdot n$ прогрессий

Тогда $m = (k-n) + (k-2n) + \ldots + k - [\frac{k}{n}]\cdot n = k\cdot[\frac{k}{n}] - n\cdot\frac{[\frac{k}{n}]\cdot([\frac{k}{n}]+1)}{2} = [\frac{k}{n}](k - \frac{1}{2}n([\frac{k}{n}]+1)) < \frac{k}{n}(k - \frac{1}{2}\cdot n\cdot\frac{k}{n}) = \frac{k^2}{2n}$ - это оценка сверху.

$P(A_i$ - одноцветная$) = 2\cdot\frac{1}{2^{n+1}} = \frac{1}{2^n}$ (2 - выбор цвета).

$P($какое-то A_i - одноцветно$) = \sum_{i=1}^{m} P(A_i$ - одноцветно$) = \frac{m}{2^n} < \frac{k^2}{2n}\cdot\frac{1}{2^n} = (\frac{k}{\sqrt{2^{n+1}\cdot n}})^2 \leqslant 1$ (если так, то найдётся, на которой не выполнится) $\square$

2. Общая теория вероятностей

2.1. Колмогоровская модель теории вероятности

***Определение* 2.1.** $(\Omega, \mathcal{F}, P)$ - вероятностное пространство.

Ω - множество или пространство элементарных исходов.

$\mathcal{F}$ - σ-алгебра подмножеств Ω. Элементы $\mathcal{F}$ - случайный события.

P - мера на $\mathcal{F}$ с условием $P(\Omega) = 1$.

`Замечание.` Если Ω не более чем счётно, то можно взять $\mathcal{F} = 2^{\Omega}$

***Определение* 2.2.** Условная вероятность. A - событие, такое, что $P(A) > 0$. Тогда $P(B|A) = \frac{P(B \cap A)}{P(A)}$, где $A, B \in \mathcal{F}$.

***Определение* 2.3.** Независимые события A и B. Если $P(A \cap B) = P(A) \cdot P(B)$

***Определение* 2.4.** Независимость в совокупности $A_1, A_2 \ldots A_n$. $P(A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot \ldots \cdot P(A_{i_k})$ для всевозможных наборов индексов.

***Определение* 2.5.** Последовательность событий $A_1, A_2 \ldots$ независимы - любой конечный набор событий независим в совокупности.

Лемма. Бореля-Кантелли

$A_1, A_2, \ldots$ случайные события.

1. Если $\sum_{n=1}^{\infty} P(A_n) < +\infty$, то вероятность, что случилось бесконечное число из них равна 0.
2. Если $A_1, A_2, \ldots$ независимы и $\sum_{n=1}^{\infty} P(A_n) = +\infty$, тогда
   $P(\text{случилось бесконечное число из } A_n) = 1$.

Доказательство. $B = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$ - это переформулировка события из условия в терминах множеств.

$\omega \in B \iff \omega \in \bigcup_{k=n}^{\infty} A_k \ \forall n \iff w \in A_k$ для бесконечного количества индексов k.

Док-во этого факта:

1. $\Leftarrow$: Лежит в каждом объединении, значит лежит в B.
2. $\Rightarrow$: ω лежит в пересечении. Пусть лежит в конечном - возьмём самый большой номер и получим противоречие.

Док-во теоремы:

1. Хотим доказать, что $P(B) = 0$
   $B \subset \bigcup_{k=n}^{\infty} A_k \Rightarrow P(B) \leqslant P(\bigcup_{k=n}^{\infty} A_k) \leqslant \sum_{k=n}^{\infty} P(A_k)$ - это хвост сходящегося ряда, а он стремится к нулю.
2. Давайте смотреть на $\bar{A}_1, \bar{A}_2, \ldots$ - независимые события.
   $P(\bigcap_{k=1}^{n} \bar{A}_k) \overbrace{=}^{\text{независимость}} \prod_{k=1}^{n} P(\bar{A}_k) \to_{n\to\infty} \prod_{k=1}^{\infty} P(\bar{A}_k) = \prod_{k=1}^{\infty}(1 - P(A_k))$
   А ещё $P(\bigcap_{k=1}^{n} \bar{A}_k) \to P(\bigcap_{k=1}^{\infty} \bar{A}_k)$, так как множества вложены в друг друга и есть монотонность меры.

Значит $P(\bigcap_{k=n}^{\infty}\bar{A}_k) = \prod_{k=n}^{\infty}(1-P(A_k)) \overset{\text{логарифмируем}}{\iff} \ln P(\bigcap_{k=n}^{\infty}\bar{A}_k) =$
$= \sum_{k=n}^{\infty}\ln(1-P(A_k)) \overset{\ln(1-t)\leqslant -t}{\leqslant} \sum_{k=n}^{\infty}(-P(A_k)) = -\infty$ - хвост расходящегося ряда.

А значит мы логарифмировали $0 \Rightarrow P(\bigcap_{k=n}^{\infty}\bar{A}_k) = 0 \Rightarrow P(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\bar{A}_k) = 0 \Rightarrow P(\bar{B}) = 0 \overset{(*)}{\Rightarrow} P(B) = 1$

$(*)\, \overline{\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\bar{A}_k} = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k = B$

□

Теорема 2.1. Закон нуля и единицы

Если $A_1, A_2 \ldots$ независимы, то $P(B) = 0$ или $P(B) = 1$. При $B = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k$.

Пример. Испытания Бернулли, успех с вероятностью p,

P(ОРО встречается бесконечное число раз) = ?.

A_n = случилось ОРО на позициях $n, n+1, n+2$.

Тогда $A_1, A_4, A_7, \ldots$ независимы. $P(A_j) = pqp = p^2q > 0$.

Лемма Бореля-Кантелли говорит: бесконечное кол-во A_{3k+1} случится, если $\sum_{k=1}^{\infty}P(A_{3k+1}) = +\infty \implies P$(ОРО встречается бесконечное число раз) = 1.

2.2. Случайные величины

***Определение* 2.6.** $(\Omega, \mathcal{F}, P)$ - вероятностное пространство.

$\xi : \Omega \to \mathbb{R}$ - случайная величина, если это измеримая функция.

***Определение* 2.7.** Распределение случайное величины

P_ξ - вероятностная мера на борелевских подмножествах $\mathbb{R}$

A – борелевское мн-во, $P_\xi(A) = P(\omega \in \Omega \ : \ \xi(\omega) \in A)$

***Определение* 2.8.** Случаный величины ξ и η одинаково распределены, если $P_\xi = P_\eta$

`Замечание.` P_ξ однозначно определяются своими значениями на ячейках.

$P_\xi(a,b] = P_\xi(-\infty, b] - P_\xi(-\infty, a] = P(\xi \leqslant b) - P(\xi \leqslant a)$

***Определение* 2.9.** Функция распределения случайной величины

$F_\xi(x) = P(\xi \leqslant x)$

Свойства.

1. Функция распределения однозначно определяет распределение случайной величины.

   Доказательство. Функция распределения однозначно задаёт значения на ячейках □

2. $0 \leqslant F_\xi(x) \leqslant 1\, \forall x \in \mathbb{R}$

3. $\lim_{x\to-\infty} F_\xi(x) = 0$

   $\lim_{x\to+\infty} F_\xi(x) = 1$

   Доказательство. берём $x_n \to -\infty, A_n = \{\xi \leqslant x_n\}$ Тогда $A_{n+1} \subset A_n$. Тогда $\lim_{n\to\infty} P(A_n) = P(\bigcap_{n=1}^{\infty} A_n) = P(\varnothing) = 0$ □

4. F_ξ монотонно возрастает

5. Непрерывность справа: $\lim_{y\to x+} F_\xi(y) = F_\xi(x)$

   Доказательство. берём y_n убывающие и $y_n \to x$. Тогда $A_n = \{\xi \leqslant y_n\}$. $A_{n+1} \subset A_n$. А тогда $\lim P(A_n) = P(\bigcap_{n=1}^\infty A_n) = P(\xi \leqslant x) = F_\xi(x)$. Но с другой стороны $\lim P(A_n) = \lim P(\xi \leqslant y_n) = \lim F_\xi(y_n)$ □

6. $\lim_{y\to x-} F_\xi(y) = P(\xi < x)$

   Доказательство. берём y_n возрастающие и $y_n \to x$. $B_n = \{\xi \leqslant y_n\}$ и $B_n \subset B_{n+1}$. $\lim P(B_n) = P(\bigcup B_n) = P(\xi < x)$. Но с другой стороны $\lim P(B_n) = \lim F_\xi(y_n)$ □

7. $F_{\xi+a}(x) = F_\xi(x-a)$

   Доказательство. $\{\xi + a \leqslant x\} = \{\xi \leqslant x - a\}$ □

8. $F_{c\xi} = F_\xi(\frac{x}{c})$

   Доказательство. $\{c\xi \leqslant x\} = \{\xi \leqslant \frac{x}{c}\}$ □

`Замечание.` Фукнция, обладающая свойствами 3, 4, 5 - это фукнция распределения некоторой случайной величины.

Доказательство. пусть g - такая функция. Тогда $\nu_g(a,b] = g(b) - g(a)$. $\Omega = \mathbb{R}$, $\mathcal{F}$ - измеримо по Лебегу, случайная величина $\xi(w) = w$. Тогда $F_\xi = g$ □

***Определение* 2.10.** Случайная величина имеет дискретное распределение, если её множество значений не более чем счётное.

`Замечание.`

1. $\xi \to \{y_1, y_2, \ldots\}$

   Если $x \neq y_k$, то $P(\xi = x) = 0$, т.е. $P_\xi(\{x\}) = 0$

2. $P_\xi(A) = \sum_{k: y_k \in A} P(\xi = y_k)$. Тут счётное число слагаемых, поэтому сумма корректно определена.

   Распределение однозначно определяется набором вероятностей $P(\xi = y_k)$

3. $F_\xi(x) = \sum_{k: y_k \leqslant x} P(\xi = y_k)$

***Определение* 2.11.** Случайная величина имеет непрерывное распределение, если $\forall x \in \mathbb{R} : P(\xi = x) = 0$

`Замечание.`

1. ξ – имеет непрерывное распределение $\iff F_\xi$ непрерывна во всех точках.

   $P(\xi < x) = \lim_{y\to x-} P(\xi \leq y) = \lim_{y\to x-} F_\xi(y)$

   $0 = P(\xi = x) = P(\xi \leq x) - P(\xi < x) = F_\xi(x) - \lim_{y\to x-} F_\xi(y) \Rightarrow F_\xi(x) = \lim_{y\to x-} F_\xi(y)$

2. Непрерывные распределения бывают не очень хорошими, например Канторова лестница.

***Определение* 2.12.** Случайная величина имеет абсолютно непрерывное распределение, если существует $p_\xi(t) \geqslant 0$, измеримая, т.ч. $F_\xi(x) = \int_{-\infty}^x p_\xi(t)\,dt$ ($p_\xi(t)$ – плотность распределения).

Свойства.

1. $A \subset \mathbb{R}$ – борелевское, то $P_\xi(A) = \int_A p_\xi(t)\,dt$

Доказательство. слева мера и справа написаны меры. На лучах они совпадают по определению, значит совпадают на ячейках, а значит и совпадают везде □

$P_\xi(a, b] = F_\xi(b) - F_\xi(a) = \int_a^b p_\xi(t)\,dt$

2. $\int_{-\infty}^{+\infty} p_\xi(t)\,dt = 1$

3. p_ξ определена однозначно с точностью до почти везде (из теории меры)

4. F_ξ почти везде диффиренцируема и $F'_\xi(x) = p_\xi(x)$

   Доказательство. без доказательства □

Пример. Вероятностные распределения

1. Биномиальное распределение: $\xi \sim Binom(p, n), 0 < p < 1$
   $\xi : \Omega \to \{0, 1, \ldots n\}$. $P(\xi = k) = \binom{n}{k} p^k (1-p)^{n-k}$

2. Распределение Пуассона: $\xi \sim Poisson(\lambda), \lambda > 0$.
   $\xi : \Omega \to \{0, 1, \ldots\}$. $P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}$

3. Геометрическое распределение: $\xi \sim Geom(p), 0 < p < 1$.
   $\xi : \Omega \to \{1, 2, \ldots\}$. $P(\xi = k) = p(1-p)^{k-1}$.

4. Дискретное равномерное распределение:
   $\xi : \Omega \to \{1, 2, \ldots n\}$. $P(\xi = k) = \frac{1}{n}$

5. Непрерывно равномерное распределение: $\xi \sim U([a, b])$
   $\xi : \Omega \to [a, b]$. $p_\xi(t) = \frac{1}{b-a} \cdot \mathbb{1}_{[a,b]}(t)$

6. Нормальное распределение: $\xi \sim \mathcal{N}(a, \sigma^2), a \in \mathbb{R}, \sigma > 0$
   $\xi : \Omega \to \mathbb{R}$. $p_\xi(t) = \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{(t-a)^2}{2\sigma^2}}$
   Стандартное нормальное распределение: $\mathcal{N}(0, 1)$

7. Экспонециальное распределение: $\xi \sim Exp(\lambda), \lambda > 0$.
   $\xi : \Omega \to [0, +\infty]$. $p_\xi(t) = \begin{cases} \lambda e^{-\lambda t}, & \text{при } t \geqslant 0 \\ 0, & \text{в других точках} \end{cases}$

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

   На самом деле это функция распределения стандартной нормальной случайной величины.

2. Если $\nu \sim \mathcal{N}(0, 1)$, то $\xi = \sigma\nu + a$. $\xi \sim \mathcal{N}(a, \sigma^2)$
   $F_\xi(x) = P(\sigma\nu + a \leqslant x) = P(\nu \leqslant \frac{x-a}{\sigma}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{t^2}{2}}\,dt$
   Замена $t = \frac{s-a}{\sigma}$. Тогда $dt = \frac{ds}{\sigma}$
   Тогда: $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{t^2}{2}}\,dt = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{(s-a)^2}{2\sigma^2}}\,ds$

2.3. Совместное распределение

***Определение* 2.13.** Совместное (многомерное) распределение.

$\bar{\xi} = (\xi_1, \xi_2, \ldots, \xi_n) : \Omega \to \mathbb{R}^n$

$P_{\bar{\xi}}(A) = P(\bar{\xi} \in A)$, где A - борелевское подмножество $\mathbb{R}^n$

`Замечание.` $P_{\bar{\xi}}$ однозначно определяет распределение P_{ξ_k}, но не наоборот

Пример. $\xi, \eta : \Omega \to \{0, 1\}$ с равными вероятностями.

Если это были независимые подбрасывания: $(\xi, \eta) : \Omega \to \{(0,0), (0,1), (1,0), (1,1)\}$ с равными вероятностями.

Если $\xi = \eta$, то $(\xi, \eta) : \Omega \to \{(0,0), (1,1)\}$.

То есть получили 2 разных совместных распределения, при это координатное распределение только одно

***Определение* 2.14.** Случайные величины $\xi_1, \xi_2 \ldots \xi_n$ независимы, если для любых борелевских подмножеств $A_1, A_2 \ldots A_n \subset \mathbb{R}$, события $\{\xi_1 \in A_1\}, \ldots, \{\xi_n \in A_n\}$ независимы

`Замечание.` $P(\xi_1 \in A_1, \ldots, \xi_n \in A_n) = P(\xi_1 \in A_1) \cdot \ldots \cdot P(\xi_n \in A_n)$

Теорема 2.2. $\xi_1, \xi_2 \ldots \xi_n$ независимы $\Longleftrightarrow P_{\bar{\xi}} = P_{\xi_1} \times \ldots \times P_{\xi_n}$

Доказательство.

1. $\Leftarrow P(\xi_1 \in A_1, \ldots, \xi_n \in A_n) = P_{\xi_1, \ldots, \xi_n}(A_1 \times \ldots \times A_n) = P_{\xi_1}(A_1) \cdot \ldots \cdot P_{\xi_n}(A_n)$
2. $\Rightarrow$. Достаточно проверить совпадение на ячейках, то есть, что $P(\bar{\xi} \in (a, b]) = P_{\xi_1}(a_1, b_1] \cdot \ldots \cdot P_{\xi_n}(a_n, b_n]$. А это просто определение независимости.

□

***Определение* 2.15.** Совместная (многомерная) функция распределения.

$\bar{\xi} = (\xi_1 \ldots \xi_n)$. $F_{\bar{\xi}} : \mathbb{R}^n \to \mathbb{R}$. и $F_{\bar{\xi}}(\bar{x}) = P(\xi_1 \leqslant x_1, \ldots, \xi_n \leqslant x_n)$

Свойства.

1. $0 \leqslant F_{\bar{\xi}} \leqslant 1$
2. Монотонно возрастает по каждой координате
3. $\lim_{x_i \to -\infty} F_{\bar{\xi}}(\bar{x}) = 0$

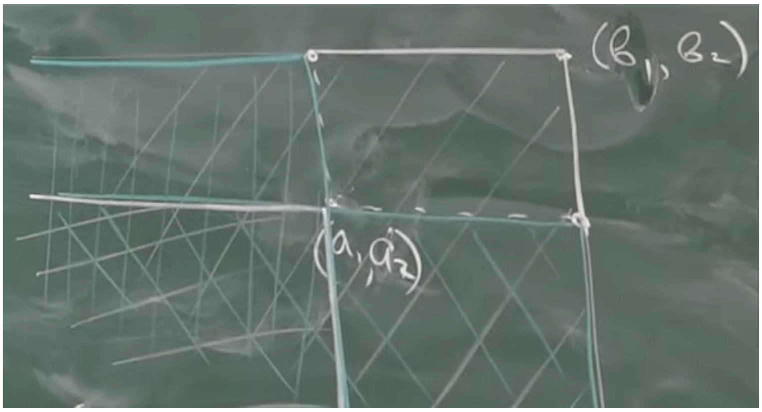
   $\lim_{x_1, \ldots, x_n \to +\infty} F_{\bar{\xi}}(\bar{x}) = 1$
4. $\lim_{x_i \to +\infty} F_{\bar{\xi}}(\bar{x}) = F_{\xi_1, \ldots, \xi_{i-1}, \xi_{i+1}, .., \xi_n}(x_1, .., x_{i-1}, x_{i+1}, .., x_n)$

***Определение* 2.16.** Совместная плотность $p_{\bar{\xi}}(\bar{t})$ - неотрицательная измеримая функция, такая, что $F_{\bar{\xi}}(\bar{\xi}) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} p_{\bar{\xi}}(\bar{t})\, dt_n \ldots dt_1$

Теорема 2.3. $\xi_1 \ldots \xi_n$ независимы $\Longleftrightarrow F_{\bar{\xi}}(\bar{x}) = F_{\xi_1}(x_1) \cdot \ldots \cdot F_{\xi_n}(x_n),\ \forall \bar{x} \in \mathbb{R}^n$

Доказательство.

1. Докажем $\Rightarrow$. Независимость $\Rightarrow (*) P_{\bar{\xi}} = P_{\xi_1} \times \ldots \times P_{\xi_n} \Rightarrow P_{\bar{\xi}}((-\infty, x_1] \times \ldots \times (-\infty, x_n]) = P_{\xi_1}(-\infty, x_1] \cdot \ldots \cdot P_{\xi_n}(-\infty, x_n]$
2. Хотим проверить совпадение на ячейках, чтобы доказать $(*)$ ещё и в другую сторону (приведем выкладки для $n = 2$, для больших n рассуждения не меняются).



$$P_{\bar\xi}((a_1,b_1]\times(a_2,b_2]) = F_{\bar\xi}(b_1,b_2)+F_{\bar\xi}(a_1,a_2)-F_{\bar\xi}(a_1,b_2)-F_{\bar\xi}(a_2,b_1) =$$
$$= (F_{\xi_1}(b_1)-F_{\xi_1}(a_1))\cdot(F_{\xi_2}(b_2)-F_{\xi_2}(a_2)) = P_{\xi_1}(a_1,b_1]\cdot P_{\xi_2}(a_2,b_2]$$

□

Следствие. $\xi_1\dots\xi_n$ - абсолютно непрерывные случайные величины. Тогда $\xi_1\dots\xi_n$ независимы $\iff p_{\bar\xi}(\bar t) = p_{\xi_1}(t_1)\cdot\ldots\cdot p_{\xi_n}(t_n)$

В частности, в случае независимости $\bar\xi$ абсолютно непрерывна.

Доказательство. 1. Докажем $\Rightarrow$.

Независимость $\Rightarrow F_{\bar\xi}(\bar x) = F_{\xi_1}(x_1)\cdot\ldots\cdot F_{\xi_n}(x_n) = \int_{-\infty}^{x_1} p_{\xi_1}(t_1)\,dt_1\cdot\ldots\cdot\int_{-\infty}^{x_n} p_{\xi_n}(t_n)\,dt_n = \int_{-\infty}^{x_1}\ldots\int_{-\infty}^{x_n} p_{\xi_1}(t_1)\ldots p_{\xi_n}(t_n)\,dt_n\ldots dt_1$.

Запихали всё под один интеграл, то что под интегралом и есть совместная плотность.

2. Докажем $\Leftarrow$.

Просто проинтегрируем равенство.

$$\int_{-\infty}^{x_1}\ldots\int_{-\infty}^{x_n} p_{\bar\xi}(\bar t)\,dt_n\ldots dt_1 \underbrace{=}_{\text{по т. Тонелли можно выносить интегралы}} \int_{-\infty}^{x_1}\ldots\int_{-\infty}^{x_n} p_{\xi_1}(t_1)\ldots p_{\xi_n}(t_n)\,dt_n\ldots dt_1 = F_{\xi_1}(x_1)\cdot\ldots\cdot F_{\xi_n}(x_n)$$

т. Тонелли можно использовать, так как мы интегрируем неотрицательную функцию.

□

`Замечание.` Напоминание.

Свертка последовательностей: $\{a_n\},\{b_n\}$ это $\{c_n\}$, такая что $c_n = a_0b_n + a_1b_{n-1}+\ldots+a_nb_0$.

Мотивировка: $(\sum_{n=0}^\infty a_nz^n)\cdot(\sum_{n=0}^\infty b_nz^n) = \sum_{n=0}^\infty c_nz^n$ (при наличии хоть каких-нибудь кругов сходимости у обоих рядов).

`Замечание.` **Свертки мер**

μ и ν - конечные меры на борелевских подмножествах $\mathbb{R}$.

$\mu*\nu(A) = \int_{\mathbb{R}}\mu(A-x)\,d\nu(x)$ - это свертка мер, где $(A-x) := \{a-x \mid a\in A\}$.

Свойства. **Свойства свёртки**

1. $\mu*\nu(A) = \int_{\mathbb{R}^2}\mathbb{1}_A(x+y)d\mu(x)\,d\nu(y)$

   Доказательство. $\mu*\nu(A) = \int_{\mathbb{R}}\mu(A-x)\,d\nu(x) \overset{\mu(A-x)=\int_{\mathbb{R}}\mathbb{1}_{A-x}d\mu(y)}{=} \int_{\mathbb{R}}\int_{\mathbb{R}}\mathbb{1}_{A-x}(y)d\mu(y)\,d\nu(x)$ □

2. $\mu * \nu = \nu * \mu$

3. $\mu_1 * \ldots * \mu_n(A) = \int_{\mathbb{R}^n} \mathbb{1}_A(x_1 + \ldots + x_n)\, d\mu_1(x_1) \ldots\, d\mu_n(x_n)$

4. $(\mu_1 * \mu_2) * \mu_3 = \mu_1 * (\mu_2 * \mu_3)$

5. $(\mu_1 + \mu_2) * \nu = \mu_1 * \nu + \mu_2 * \nu$

6. δ_x - мера с единичной нагрузкой в точке x. Тогда $\mu * \delta_0 = \mu$.

   Получили линейное пространство относительно $+$ и $*$

   Доказательство. $\mu * \delta_0(A) = \delta_0 * \mu(A) = \int_{\mathbb{R}} \delta_0(A - x)\, d\mu(x) \overset{\delta_0 = 1 \Leftrightarrow 0 \in A - x \Leftrightarrow x \in A}{=} \int_{\mathbb{R}} \mathbb{1}_A d\mu(x) = \mu A$ □

Теорема 2.4. Пусть μ и ν имеют плотности p_μ и p_ν

Тогда $\mu * \nu$ имеет плотность $p(t) = \int_{\mathbb{R}} p_\mu(t - s) p_\nu(s)\, ds$

Доказательство. Возьмём функцию, определяемую этой формулой и проверим, что это плотность.

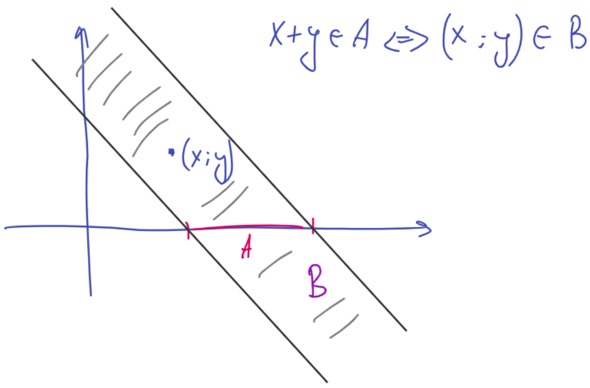
То есть проверим, что $\int_A p(x)dx = \mu * \nu(A)$.

$\int_A p(t)\, dt = \int_A \int_{\mathbb{R}} p_\mu(t - s) p_\nu(s)\, ds\, dt = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_A(t) p_\mu(t - s) p_\nu(s)\, ds\, dt = (*)$.

Положим $u = t - s$. Тогда $(*) = \int_{\mathbb{R}^2} \mathbb{1}_A(u + s) p_\mu(u) p_\nu(s)\, ds\, du = \int_{\mathbb{R}^2} \mathbb{1}_A(u + s)\, d\nu(s)\, d\mu(u) = \mu * \nu(A)$ □

Теорема 2.5. Если ξ и η независимые случайный величины, то $P_{\xi+\eta} = P_\xi * P_\eta$

Доказательство. Нужно взять какое-то борелевское множество и понять как устроено там распределение суммы.



Пусть $B = \{(x, y) : x + y \in A\}$

$P_{\xi+\eta}(A) = P(\xi + \eta \in A) = P((\xi, \eta) \in B) = P_{\xi,\eta}(B) = \int_{\mathbb{R}^2} \mathbb{1}_B(x, y) dP_{\xi,\eta}(x, y) =$

$\underbrace{=}_{\text{т.к. } \xi \text{ и } \eta \text{ независимые}} \int_{\mathbb{R}^2} \mathbb{1}_B(x, y) dP_\xi(x)\, dP_\eta(y) = \int_{\mathbb{R}^2} \mathbb{1}_A(x + y) dP_\xi(x)\, dP_\eta(y) = P_\xi * P_\eta(A)$ □

Пример. 1. **Свертка с дисретктным распределением**

$\nu = \sum_{k=1}^{\infty} p_k \delta_{x_k}$.

$$\delta_{x_k}(A) = \begin{cases} 1, \ x \in A \\ 0, \ otherwise \end{cases} \tag{1}$$

Тогда $\mu * \nu(A) = \int_{\mathbb{R}} \mu(A-x)\, d\nu(x) = \sum_{k=1}^{\infty} \mu(A-x_k)p_k$

2. $\xi_i \sim Poisson(\lambda_i)$. ξ_1 и ξ_2 независимы.

   $P_{\xi_1+\xi_2}(\{n\}) = \sum_{k=0}^{+\infty} P_{\xi_1}(\{n-k\}) \cdot \frac{\lambda_2^k e^{-\lambda_2}}{k!} = \sum_{k=0}^{n} \frac{\lambda_1^{n-k} e^{-\lambda_1}}{(n-k)!} \cdot \frac{\lambda_2^k e^{-\lambda_2}}{k!} = e^{-\lambda_1} e^{-\lambda_2} \sum_{k=0}^{n} \frac{\lambda_1^{n-k}\lambda_2^k}{k!(n-k)!} = \frac{(\lambda_1+\lambda_2)^n e^{-\lambda_1-\lambda_2}}{n!}$

   $\xi_1 + \xi_2 \sim Poisson(\lambda_1 + \lambda_2)$

2.4. Математическое ожидание и дисперсия

***Определение* 2.17.** $\xi : \Omega \to \mathbb{R}$ - случайная величина ($\xi \geq 0$, либо суммируемая функция). $\mathbb{E}\xi = \int_{\mathbb{R}} \xi(\omega)\, dP(\omega)$ - математическое ожидание (среднее значение случайной величины).

Свойства. 1. $a, b \in \mathbb{R}: \ \mathbb{E}(a\xi + b\eta) = a\mathbb{E}\xi + b\mathbb{E}\eta$

2. Если $\xi \geqslant 0$, с вероятностью 1, то $\mathbb{E}\xi \geqslant 0$ (по сути написано, что если функция почти везде неотрицательна, то интеграл неотрицателен).

3. Если $\xi \geqslant \eta$ с вероятностью 1, то $\mathbb{E}\xi \geqslant \mathbb{E}\eta$

4. $\mathbb{E}\xi = \int_{\mathbb{R}} x\, dP_\xi(x)$

5. Если $f : \mathbb{R}^n \to \mathbb{R}$ - измерима относительно борелевской $\sigma-$алгебры.

   Тогда $\mathbb{E}f(\xi_1, \xi_2 \ldots \xi_n) = \int_{\mathbb{R}^n} f(x_1, \ldots, x_n) dP_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n)$

   Доказательство: $f = \mathbb{1}_A$. Тогда $\mathbb{E}\mathbb{1}_A(\xi_1, \ldots \xi_n) = \int_\Omega \mathbb{1}_A(\xi_1(w), \ldots, \xi_n(w)) dP(\omega) = P(\omega \in \Omega : \bar{\xi} \in A) = P_{\bar{\xi}}(A) = \int_{\mathbb{R}^n} \mathbb{1}_A(x_1, \ldots, x_n) dP_{\bar{\xi}}(x_1, \ldots, x_n)$.

   Тогда по линейности верно для простых.

   Теперь берём f_j неотрицательный простые, такие, что возрастают и $\to f$. И предельный переход по теореме Леви.

6. Если ξ_1 и ξ_2 независимы, то $\mathbb{E}(\xi \cdot \eta) = \mathbb{E}\xi \cdot \mathbb{E}\eta$

   Доказательство: $\mathbb{E}(\xi\eta) = \int_{\mathbb{R}^2} xy dP_{\xi,\eta}(x, y) =$

   $$\underbrace{=}_{\text{независимость сл. вел.}} \int_{\mathbb{R}} \int_{\mathbb{R}} xy dP_\xi(x)\, dP_\eta(y) = \int_{\mathbb{R}} y \int_{\mathbb{R}} x dP_\xi(x)\, dP_\eta(y) = \mathbb{E}\xi \cdot \mathbb{E}\eta$$

7. Если $\xi \geqslant 0$, то $\mathbb{E}\xi = \int_0^{+\infty} P(\xi \geqslant t)\, dt$ - из теории меры.

8. Если $p, q > 1$ и $\frac{1}{p} + \frac{1}{q} = 1$, то $\mathbb{E}|\xi\eta| \leqslant (\mathbb{E}|\xi|^p)^{\frac{1}{p}} (\mathbb{E}|\eta|^q)^{\frac{1}{q}}$ - неравенство Гёльдера

9. Неравенство Ляпунова

   $0 < r < s$, тогда $(\mathbb{E}|\xi|^r)^{\frac{1}{r}} \leqslant (\mathbb{E}|\xi|^s)^{\frac{1}{s}}$.

   Доказательство: $p = \frac{s}{r} > 1, \ \frac{1}{q} = 1 - \frac{1}{p} = \frac{s-r}{s} < 1$.

   Тогда запишем Гельдера для ξ и $\eta = 1$:

   $\mathbb{E}|\xi|^r|1| \leq (\mathbb{E}(|\xi|^r)^p)^{\frac{1}{p}} \cdot (\mathbb{E}1^q)^{\frac{1}{q}} = (\mathbb{E}|\xi|^s)^{\frac{r}{s}}$.

`Замечание.` $\mathbb{E}(\xi\eta) = \mathbb{E}\xi \cdot \mathbb{E}\eta$ без независимости неверно.

Возьмём $\xi = \pm 1$ с вероятностями $\frac{1}{2}$. Тогда $\mathbb{E}\xi = 0$.

Также пусть $\eta = \xi$. Тогда $\xi\eta = \mathbb{E}\xi^2 = 1 \neq (\mathbb{E}\xi)^2$

Теорема 2.6. Неравенство Маркова

Если $\xi \geqslant 0, p, t > 0$, то $P(\xi \geqslant t) \leqslant \frac{\mathbb{E}\xi^p}{t^p}$.

Доказательство. Неравенство Чебышёва из теории меры. □

***Определение* 2.18.**

1. Моменты случайной величины. $\mathbb{E}(\xi^k)$ - k-ый момент.
2. Центральный момент. $\mathbb{E}(\xi - \mathbb{E}\xi)^k$ - k-ый центральный момент.
3. Абсолютный момент. $\mathbb{E}|\xi|^k$ - k-ый абсолютный момент.

***Определение* 2.19.** Медиана случайной величины. m - медиана ξ, если $P(\xi \geqslant m) \geqslant \frac{1}{2}$ и $P(\xi \leqslant m) \geqslant \frac{1}{2}$.

`Замечание.` Медиана не единственна.

Возьмём кубик. $\xi = 1, 2, \dots, 6$ с вероятностью $\frac{1}{6}$. Тогда любое число $m \in [3, 4]$ подходит.

Чаще всего всё равно берут середину, чтобы была единственность.

Пример. Есть организация из 1000 человек. 1 начальник и 999 подчиненных.

Зарплата начальника 1.000.000$, а подчинённых 1000$.

$\mathbb{E} = \frac{999}{1000} \cdot 1000 + \frac{1}{1000} \cdot 1000000 = 1999$

$m = 1000$ - медиана лучше характеризует ситуацию в этом случае.

***Определение* 2.20.** Дисперсия. $\mathbb{D}\xi = \mathbb{E}(\xi - \mathbb{E}\xi)^2$ - второй центральный момент.

Обозначение в англоязычной литературе: $Var\xi$

Свойства.

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

   Доказательство: Пусть $a = \mathbb{E}\xi$.

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

2. $\mathbb{D}\xi \geqslant 0$ и если $\mathbb{D}\xi = 0$, то $P(\xi = c) = 1$

   Доказательство: Если $\mathbb{D}\xi = 0$, то $\int_\Omega (\xi - a)^2 \, dP = 0$, значит $(\xi - a)^2 = 0$ почти везде.

3. $\mathbb{D}(\xi + a) = \mathbb{D}\xi$

   Доказательство: $\mathbb{E}(\xi + a) = \mathbb{E}\xi + a$. А тогда $(\xi + a) - \mathbb{E}(\xi + a) = \xi - \mathbb{E}\xi$

4. $\mathbb{D}(c\xi) = c^2\mathbb{D}\xi$

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

5. Если ξ и η независимы, то $\mathbb{D}(\xi + \eta) = \mathbb{D}\xi + \mathbb{D}\eta$

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

6. Аналогично предыдущему, но для n случайных величин.

   Доказательство: индукция

7. $\mathbb{E}|\xi - \mathbb{E}\xi| \leqslant \sqrt{\mathbb{D}\xi}$

   Доказательство: $\mathbb{E}|\xi - \mathbb{E}\xi| \leqslant (\mathbb{E}|\xi - \mathbb{E}\xi|^2)^{\frac{1}{2}} = \sqrt{\mathbb{D}\xi}$ - написали Ляпунова.

8. **Неравенство Чебышёва**

   $P(|\xi - \mathbb{E}\xi| \geqslant t) \leqslant \frac{\mathbb{D}\xi}{t^2}$, где $t > 0$

   Доказательство: $P(|\xi - \mathbb{E}\xi| \geqslant t) \leqslant \frac{\mathbb{E}|\xi - \mathbb{E}\xi|^2}{t^2} = \frac{\mathbb{D}\xi}{t^2}$ - неравенство Маркова для $p = 2$.

***Определение* 2.21.** Стандартное отклонение $\sigma = \sqrt{\mathbb{D}\xi}$

Пример. 1. $\xi \sim U[0,1]$.

   Тогда $\mathbb{E}\xi = \int_0^1 x\,dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2}$.

   $\mathbb{E}\xi^2 = \int_0^1 x^2\,dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}$. А тогда $\mathbb{D}\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2 = \frac{1}{12}$

2. $\xi \sim U[a,b]$.

   Если $\eta \sim U[0,1]$ и $\xi = (b-a)\eta + a \sim U[a,b]$. Тогда $\mathbb{E}\xi = \mathbb{E}((b-a)\eta + a) = \frac{a+b}{2}$

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

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

   $\mathbb{E}\xi = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} xe^{\frac{-x^2}{2}}\,dx = 0$, так как функция нечётная.

   Значит $\mathbb{D}\xi = \mathbb{E}\xi^2 = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} x^2e^{-\frac{x^2}{2}}\,dx = -\frac{e^{\frac{-x^2}{2}}x}{\sqrt{2\pi}}\Big|_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{x^2}{2}}\,dx = 1$

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

   Если $\eta \sim \mathcal{N}(0,1)$, то $\xi = \sigma\eta + a \sim \mathcal{N}(a,\sigma^2)$.

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

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

***Определение* 2.22.** Пусть $\mathbb{E}\xi^2 < +\infty$ и $\mathbb{E}\eta^2 < +\infty$.

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

Свойства. 1. $cov(\xi,\xi) = \mathbb{D}\xi$

2. $cov(\xi,\eta) = cov(\eta,\xi)$

3. $cov(c\xi,\eta) = c\cdot cov(\xi,\eta)$

4. $cov(\xi_1 + \xi_2,\eta) = cov(\xi_1,\eta) + cov(\xi_2,\eta)$

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

   Доказательство. $\mathbb{E}\xi = a, \mathbb{E}\eta = b$

   $cov(\xi,\eta) = \mathbb{E}((\xi - a)(\eta - b)) = \mathbb{E}(\xi\eta) - a\mathbb{E}\eta - b\mathbb{E}\xi + ab$ □

6. Если ξ и η независимы, то $cov(\xi,\eta) = 0$

7. $\mathbb{D}(\xi + \eta) = \mathbb{D}\xi + \mathbb{D}\eta + 2cov(\xi,\eta)$

8. $\mathbb{D}(\xi_1 + \xi_2 + \ldots + \xi_n) = \mathbb{D}\xi_1 + \mathbb{D}\xi_2 + \ldots + \mathbb{E}\xi_n + 2\sum_{i<j} cov(\xi_i, \xi_j)$.

Пример. P(успех) $= p$. Делаем n подбрасываний. $\eta =$ количество переходов от орла к решке.

Пусть $\xi_i = 1$, если на i позиции орёл, на $i+1$ позиции решка, иначе $\xi_i = 0$.

$\eta = \xi_1 + \ldots + \xi_{n-1}$. Тогда $\mathbb{E}\eta = \sum_{i=1}^{n-1} \mathbb{E}\xi_i = (n-1)pq$.

$\mathbb{D}\eta = \sum_{i=1}^{n-1} \mathbb{D}\xi_i + 2\sum_{i<j} cov(\xi_i, \xi_j)$.

Если $i + 1 < j$, то ξ_i и ξ_j независимы, поэтому в сумме почти везде нули.

Значит $\mathbb{D}\eta = \sum_{i=1}^{n-1} \mathbb{D}\xi_i + 2\sum_{i=1}^{n-1} cov(\xi_i, \xi_{i+1})$.

$\mathbb{D}\xi_i = \mathbb{E}\xi_i^2 - (\mathbb{E}\xi_i)^2 = pq - p^2q^2$.

$cov(\xi, \xi_{i+1}) = \mathbb{E}(\xi_i\xi_{i+1}) - \mathbb{E}\xi_i\mathbb{E}\xi_{i+1} = -p^2q^2$

`Замечание.` 1. $\{\xi : \mathbb{E}\xi^2 < +\infty\}$

$\langle \xi, \eta \rangle = \mathbb{E}(\xi\eta)$ - скалярное произведение.

$\mathbb{E}\xi$ - ортогональная проекция на константы.

2. $\langle \xi, \eta \rangle = cov(\xi, \eta)$ - тоже скалярное произведение.

Норма - это стандартное отклонение.

Теорема 2.7. Выбор двудольного подграфа

Есть граф G с n вершинами и m рёбрами. Хотим стереть некоторое количество рёбер(как можно меньше) так, чтобы остался двудольный подграф.

Тогда G содержит двудольный подграф с $\geqslant \frac{m}{2}$ рёбрами.

Доказательство. A - те вершины, на которых выпал орёл, B - на которых выпала решка.

Будем интересоваться матожиданием количества рёбер в такой ситуации. Пусть $xy \in E(G)$, сопоставим ребру следующую случайную величину:

$$\xi_{xy} = \begin{cases} 1, & \text{если x, y из разных долей} \\ 0, & \text{иначе} \end{cases}$$

Пусть $\eta = \sum_{xy \in E} \xi_{xy}$ - число рёбер, которое нужно оставить, при таком разбиении на доли.

$\mathbb{E}\eta = \sum_{xy \in E} \mathbb{E}\xi_{xy} = m \cdot (\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0) = \frac{m}{2}$, а значит есть реализация с $\frac{m}{2}$ рёбрами (если бы все значения кол-ва ребер были меньше $\frac{m}{2}$, то и мат. ожидание было бы меньше $\frac{m}{2}$). ☐

***Определение* 2.23.** Коэффициент корреляции. $\rho(\xi, \eta) = \frac{cov(\xi,\eta)}{\sqrt{\mathbb{D}\xi}\sqrt{\mathbb{D}\eta}} \in [-1, 1]$

***Определение* 2.24.** Если $cov(\xi, \eta) = 0$, то это некоррелирующие случайные величины.

Теорема 2.8. $v_1, v_2 \ldots v_n \in \mathbb{R}^n$ - векторы единичной длины, тогда существует расстановка знаков $\varepsilon_1 = \pm 1, \ldots, \varepsilon_n = \pm 1$, такая, что $||\varepsilon_1 v_1 + \ldots + \varepsilon_n v_n|| \leqslant \sqrt{n}$.

`Замечание.` Эта оценка не улучшаема, если все вектора попарно ортогональны, тогда длина вектора $\sqrt{n}$.

Доказательство. Пусть $\varepsilon_1 \ldots \varepsilon_n$ - независимые случайные величины, такие, что:

$$\varepsilon_i = \begin{cases} 1, & \text{с вероятностью } \frac{1}{2} \\ -1, & \text{с вероятностью } \frac{1}{2} \end{cases}$$

Введем величину $\xi = ||\varepsilon_1 v_1 + \ldots + \varepsilon_n v_n||^2$.

Тогда $\mathbb{E}\xi = \mathbb{E}\langle v, v \rangle = \mathbb{E}(\sum_{i,j=1}^n \varepsilon_i\varepsilon_j \langle v_i, v_j \rangle) = \sum_{i,j=1}^n \langle v_i, v_j \rangle \mathbb{E}\varepsilon_i\varepsilon_j = \sum_{i=1}^n \langle v_i, v_j \rangle = n$.

1. Если $i = j$, то $\mathbb{E}\varepsilon_i\varepsilon_j = \mathbb{E}\varepsilon_i^2 = 1$

2. Если $i \neq j$, то $\mathbb{E}\varepsilon_i\varepsilon_j \overset{\text{независимость}}{=} \mathbb{E}\varepsilon_i \cdot \mathbb{E}\varepsilon_j = 0$

□

Теорема 2.9. $v_1, v_2 \ldots v_n \in \mathbb{R}^n$, $||v_i|| \leqslant 1, p_i \in [0, 1]$ и $w = p_1v_1 + \ldots + p_nv_n$

Тогда существует $\varepsilon_1 \in \{0, 1\}, \ldots \varepsilon_n \in \{0, 1\}$, такие, что $v = \varepsilon_1 v_1 + \ldots + \varepsilon_n v_n$ и $||v - w|| \leqslant \frac{\sqrt{n}}{2}$

Доказательство. Пусть $\varepsilon_1 \ldots \varepsilon_n$ - независимые случайные величины.

$$\varepsilon_i = \begin{cases} 1, & \text{с вероятностью } p_i \\ 0, & \text{с вероятностью } 1 - p_i \end{cases}$$

Интересуемся $\xi = ||v - w||^2$. Тогда $\mathbb{E}\xi = \mathbb{E}(\sum_{i,j=1}^n (\varepsilon_i - p_i)(\varepsilon_j - p_j) \langle v_i, v_j \rangle) = \sum_{i,j=1}^n \langle v_i, v_j \rangle \mathbb{E}(\varepsilon_i - p_i)(\varepsilon_j - p_j) \overset{\text{пояснение ниже}}{=} \sum_{i=1}^n \langle v_i, v_i \rangle (p_i - p_i^2) \leqslant \frac{n}{4}$.

1. Если $i = j$, то $cov(\varepsilon_i, \varepsilon_j) = \mathbb{D}\varepsilon_i = p_i - p_i^2 \leqslant \frac{1}{4}$

2. Если $i \neq j$, то $cov(\varepsilon_i, \varepsilon_j) \overset{\text{независимы}}{=} 0$

□

Теорема 2.10. Харди-Рамануджана

Пусть $\nu(k)$ – число различных простых делителей в разложении k.

Хотим понять, чему будет равно это число, если мы наугад возьмем число из мн-ва $\{1, 2, \ldots, n\}$.

$P(|\nu(k) - \ln\ln n| \geqslant w(n)\sqrt{\ln\ln n}) \to_{n\to\infty} 0$, где $w(n) \to_{n\to\infty} +\infty$ и $k \in \{1, 2, \ldots, n\}$.

Доказательство. Пусть $m = \sqrt[10]{n}$. $p \leqslant m$ - простое и

$$\xi_p(k) = \begin{cases} 1, & \text{если } k \text{ делится на } p \\ 0, & \text{иначе} \end{cases}$$

$\xi = \sum_{p \leqslant m} \xi_p$ - количество различных простых $\leqslant m$. Тогда $\nu(k) - 10 \leqslant \xi(k) \leqslant \nu(k)$.

Посчитаем матожидание ξ, тогда посчитаем мат. ожидание слагаемых:

$\mathbb{E}\xi_p = \frac{[\frac{n}{p}]}{n} \leqslant \frac{\frac{n}{p}}{n} = \frac{1}{p}$. С другой стороны, $\mathbb{E}\xi_p \geqslant \frac{\frac{n}{p} - 1}{n} = \frac{1}{p} - \frac{1}{n}$.

Знаем, что $\mathbb{E}\xi = \sum_{p \leqslant m} \mathbb{E}\xi_p$, тогда

$\sum_{p \leqslant m} \frac{1}{p} - \frac{m}{n} \leqslant \sum_{p \leqslant m} \mathbb{E}\xi_p \leqslant \sum_{p \leqslant m} \frac{1}{p} = \ln\ln m + \mathcal{O}(1) = \ln\ln n + \mathcal{O}(1)$. Оценка в другую сторону аналогично, потому что $\frac{m}{n} \leqslant 1$.

Теперь считаем дисперсию для ξ_p:

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

Теперь оценим ковариацию:

$$cov(\xi_p, \xi_q) = \underbrace{\mathbb{E}(\xi_p\xi_q)}_{\text{аргумент равен 1, когда } n \,\vdots\, pq} - \mathbb{E}\xi_p\mathbb{E}\xi_q = \frac{\left[\frac{n}{pq}\right]}{n} - \frac{\left[\frac{n}{p}\right]}{n} \cdot \frac{\left[\frac{n}{q}\right]}{n} = (*).$$

Оценим $(*)$ с двух сторон:

1. $(*) \geq \frac{\frac{n}{pq} - 1}{n} - \frac{\frac{n}{p}}{n} \cdot \frac{\frac{n}{q}}{n} = -\frac{1}{n}$

2. $(*) \le \frac{\frac{n}{pq}}{n} - \frac{\frac{n}{p}-1}{n} \cdot \frac{\frac{n}{q}-1}{n} \le (\frac{1}{p}+\frac{1}{q}) \cdot \frac{1}{n}$

Теперь смотрим на сумму ковариаций (так как она фигурирует как слагаемое для $\mathbb{D}\xi$):

$$\underbrace{\frac{1}{n} \cdot \sum_{p<q\le m} \left(\frac{1}{p}+\frac{1}{q}\right)}_{=\frac{1}{2n}\sum_{p\neq q,\ p,q\le m}\left(\frac{1}{p}+\frac{1}{q}\right)\le\frac{1}{2n}2m\sum_{p\le m}\frac{1}{p}=\mathcal{O}(1)} \ge \sum_{p<q\le m} cov(\xi_p,\xi_q) \ge -\frac{m^2}{n} = \mathcal{O}(1)$$

Теперь оцениваем дисперсию для ξ:

$$\mathbb{D}\xi = \sum_{p\le m}\mathbb{D}\xi_p + 2\underbrace{\sum_{p<q\le m} cov(\xi_p,\xi_q)}_{=\mathcal{O}(1)} = \sum_{p\le m}\left(\frac{1}{p}-\frac{1}{p^2}+\mathcal{O}(\frac{1}{n})\right)+\mathcal{O}(1) =$$

$= \sum_{p\le m}\frac{1}{p} + \mathcal{O}(1) = \ln\ln m + \mathcal{O}(1) = \ln\ln n + \mathcal{O}(1)$.

Теперь применим Чебышёва.

$P(|\xi - \mathbb{E}\xi| \geqslant t) \leqslant \frac{\mathbb{D}\xi}{t^2}$. В качестве t подставим $w(n)\sqrt{\ln\ln n}$.

Тогда $P(|\nu(k)-\ln\ln n| \geqslant w(n)\sqrt{\ln\ln n}) \underbrace{<}_{(**)} P(|\xi-\mathbb{E}\xi| \geqslant w(n)\sqrt{\ln\ln n}) \leqslant \frac{\mathbb{D}\xi}{w^2(n)\ln\ln n} \to 0$.

$(**)$: такое нер-во можно писать, так как $|\nu(k)-\xi(k)| \le 10$ и $\mathbb{E}\xi = \ln\ln n + \mathcal{O}(1)$.

□

Теорема 2.11. Эрдёша-Каца

$\lim_{n\to\infty}\frac{\#\{k\le n : a\leqslant\frac{|\nu(k)-\ln\ln n|}{\sqrt{\ln\ln n}}\leqslant b\}}{n} \to \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{t^2}{2}}\,dt$

2.5. Сходимость последовательностей случайных величин

Теорема 2.12. $\xi_1, \xi_2, \ldots$ - независимые случайные величины, $f_i : \mathbb{R}^{n_i} \to \mathbb{R}$ - измерима, относительно борелевской σ-алгребры.

Тогда $f_1(\xi_1,\ldots\xi_{n_1}), f_2(\xi_{n_1+1},\ldots,\xi_{n_1+n_2})$ - независимые случаные величины.

Доказательство. $f : \mathbb{R}^m \to \mathbb{R}$ и $g : \mathbb{R}^n \to \mathbb{R}$, $\xi_1 \ldots \xi_m$ и $\eta_1 \ldots \eta_n$ независимые случайные величины.

Возьмём $\tilde{A}$ и $\tilde{B} \in \mathbb{R}$ борелевские.

Надо доказать, что

$P(f(\xi_1\ldots\xi_m)\in\tilde{A}) \cdot P(g(\eta_1\ldots\eta_n)\in\tilde{B}) = P(f(\xi_1,\ldots,\xi_m)\in\tilde{A}, g(\eta_1\ldots\eta_n)\in\tilde{B})$.

$P((\xi_1,\ldots,\xi_m)\in\underbrace{f^{-1}(\tilde{A})}_{=:A}) \cdot P((\eta_1,\ldots,\eta_n)\in\underbrace{g^{-1}(\tilde{B})}_{=:B}) = P((\xi_1,\ldots,\xi_m)\in A,\ (\eta_1,\ldots,\eta_n)\in B)$

Поймём это для ячеек.

$A = (a,b] :\ (\xi_1,\ldots,\xi_m)\in(a,b] \Leftrightarrow \forall k :\ \xi_k\in(a_k,b_k]$

$B = (c,d] :\ (\eta_1,\ldots,\eta_n)\in(c,d] \Leftrightarrow \forall k :\ \eta_k\in(c_k,d_k]$

Мы знаем, что

$P((\xi,\ldots,\xi_m)\in A,\ (\eta_1,\ldots,\eta_n)\in B) = P(\forall k :\ \xi_k\in(a_k,b_k],\ \forall k :\ \eta_k\in(c_k,d_k]) = \prod_{k=1}^m P(\xi_k\in(a_k,b_k]) \cdot \prod_{k=1}^n P(\eta_k\in(c_k,d_k])$

Но $\prod_{k=1}^m P(\xi_k\in(a_k,b_k]) = P((\xi_1,\ldots,\xi_m)\in A)$, а $\prod_{k=1}^n P(\eta_k\in(c_k,d_k]) = P((\eta_1,\ldots,\eta_n)\in B)$.

То есть доказали на ячейках, а значит и доказали теорему.

□

***Определение* 2.25.** $\xi, \xi_1, \xi_2, \ldots : \Omega \to \mathbb{R}$.

1. ξ_n сходится к ξ почти наверное, если $P(w \in \Omega : \lim_{n\to\infty} \xi_n(w) = \xi(w)) = 1$
2. ξ_n сходится к ξ в среднем порядка $r > 0$, если $\mathbb{E}(|\xi_n - \xi|^r) \to_{n\to\infty} 0$
3. ξ_n сходится к ξ по вероятности, если $\forall \varepsilon > 0,\ P(\omega \in \Omega : |\xi_n(\omega) - \xi(\omega)| \geqslant \varepsilon) \to_{n\to\infty} 0$
4. $\xi_n : \Omega_n \to \mathbb{R}$.

   ξ_n сходится к ξ по распределению, если $\lim_{n\to\infty} F_{\xi_n}(x) = F_\xi(x)$ во всех точках непрерывности F_ξ.

Связь между сходимостями:

1. $1 \Rightarrow 3$: теорема Лебега из теории меры
2. $2 \Rightarrow 3$:

   Пишем нер-во Маркова

   $P(|\xi_n - \xi| \geq \varepsilon) \leq \frac{\mathbb{E}|\xi_n - \xi|^r}{\varepsilon^r} \to 0$.
3. $2 \not\Rightarrow 1$:

   $\Omega := [0, 1),\ P := \lambda$ (мера Лебега).

   Возьмем такую последовательность функций, для которой нет такого следствия:

   $\mathbb{1}_{[0,1)}, \mathbb{1}_{[0,\frac{1}{2})}, \mathbb{1}_{[\frac{1}{2},1)}, \mathbb{1}_{[0,\frac{1}{3})}, \mathbb{1}_{[\frac{1}{3},\frac{2}{3})}, \mathbb{1}_{[\frac{2}{3},1)}, \ldots$

   Тогда $\mathbb{E}|\xi_n|^r = \mathbb{E}\ \mathbb{1}_{[\frac{k}{m},\frac{k+1}{m})} = \frac{1}{m} \to \infty$, то сх-ть в среднем есть.

   Сх-ти почти наверное нет, потому что ни в одной точке нет сх-ти, так как у $\xi_n(\omega)$ сколько угодно далеко есть как значения 1, так и значения 0.

   Иными словами $\lim \xi_n(\omega)$ не существует.
4. $3 \not\Rightarrow 1$:

   верно, так как $2 \Rightarrow 3$, но $2 \not\Rightarrow 1$.
5. $1 \not\Rightarrow 2$:

   $\Omega := [0, 1],\ P := \lambda$ (мера Лебега).

   $\xi_n(\omega) = n^{\frac{1}{r}} \cdot \mathbb{1}_{[0,\frac{1}{n})}$.

   Тогда $\xi_n(\omega) \to 0$ при $\omega \neq 0$, тогда ξ_n сх-ся к $\xi \equiv 0$ почти наверное.

   $\mathbb{E}|\xi_n|^r = \mathbb{E}\left(n\mathbb{1}_{[0,\frac{1}{n})}\right) = 1 \not\to 0$
6. $3 \not\Rightarrow 2$:

   верно, так как $1 \not\Rightarrow 2$.

7. $3 \Rightarrow 4$:

   Хотим понять, что $P(\xi_1 \le x) \to P(\xi \le x)$ если x – точка непрерывности F_ξ.

   $\{\xi_n \le x\} \subset \{\xi \le x + \varepsilon\} \cup \{|\xi_n - \xi| > \varepsilon\}$

   $P(\xi_n \le x) \subset P(\xi \le x + \varepsilon) + P(|\xi_n - \xi| > \varepsilon)$

   Напишем верхние пределы для этого нер-ва:

   $\overline{\lim} P(\xi_n \le x) \le \underbrace{P(\xi \le x + \varepsilon)}_{\text{это просто } const} + \underbrace{\overline{\lim} P(|\xi_n - \xi| > \varepsilon)}_{\to 0\text{, т.к. сх-ть по вероятности}}$

   Теперь надо подпереть чем-то снизу:

   $\{\xi_n \le x\} \supset \{\xi \le x - \varepsilon\} \setminus \{|\xi_n - \xi| > \varepsilon\}$

   $P(\xi_n \le x) \ge P(\xi \le x - \varepsilon) - P(|\xi_n - \xi| > \varepsilon)$

   Теперь пишем нижние пределы:

   $\underline{\lim} P(\xi_n \le x) \ge \underbrace{P(\xi \le x - \varepsilon)}_{=const} - \underbrace{P(|\xi_n - \xi| > \varepsilon)}_{\to 0}$

   Тогда получаем, что

   $F_\xi(x - \varepsilon) \le \underline{\lim} F_{\xi_n}(x) \le \overline{\lim} F_{\xi_n}(x) \le F_\xi(x + \varepsilon)$

   Теперь устремим $\varepsilon \to 0$, тогда

   $F_\xi(x) \le \underline{\lim} F_{\xi_n}(x) \le \overline{\lim} F_{\xi_n}(x) \le F_\xi(x)$

   Тогда $\lim_{n\to\infty} F_{\xi_n}(x) = F_\xi(x)$ – доказали стрелочку.

8. $4 \not\Rightarrow 3, 2, 1$:

   из-за разных определений (где-то одно вероятностное пр-во, а где-то их может быть много разных)

Теорема 2.13. Закон больших чисел

$\xi_1, \xi_2, \ldots$ - попарно некоррелируемые случайные величины и $\mathbb{D}\xi_n = o(n)$.

$S_n = \xi_1 + \xi_2 + \ldots + \xi_n$.

Тогда $\frac{S_n}{n} - \mathbb{E}\frac{S_n}{n} \underset{P}{\to} 0$. То есть вероятность того, что $P(|\frac{S_n}{n} - \mathbb{E}\frac{S_n}{n}| \geqslant \varepsilon) \to 0$

Следствие. Если $\mathbb{D}\xi_n$ ограничены, то такой же вывод.

Доказательство. $P(\left|\frac{S_n}{n} - \mathbb{E}\frac{S_n}{n}\right| \geqslant \varepsilon) \leqslant \frac{\mathbb{D}\frac{S_n}{n}}{\varepsilon^2} = \frac{\mathbb{D}S_n}{\varepsilon^2 n^2} = \frac{\sum_{k=1}^n \mathbb{D}\xi_k}{\varepsilon^2 n^2} \to_{\text{Штольц}} \lim_{n\to\infty} \frac{\mathbb{D}\xi_n}{\varepsilon^2(2n-1)} = 0.$ □

***Следствие.* ЗБЧ в форме Чебышёва**

$\xi_1, \xi_2, \ldots$ независимые, одинаково распределенённые случайные величины с конечной дисперсией и $a = \mathbb{E}\xi_1$.

Тогда $P(\left|\frac{S_n}{n} - a\right| \geqslant \varepsilon) \to 0$ или же $\frac{S_n}{n} \underset{P}{\to} a$

Доказательство. Мат. ожидание всех случайных величин равно, они одинаково распределены. Поэтому $\mathbb{E}\frac{S_n}{n} = \mathbb{E}\frac{\xi_1 + \ldots + \xi_n}{n} = a$. Поэтому все условия предыдущей теоремы выполнены □

***Следствие.* ЗБЧ для схем Бернулли**

Есть схема Бернулли с вероятностью успеха $p \in (0, 1)$.

Тогда $P(|\frac{S_n}{n} - p| \ge \varepsilon) \to 0$, где S_n число успехов при n подбрасываниях.

Теорема 2.14. Усиленный ЗБЧ

$\xi_1, \xi_2, \ldots$ - независимые случайные величины. $\mathbb{E}(\xi_n - \mathbb{E}\xi_n)^4 \leqslant C$.

Тогда $\frac{S_n}{n} - \mathbb{E}\frac{S_n}{n} \to 0$ почти наверное.

Доказательство. $\frac{S_n}{n} - \mathbb{E}\frac{S_n}{n} = \frac{1}{n}(S_n - \mathbb{E}S_n) = \frac{1}{n}(\sum_{k=1}^{n}(\xi_k - \mathbb{E}\xi_k))$. Задвинем все матожидания в ноль.

Тогда по условию $\mathbb{E}\xi_n^4 \leqslant C$ и надо доказать, что $\frac{S_n}{n} \to 0$ почти наверное.

Пусть $A_n = \{\left|\frac{S_n}{n}\right| \geqslant \varepsilon\}$. Нам нужно понять, что бесконечное количество A_n случаются с нулевой вероятностью, то есть что $P(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k) = 0$.

Из леммы Бореля-Кантелли, если $\sum_{k=1}^{\infty} P(A_n) < +\infty$, то нужное нам условие выполнено.

Напишем неравенство Маркова: $P(A_n) = P\left(\frac{S_n^4}{n^4} \geq \varepsilon^4\right) \leqslant \frac{\mathbb{E}\frac{S_n^4}{n^4}}{\varepsilon^4} = \frac{\mathbb{E}S_n^4}{n^4\varepsilon^4}$. Достаточно доказать, что $\mathbb{E}S_n^4 = \mathcal{O}(n^2)$, тогда ряд сойдётся. Раскроем все скобки.

$\mathbb{E}(\xi_1 + \ldots + \xi_n)^4 = \sum_{i=1}^{n}\mathbb{E}\xi_i^4 + 4\sum_{i\neq j}\mathbb{E}\xi_i^3\xi_j + 6\sum_{i\neq j}\mathbb{E}\xi_i^2\xi_j^2 + 12\sum_{i\neq j\neq k}\mathbb{E}\xi_i^2\xi_j\xi_k + 24\sum_{\ldots}\mathbb{E}\xi_i\xi_j\xi_k\xi_m$

1. $\mathbb{E}\xi_i\xi_j\xi_k\xi_m = 0$
2. $\mathbb{E}\xi_i^2\xi_j\xi_k = 0$

Итого получаем $\sum_{i=1}^{n}\mathbb{E}\xi_i^4 + 6\sum\mathbb{E}x_i^2\mathbb{E}\xi_j^2 = (*)$. По неравенству Ляпунова $\mathbb{E}\xi_i^2 \leqslant \sqrt{\mathbb{E}\xi_i^4} \leqslant \sqrt{C}$.

Значит $(*) = nC + 6n(n-1)\sqrt{C}\sqrt{C} \leqslant 6Cn^2 = \mathcal{O}(n^2)$, значит ряд сходится и лемма Бореля-Кантелли выполняется. □

Следствие. **Усиленный ЗБЧ для схем Бернулли**

В схеме Бернулли с вероятностью успеха p : $\frac{S_n}{n} \to p$ почти наверное.

Доказательство. Нужно проверить, что $\mathbb{E}(\xi_i - p)^4$ - конечно, раскроем скобки, получим какие-то константы и ξ_i^4. □

Теорема 2.15. Усиленный ЗБС в форме Колмогорова

$\xi_1, \xi_2, \ldots$ - независимо, одинаково распределённые случайные величины.

Тогда $\frac{S_n}{n} \to a \in \mathbb{R}$ почти наверное $\Leftrightarrow a = \mathbb{E}\xi_1$

Метод Монте-Карло

Φ - ограниченная фигура на плоскости. Хотим примерно узнать её площадь.

Берём случайную точку в прямоугольнике и выясняем, попала она в фигуру или нет.

$$\xi_i = \begin{cases} 1, & \text{точка попала в } \Phi \\ 0, & \text{иначе} \end{cases}$$

Вероятность успеха $\frac{Area(\Phi)}{Area(\text{прямоугольника})}$. Тогда усиленный ЗБЧ говорит, что $\frac{S_n}{n} \to p$ почти наверное.

Теорема 2.16. $\xi_1, \xi_2, \ldots$ последовательность случайных величин, $\xi_n \to_P a \in \mathbb{R}$. f ограниченная функция, непрерывная в точке a.

Тогда $\mathbb{E}f(\xi_n) \to f(a)$

Доказательство. $|\mathbb{E}f(\xi_n) - f(a)| = |\mathbb{E}(f(\xi_n) - f(a))| \leqslant \mathbb{E}|f(\xi_n) - f(a)| = \mathbb{E}|f(\xi_n - f(a))| \cdot \mathbb{1}_{\{\xi_n - a < \varepsilon\}} + \mathbb{E}|f(\xi_n - f(a))| \cdot \mathbb{1}_{\{\xi_n - a \geqslant \varepsilon\}} = (*)$.

Пусть f ограничена константой M.

$\mathbb{E}|f(\xi_n - f(a))| \cdot \mathbb{1}_{\{\xi_n - a \geqslant \varepsilon\}} \leqslant 2M\mathbb{E}\mathbb{1}_{\{\xi_n - a \geqslant \varepsilon\}}$

$|f(\xi_n - f(a))| \cdot \mathbb{1}_{\{\xi_n - a < \varepsilon\}} \leqslant \sup_{|x-a|<\varepsilon} |f(x) - f(a)|$

Тогда $(*) \leqslant \sup_{|x-a|<\varepsilon} |f(x) - f(a)| + 2MP(|\xi_n - a| \geqslant \varepsilon)$.

$\overline{\lim}|\mathbb{E}f(\xi_n) - f(a)| \leqslant \sup_{|x-a|<\varepsilon} |f(x) - f(a)| + 2M\overline{\lim}P(|\xi_n - a| \geqslant \varepsilon) \leqslant \sup_{|x-a|<\varepsilon} |f(x) - f(a)| \to 0$ при $\varepsilon \to 0$.

Тогда $0 \leqslant \underline{\lim} \leqslant \overline{\lim} \leqslant 0 \Rightarrow \lim |\mathbb{E}f(\xi_n) - f(a)| = 0$ □

`Замечание.` В условии теоремы $|\mathbb{E}f(\xi_n) - f(a)| \leqslant \sup_{|x-a|<\varepsilon} |f(x) - f(a)| + 2MP(|\xi_n - a| \geqslant \varepsilon)$

Теорема 2.17. Вейерштрасса

$f \in C[a, b]$, то существует последовательность многочленов P_n, такая, что $P_n \rightrightarrows f$ на $[a, b]$

Доказательство. Можно считать, что всё на $[0, 1]$. Рассмотрим схему Бернулли с вероятностью успеха p. Тогда $\frac{S_n}{n} \to p$. Подставим $\xi_n = \frac{S_n}{n}$ в замечание.

$|\mathbb{E}f(\frac{S_n}{n}) - f(p)| \leqslant \sup_{|x-a|<\varepsilon} |f(x) - f(a)| + 2MP(|\frac{S_n}{n} - p| \geqslant \varepsilon) = (*)$

Из неравенства Чебышёва $P(|\frac{S_n}{n} - p| \geqslant \varepsilon) \leqslant \frac{\mathbb{D}\frac{S_n}{n}}{\varepsilon^2} = \frac{np(1-p)}{n^2\varepsilon^2} \leqslant \frac{1}{4n\varepsilon^2}$.

И тогда $(*) \leqslant \sup_{|x-y|<\varepsilon} |f(x) - f(y)| + \frac{M}{2n\varepsilon^2}$. При $n = \frac{1}{\varepsilon^3}$ правое слагаемое оценивается ε', а первое слагаемое мало из равномерной непрерывности.

Значит $\mathbb{E}f(\frac{S_n}{n}n) - f(p) \rightrightarrows 0$. $\mathbb{E}f(\frac{S_n}{n}) = \sum_{k=0}^n f(\frac{k}{n}) \cdot \binom{n}{k} p^k(1-p)^{n-k}$ - многочлен Бернштейна. □

***Определение* 2.26.** Многочлен Бернштейна $B_n(x) = \sum_{k=0}^n f(\frac{k}{n})\binom{n}{k}x^k(1-x)^{n-k}$

Следствие.

1. $B_n(0) = f(0)$
2. $B_n(1) = f(1)$
3. $B_n'(0) = n(f(\frac{1}{n}) - f(0))$
   $B_n'(1) = n(f(1) - f(\frac{n-1}{n}))$
   Доказательство: $B_n'(x) = \sum_{k=0}^n f(\frac{k}{n})\binom{n}{k}(kx^{k-1}(1-x)^{n-k} - (n-k)x^k(1-x)^{n-k-1}) =$
   $= \sum_{k=0}^n f(\frac{k}{n})\binom{n}{k}x^{k-1}(1-x)^{n-k-1}(k - nx)$
4. $B_n'(x) = \sum_{k=0}^n f(\frac{k}{n})f(\frac{k}{n})\binom{n}{k}x^{k-1}(1-x)^{n-k-1}(k - nx)$
5. $B_n(\alpha f + \beta g, x) = \alpha B_n(f, x) + \beta B_n(g, x)$

Кривые Безье

$\sum_{k=0}^n a_k\binom{n}{k}t^k(1-t)^{n-k}, a_k \in \mathbb{R}^2$. Получается отображение $\gamma : [0, 1] \to \mathbb{R}^2$.

1. $n = 1 : a(1-t) + bt$ - отрезок соединяющий точки a и b.
2. $n = 2 : a(1-t)^2 + 2bt(1-t) + ct^2$. Мы знаем, что $B'(0) = 2(b-a)$ и $B'(1) = 2(c-b)$. Это кривая из точки a в c, параметр b задаёт касательную в a и c.
3. $n = 3 : a(1-t)^3 + 3bt(1-t)^2 + 3ct^2(1-t) + dt^3$.

   Здесь $B(0) = a, B(1) = d, B'(0) = 3(b-a), B'(1) = 3(d-c)$. Кривая выходит из точки a с касательной $3(b-a)$, а заходит в точку d с касательной $3(d-c)$.

2.6. Производящие функции

***Определение* 2.27.** $\xi : \Omega \to \{1, 2, \ldots\}$ - случайная величина.

$G_\xi(z) = \sum_{n=0}^{\infty} P(\xi = n) z^n$ - производящая фукнция

Свойства.

1. G_ξ однозначно определяет распределение
2. $G_\xi(1) = 1$ и G_ξ сходится в круге $|z| < 1$.
3. $G_\xi(x) = \mathbb{E}x^\xi$, где $x \in \mathbb{R}$

   Доказательство: $\mathbb{E}x^\xi = \sum_{n=0}^{\infty} x^n \cdot P(\xi = n) = G_\xi(x)$
4. $G'_\xi(1) = \mathbb{E}\xi$

   Доказательство: $G'_\xi(x) = \sum_{n=1}^{\infty} P(\xi = n) n x^{n-1}$ - если подставить единицу - получим матожидание.
5. $\mathbb{E}\xi^2 = G''_\xi(1) + G'_\xi(1)$

   Доказательство: $G''_\xi(x) = \sum_{n=2}^{\infty} P(\xi = n) n(n-1) x^{n-2}$ - если подставить единицу - получим матожидание.
6. $\mathbb{D}\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2 = G''_\xi(1) + G'_\xi(1) - (G'_\xi(1))^2$
7. G_ξ возрастает и выпукла на $[0, 1]$
8. Если ξ и η независимы, то $G_{\xi+\eta}(z) = G_\xi(z) \cdot G_\eta(z)$

   Доказательство: x^ξ и x^η независимы, а тогда $\mathbb{E}(x^\xi \cdot x^\eta) = \mathbb{E}x^\xi \cdot \mathbb{E}x^\eta$

Пример.

1. Равномерное распределение на $\{0, 1, \ldots, n-1\}$.

   Тогда $G_\xi(z) = \frac{1}{n}(1 + z + z^2 + \ldots z^{n-1}) = \frac{1-z^n}{1-z} \cdot \frac{1}{n}$. Пусть хотим посчитать матожидание и диспресию, но единицу то подставить нельзя в свернутую формулу. Решается эта проблема так:

   Давайте скажем, что $z = 1 + y$. Тогда $G_\xi(1+y) = \frac{(1+y)^n - 1}{ny} = 1 + \binom{n}{2}\frac{y}{n} + \binom{n}{3}\frac{y^2}{n} \ldots$

   Тогда $G'_\xi(1) = \frac{\binom{n}{2}}{n} = \frac{n-1}{2}$, $\mathbb{E}\xi^2 = G''_\xi(1) + G'_\xi(1) = 2\frac{n(n-1)(n-2)}{6n} + \frac{n-1}{2} = \frac{n-1}{2}(\frac{2n-4}{3} + 1) = \frac{n-1}{2} \cdot \frac{2n-1}{3}$. И тогда $\mathbb{D}\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2 = \frac{n-1}{2} \cdot \frac{n+1}{6} = \frac{n^2-1}{12}$

2. **Задача Галилея**

   Есть 3 правильных кубика, бросили и посчитали сумму значений. Интересуемся вероятностью того, что в сумме выпало 10.

   $P(\text{в сумме } 10) = ?$

   ξ_i - значение на i-том кубике. Тогда $G_{\xi_i}(z) = \frac{1}{6}(z + z^2 + \ldots + z^6) = \frac{z(1-z^6)}{1-z} \cdot \frac{1}{6}$. Кубика у нас три, поэтому нас интересует $G_{\xi_1+\xi_2+\xi_3} = G_{\xi_1} \cdot G_{\xi_2} \cdot G_{\xi_3} = \left(\frac{z(1-z^6)}{1-z} \cdot \frac{1}{6}\right)^3 = (*)$

   $\frac{1}{(1-z)^3} = \sum_{n=0}^{\infty} \binom{n+2}{n} z^n$. Тогда $(*) = \frac{1}{6^3}(z^3 - 3z^9 + 3z^15 - z^21) \cdot \sum_{n=0}^{\infty} \binom{n+2}{n} z^n$. Коэффициент при z^{10} будет такой $\frac{1}{6^3}(1 \cdot \binom{9}{7} - 3 \cdot \binom{3}{1}) = \frac{1}{6^3}(36 - 3^2) = \frac{1}{8}$

3. Метод характеристических функций

3.1. Характеристические функции случайных величин

***Определение* 3.1.** Комплекснозначная случайная величина $\xi = \operatorname{Re}\xi + i\operatorname{Im}\xi$, где $\operatorname{Re}\xi$ и $\operatorname{Im}\xi$ вещественнозначные случайные величины.

***Определение* 3.2.** $\xi : \Omega \to \mathbb{C}$

$\mathbb{E}\xi = \mathbb{E}\operatorname{Re}\xi + i\mathbb{E}\operatorname{Im}\xi$

Свойства.

1. $\mathbb{E}(i\xi) = i\mathbb{E}\xi$
2. Комплексная линейность $\mathbb{E}(\alpha\xi + \beta\eta) = \alpha\mathbb{E}\xi + \beta\mathbb{E}\eta$, где $\alpha, \beta \in \mathbb{C}, \xi, \eta : \Omega \to \mathbb{C}$

   Доказательство: $\mathbb{E}(\alpha\xi) = \mathbb{E}(a+ib)\xi = \mathbb{E}(a\xi) + \mathbb{E}(b\xi i) = (a+bi)\mathbb{E}\xi$
3. $\overline{\mathbb{E}\xi} = \mathbb{E}\overline{\xi}$
4. $|\mathbb{E}\xi| \leqslant \mathbb{E}|\xi|$

   Доказательство: Возьмём $c \in \mathbb{C}, |c| = 1$, такой, что $\mathbb{E}(c\xi) = |\mathbb{E}\xi|$, то есть $c = \frac{\overline{\mathbb{E}\xi}}{|\mathbb{E}\xi|}$

   Тогда $|\mathbb{E}\xi| = \mathbb{E}(c\xi) = \mathbb{E}(\operatorname{Re}(c\xi)) \leqslant \mathbb{E}|\operatorname{Re}(c\xi)| \leqslant \mathbb{E}|c\xi| = \mathbb{E}|\xi|$

***Определение* 3.3.** Ковариация $cov(\xi, \eta) = \mathbb{E}(\xi - \mathbb{E}\xi)\overline{(\eta - \mathbb{E}\eta)}$

***Определение* 3.4.** Дисперсия $\mathbb{D}\xi = \mathbb{E}|\xi - \mathbb{E}\xi|^2$

$cov(\xi, \xi) = \mathbb{D}\xi$

***Определение* 3.5.** $\xi : \Omega \to \mathbb{R}$. Назовём характеристической функцией ξ:

$\phi_\xi(t) = \mathbb{E}e^{it\xi}$, где $t \in \mathbb{R}$

Свойства.

1. $\phi_\xi(0) = 1$ и $|\phi_\xi(t)| \leqslant 1$

   Доказательство: $|\phi_\xi(t)| \leqslant |\mathbb{E}e^{it\xi}| \leqslant \mathbb{E}|e^{it\xi}| = 1$
2. $\phi_{a\xi+b}(t) = e^{ibt}\phi_\xi(at)$

   Доказательство: $\phi_{a\xi+b}(t) = \mathbb{E}e^{i(a\xi+b)t} = \mathbb{E}e^{ibt}e^{i\xi at} = e^{ibt}\mathbb{E}e^{i\xi(at)} = \phi_\xi(at)e^{ibt}$
3. Если ξ и η независимы, то $\phi_{\xi+\eta}(t) = \phi_\xi(t) \cdot \phi_\eta(t)$

   Доказательство: $e^{i\xi t}$ и $e^{i\eta t}$ независимы и пишем произведение матожиданий
4. $\overline{\phi_\xi(t)} = \phi_\xi(-t)$

   Доказательство: $\overline{\phi_\xi(t)} = \overline{\mathbb{E}e^{i\xi t}} = \mathbb{E}\overline{e^{i\xi t}} = \mathbb{E}e^{-i\xi t} = \phi_\xi(-t)$
5. ϕ_ξ равномерно непрерывна на $\mathbb{R}$

   Доказательство: $|\phi_\xi(t+h) - \phi_\xi(t)| = |\mathbb{E}e^{i(t+h)\xi} - \mathbb{E}e^{it\xi}| = |\mathbb{E}\left(e^{it\xi}(e^{ih\xi} - 1)\right)| \leq$

   $\leq \mathbb{E}\left|e^{it\xi}(e^{ih\xi} - 1)\right| = \mathbb{E}|e^{ih\xi} - 1| = \int_{\mathbb{R}} |e^{ihx} - 1| dP_\xi(x) \to_{(*)} 0.$

   $(*)$: Знаем, что $e^{ihx} \to_{h\to 0} 1$, хотим понять, что можно вносить предел под знак интеграла. Это сделать можно по теорема Лебега, где суммируемая мажоранта будет 2.

Пример. $\xi \sim \mathcal{N}(a, \sigma^2)$. Хотим посчитать характеристическую функцию.

Возьмём $\eta \sim \mathcal{N}(0, 1)$. Тогда $\xi = \sigma\eta + a$ - имеет нужное нам распределение.

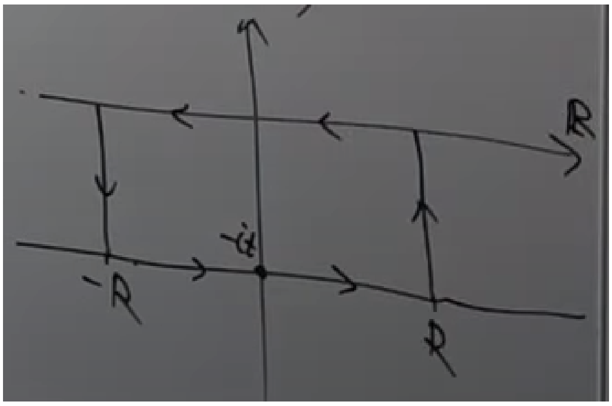
$\phi_{\sigma\eta+a}(t) = e^{ita}\phi_\eta(\sigma t)$

Считаем для η:

$\phi_\eta(t) = \mathbb{E}e^{i\eta t} = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{itx} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \int_{\mathbb{R}} e^{-\frac{(x-it)^2}{2}} dx.$

Посчитаем $I := \int_{\mathbb{R}} e^{-\frac{(x-it)^2}{2}} dx = \int_{-it+\mathbb{R}} e^{-\frac{z^2}{2}} dz = \int_{\Gamma_R} e^{-\frac{z^2}{2}} dz = 0.$

С другой стороны это также равно: $\int_{\Gamma_R} e^{-\frac{z^2}{2}} dz = \underbrace{-\int_{-R}^{R}}_{\to\sqrt{2\pi}} + \underbrace{\int_{-R-it}^{R-it}}_{\to I} + \underbrace{\int_{R-it}^{R}}_{(*):\to 0} + \underbrace{\int_{-R}^{-R-it}}_{(*):\to 0}$



$(*) : \left|\int_{R-it}^{R} e^{-\frac{z^2}{2}} dz\right| =_{z=R+iy} \left|i\int_{-t}^{0} e^{-\frac{(R+iy)^2}{2}} dy\right| \le \int_{-t}^{0} |\ldots| dy = \int_{-t}^{0} e^{\frac{-R^2+y^2}{2}} dy \le te^{\frac{t^2}{2}} e^{\frac{-R^2}{2}} \to_{R\to\infty} 0.$

Тогда получаем, что $I = \sqrt{2\pi}$, и $\phi_\eta(t) = e^{-\frac{t^2}{2}}$.

Теперь находим $\phi_\xi(t) = e^{iat}\phi_\eta(\sigma t) = e^{-\frac{\sigma^2 t^2}{2} + iat}$.

Теорема 3.1. Пусть $\mathbb{E}|\xi|^n < +\infty$.

Тогда при $k \leqslant n$ верно, что $\varphi^{(k)}(t) = \mathbb{E}((i\xi)^k e^{i\xi t})$.

В частности, $\varphi^{(k)}(0) = i^k \mathbb{E}\xi^k$.

Тут имеется в виду k-ая производная.

Следствие. Если $\mathbb{E}\xi^2 < +\infty$, то $\mathbb{E}\xi = -i\varphi'(0)$ и $\mathbb{D}\xi = -\varphi''(0) + (\varphi'(0))^2$

Доказательство. Теоремы.

Индукция по k

База $k = 0$: определение φ.

Переход $k \to k+1$:

$\varphi^{(k+1)}(t) = \lim_{h\to 0} \frac{\varphi^{(k)}(t+h) - \varphi^{(k)}(t)}{h} =$

$= \lim_{h\to 0} \frac{\mathbb{E}(i\xi)^k e^{i\xi(t+h)} - \mathbb{E}(i\xi)^k e^{i\xi(t)}}{h} = \lim_{h\to 0} \mathbb{E}((i\xi)^k e^{it\xi} \cdot \frac{e^{ih\xi}-1}{h}) = \mathbb{E}((i\xi)^k e^{it\xi} \cdot \lim_{h\to 0} \frac{e^{ih\xi}-1}{h})$, а предел – это $i\xi$.

Почему можно было запихать предел под матожидание?

$\lim \int_{\mathbb{R}} (ix)^k e^{itx} \frac{e^{ihx}-1}{h} dP_\xi(x) = \int_R \lim_{h\to 0}((ix)^k e^{itx} \frac{e^{ihx}-1}{h}) dP_\xi(x)$ – нужна суммируемая мажоранта.

$\left|(ix)^k e^{itx} \frac{e^{ihx}-1}{h}\right| = |x|^k \left|\frac{e^{ihx}-1}{h}\right| = (*).$

1. Если $|xh| \geqslant 1$, то $\left|\frac{e^{ihx}-1}{h}\right| \leqslant \frac{2}{|h|} \leqslant 2|x|$ и тогда $(*) \le 2|x|^{k+1}$.

2. Если $|xh| < 1$, то $e^{ihx} = 1 + \mathcal{O}(1 + ihx) \Rightarrow \left|\frac{e^{ihx}-1}{h}\right| = \left|\frac{\mathcal{O}(hx)}{h}\right| = \mathcal{O}(x)$ и тогда $(*) = \mathcal{O}(|x|^{k+1})$.

Но $\int_{\mathbb{R}} |x|^{k+1} dP_\xi(x) = \mathbb{E}|\xi|^{k+1} < +\infty$ по условию, тогда мажоранту подобрали правильную. $\square$

Теорема 3.2. Если существует $\varphi''_\xi(0)$ и конечна, то $\mathbb{E}\xi^2 < +\infty$

Замечание. Если существует $\varphi_\xi^{(2n)}$ и конечна, то $\mathbb{E}\xi^{2n} < +\infty$

Доказательство. $\mathbb{E}\xi^2 = \int_\mathbb{R} x^2 dP_\xi(x) = (*)$ – хотим доказать, что этот интеграл конечен.

Заметим, что $x = \lim_{t\to 0} \frac{\sin(tx)}{t}$ и подставим вместо x.

Тогда:

$(*) = \int_\mathbb{R} \lim_{t\to 0} \frac{\sin^2(tx)}{t^2} dP_\xi(x) \leqslant \underline{\lim}_{t\to 0} \int_\mathbb{R} -\frac{e^{2itx}+e^{-2itx}-2}{4t^2} dP_\xi(x) = (*)$ – лемма Фату и расписали синус.

$(*) = \underline{\lim}_{t\to 0} -\frac{\varphi_\xi(2t)-\varphi_\xi(-2t)-2}{4t^2} = (*)$.

Причём $\varphi_\xi(u) = \varphi_\xi(0) + \varphi'_\xi(0)\cdot u + \frac{\varphi''_\xi(0)u^2}{2} + o(u^2)$.

Тогда $\varphi_\xi(2t) + \varphi_\xi(-2t) = 2 + \frac{\varphi''_\xi(0)((2t)^2+(-2t)^2)}{2} + o(t^2)$, а тогда $(*) = \underline{\lim}_{t\to 0}(-\varphi''_\xi(0) + o(1))$ $\square$

Теорема 3.3. Формула обращения

Пусть $a < b$ и $P(\xi = a) = P(\xi = b) = 0$.

Тогда $P(\xi \in [a,b]) = \lim_{T\to+\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-iat}-e^{ibt}}{it} \varphi_\xi(t)\, dt$

То есть $v.p. \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-iat}-e^{ibt}}{it} \varphi_\xi(t)\, dt$

Доказательство. $\xi = \frac{a+b}{2} + \frac{b-a}{2}\eta$, тогда $P(\xi \in [a,b]) = P(\eta \in [-1,1])$, в частности $P_\eta(\{\pm 1\}) = 0$.

$\varphi_\xi(t) = e^{i\frac{a+b}{2}t}\varphi_\eta(\frac{b-a}{2}t)$ - подставим в наш интеграл.

$\int_{-T}^{T} \frac{e^{-iat}-e^{-ibt}}{it}\varphi_\xi(t)\, dt = \int_{-T}^{T} \frac{e^{-iat}-e^{-ibt}}{it} e^{i\frac{a+b}{2}t}\varphi_\eta(\frac{b-a}{2}t)\, dt =$

$= \int_{-T}^{T} \frac{e^{-i\frac{a-b}{2}t} - e^{-i\frac{b-a}{2}t}}{it}\varphi_\eta(\frac{b-a}{2})\, dt = \int_{-\frac{b-a}{2}T}^{\frac{b-a}{2}T} \frac{e^{is}-e^{-is}}{is}\varphi_\eta(s)\, ds$, здесь замена $s = \frac{b-a}{2}t$

Можно считать, что $a = -1$, а $b = 1$

$\int_{-T}^{T} \frac{e^{it}-e^{-it}}{it}\varphi_\xi(t)\, dt = \int_{-T}^{T}\int_\mathbb{R} \frac{e^{it}-e^{-it}}{it} e^{itx} dP_\xi(x)\, dt = (*)$ - давайте переставим местами интегралы.

Нужна суммируемость того, что под интегралом, а она есть, всё ограничено какой-то суммируемой константой.

$(*) = \int_\mathbb{R}\int_{-T}^{T} \frac{e^{it}-e^{-it}}{it} e^{itx}\, dt\, dP_\xi(x)$. Пусть $\Phi_T(x) = \int_{-T}^{T} \frac{e^{it}-e^{-it}}{it} e^{itx}\, dt$

$\lim_{T\to+\infty} \int_{-T}^{T} \frac{e^{it}-e^{-it}}{it}\varphi_\xi(t)\, dt = \lim_{T\to+\infty}\int_\mathbb{R} \Phi_T(x) dP_\xi(x) = \int_\mathbb{R} \lim_{T\to+\infty}\Phi_T(x)\, dP_\xi(x)$ - хотим понять, почему можно внести предел под интеграл, но разберемся с этим позже.

$\lim_{T\to+\infty}\Phi_T(x) = \lim_{T\to+\infty}\int_{-T}^{T}\int_{-1}^{1} e^{iut}\, du\, e^{itx}\, dt = \lim_{T\to+\infty}\int_{-1}^{1}\int_{-T}^{T} e^{it(u+x)}\, dt\, du = (*)$.

Заметим, что $\left.\frac{e^{it(u+x)}}{i(u+x)}\right|_{t=-T}^{t=+T} = \frac{2\sin((u+x)T)}{u+x}$

Тогда $(*) = \lim_{T\to+\infty}\int_{-1}^{1} \frac{2\sin((u+x)T)}{u+x}\, du = (*)$. Сделаем замену $y = (u+x)T$, тогда $dy = T\cdot du$.

Тогда $(*) = \lim_{T\to+\infty}\int_{(-1+x)T}^{(1+x)T} \frac{2\sin y}{y}\, dy = \begin{cases} 0, & \text{при } x > 1 \\ 0, & \text{при x < -1} \\ \int_\mathbb{R} \frac{2\sin y}{y}\, dy = 2\pi, & \text{иначе} \end{cases}$

Получили $2\pi\int_\mathbb{R} \mathbb{1}_{[-1,1]}(x)\, dP_\xi(x) = 2\pi P_\xi([-1,1])$.

Вспомним, что мы не доказали по дороге один переход. Нужно понять, почему $\int_a^b \frac{\sin y}{y}\, dy$ ограничен - интеграл по лучу сходится, значит первообразная в бесконечностях имеет предел, значит в середине тоже ограничена, потому что непрервность - обоснование примерно такое.

$\square$

Следствие. 1. Если $\varphi_\xi(t) = \varphi_\eta(t)$, то $P_\xi = P_\eta$

Доказательство: Рассмотрим $A = \{a \in \mathbb{R} : a$ - точка непрервности функции распределения$\}$.

Тогда $\mathbb{R} \setminus A$ - не более чем счётное. Если $a < b$ и $a, b \in A$, то $P_\xi([a,b]) = P_\eta([a,b])$

(a) Пусть $a \in \mathbb{R}, b \in A$:
Рассмотрим $a_n \in A$, такие, что $a_n \to a$ и убывают.
$P_\xi((a,b]) = \lim_{n\to\infty} P_\xi([a_n,b_n]) = \lim P_\eta([a_n,b_n]) = P_\eta((a,b])$.

(b) Пусть $a < b$ произвольные:
Возьмём $b_n \in A$, такие, что $b_n \to b$ и убывают. Тогда $P_\xi((a,b]) = \lim_{n\to\infty} P_\xi(a,b_n] = \lim P_\eta(a,b_n] = P_\eta(a,b] \Rightarrow P_\xi = P_\eta$ на ячейках, а тогда по единственности продолжения везде совпадают.

2. Если $\int_{\mathbb{R}} |\varphi_\xi(t)|\,dt < +\infty$, то ξ имеет плотность распределения $p_\xi(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}\varphi_\xi(t)\,dt$ - преобразование Фурье.

Доказательство: Из суммируемости $\varphi_\xi(t) \Rightarrow P_\xi((a,b]) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{e^{-iat}-e^{-ibt}}{it}\varphi_\xi(t)\,dt$.

Проверим, что $P_\xi(a,b] = \int_a^b p_\xi(x)\,dx$.

$\int_a^b p_\xi(x)\,dx = \frac{1}{2\pi}\int_a^b\int_{\mathbb{R}} e^{-itx}\varphi_\xi(t)\,dt\,dx = (*)$. Под внутренним интегралом суммируемая функция, значит можно переставлять местми интегралы.

Тогда $(*) = \frac{1}{2\pi}\int_{\mathbb{R}}\int_a^b e^{-itx}\,dx\varphi_\xi(t)\,dt$

Теорема 3.4. $\xi_k \sim \mathcal{N}(a_k, \sigma_k^2)$, $c_k \in \mathbb{R}$ не все нулевые и ξ_k – независимы и $\xi = a_0 + \sum_{k=1}^n c_k\xi_k$.
Тогда $\xi \sim \mathcal{N}(a, \sigma^2)$, где $a = a_0 + \sum_{k=1}^n c_k a_k$ и $\sigma^2 = \sum_{k=1}^n c_k^2\sigma_k^2$.

Доказательство. $\varphi_\xi(t) = \varphi_{a_0}(t)\varphi_{c_1\xi_1}(t)\ldots\varphi_{c_n\xi_n}(t) =$

$$= e^{ita_0}(t)\varphi_{\xi_1}(c_1t)\ldots\varphi_{\xi_n}(c_nt) = e^{ita_0}e^{ia_1c_1t}e^{-\frac{(c_1\sigma_1 t)^2}{2}}\ldots e^{ia_nc_nt}e^{-\frac{(c_n\sigma_n t)^2}{2}} = e^{ita}e^{-\frac{\sigma^2t^2}{2}}$$ $\square$

3.2. Сходимость по распределению

`Замечание.` 1. Точек, где нет непрерывности F_ξ не более чем счётное множество

2. Если $F_{\xi_n}(b) - F_{\xi_n}(a) \to F_\xi(b) - F_\xi(a)$ для всех a, b, за исключением счётного множества.
Тогда $F_{\xi_n}(b) \to F_\xi(b)$ за исключением счётного множества.

Доказательство. Рассмотрим $F(x)$, функцию распределения. Возьмём хорошие a, т.ч. $F(a) < \varepsilon$ и b, т.ч. $F(b) > 1 - \varepsilon$.

Тогда $(F_n(b) - F_n(a)) - (F(b) - F(a)) \to 0 \implies |(F_n(b) - F_n(a)) - \underbrace{(F(b) - F(a))}_{>1-2\varepsilon}| < \varepsilon \implies$

$F_n(b) - F_n(a) > 1 - 3\varepsilon \implies F_n(a) < 3\varepsilon$ при больших n.

Возьмём хорошее x, $|F_n(x) - F(x)| \leqslant |(F_n(x) - F_n(a)) - (F(x) - F(a))| + \underbrace{F_n(a)}_{<3\varepsilon} + \underbrace{F(a)}_{<\varepsilon} < 5\varepsilon$

при больших n $\square$

3. $D \subset \mathbb{R}$ не более чем счётное и $U \subset \mathbb{R}$ - открытое.
Тогда $U = \bigcup_{k=1}^\infty (a_k, B_k]$, где $a_k, b_k \notin D$

Доказательство. Нарезаем открытое множество с шагом 1, тем ячейки, которые целиком попали - берём. Те, что не попали - бьём пополам и так далее. □

4. ξ и η независимые и η имеет непрерывное распределение.

   Тогда $\xi + \eta$ имеет непрерывное распределение.

   Доказательство. $P_{\xi+\eta} = P_\xi * P_\eta$

   $P_{\xi+\eta}(\{a\}) = \int_{\mathbb{R}} \underbrace{P_\eta(\{a-x\})}_{=0,\text{т.к. непрерывность}} dP_\xi(x)$ □

***Определение* 3.6.** Множество $B \subset \mathbb{R}$ - регулярное, относительно P_ξ, если $P_\xi(Cl\, B \setminus Int\, B) = 0$, то есть $P(\xi \in Cl\, B \setminus Int\, B) = 0$

Теорема 3.5. $\xi, \xi_1, \xi_2, \ldots$ - случаный величины, $F, F_1, F_2, \ldots$ - их функции распределения, а $\varphi, \varphi_1, \varphi_2, \ldots$ - их характеристичечкие функции. Следующие условия равносильны:

1. ξ_n сходится к ξ по распределению
2. Для любого U открытого $\underline{\lim} P(\xi_n \in U) \geqslant P(\xi \in U)$
3. Для любого A замкнутого $\overline{\lim} P(\xi_n \in A) \leqslant P(\xi \in A)$
4. Для любого B регулярного борелевского $\lim P(\xi_n \in B) = P(\xi \in B)$
5. Для любого B регулярного борелевского $\lim \mathbb{E}\mathbb{1}_B(\xi_n) = \mathbb{E}\mathbb{1}_B(\xi)$
6. Для любой f непрерывной на прямой и ограниченной $\lim \mathbb{E}f(\xi_n) = \mathbb{E}f(\xi)$
7. φ_n сходится к φ поточечно

Доказательство.

1. $2 \iff 3$

   Если $A = \mathbb{R} \setminus U$, тогда $P(\xi_n \in A) = 1 - P(\xi_n \in U)$.

   $P(\xi \in A) > \overline{\lim} P(\xi_n \in A) = 1 - \underline{\lim} P(\xi_n \in U) \leqslant 1 - P(\xi \in U) = P(\xi \in A)$

2. $2 \cup 3 \implies 4$

   Мы знаем, что $U = \{\xi_n \in Int\, B\}\{\xi_n \in B\} \subset \{\xi_n \in Cl\, B\} = A$

   Тогда $P(\xi_n \in U) \leqslant P(\xi_n \in B) \leqslant P(\xi_n \in A) \implies \underbrace{\overline{\lim} P(\xi_n \in B)}_{\geqslant \underline{\lim} P(\xi_n \in B) \geqslant \underline{\lim} P(\xi_n \in U) \geqslant P(\xi \in U) = P(\xi \in B)} \leqslant$

   $\overline{\lim} P(\xi_n \in A) \leqslant P(\xi \in A) = P(\xi \in B)$

3. $4 \iff 5$

   $\mathbb{E}\mathbb{1}_B(\xi_n) = P(\mathbb{1}_B(\xi_n) = 1) = P(\xi_n \in B)$

4. $6 \implies 7$

   $\varphi_\eta(t) = \mathbb{E}e^{et\eta} = \mathbb{E}\cos(t\eta) + i\mathbb{E}\sin(t\eta)$

   Тогда $\varphi_n(t) = \mathbb{E}\cos(t\xi_n) + i\mathbb{E}\sin(t\xi_n) \to \mathbb{E}\cos(t\xi) + i\mathbb{E}\sin(t\xi) = \varphi(t)$

5. $1 \implies 2$

   Берём открытое U, по замечанию $U = \bigcup\limits_{k=1}^{\infty}(a_k, b_k]$, где a_k, b_k - точки непрерывности F.

   $\{\xi_n \in U\} \supset \{\xi_n \in \bigcup\limits_{k=1}^{m}(a_k, b_k]\} \implies P(\xi_n \in U) \geqslant \sum_{k=1}^{m} P(\xi_n \in (a_k, b_k])$

   $\underline{\lim} P(\xi_n \in U) \geqslant \underline{\lim} \sum_{k=1}^{m} \geqslant \sum_{k=1}^{m} \underline{\lim} P(\xi_n \in (a_k, b_k]) \overset{*}{=} \sum_{k=1}^{m} P(\xi \in (a_k b_k]) \overset{m\to\infty}{\to} \sum_{k=1}^{\infty} P(\xi \in (a_k, b_k]) = P(\xi \in U)$

   А значит $\underline{\lim} P(\xi_n \in U) \geqslant P(\xi \in U)$

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

6. $5 \implies 6$

   Пусть $|f| \leqslant M$ и $D = \{x \in \mathbb{R} : P(f(\xi) = x) > 0\} = \{x : P_\xi(f^{-1}(x) > 0)\}$. Это не более чем счётное множество. Потому что для разных x - это дизъюнктные. Множеств с вероятностью $\frac{1}{2}$ - не больше двух, с вероятностью $\frac{1}{3}$ не больше трёх и так далее.

   Пусть $-M = t_0 < t_1 < \ldots < t_m = M$, так, что $t_y \notin D$ и мелкость $< \varepsilon$.

   Заведём множества $A_j = \{x \in \mathbb{R} : t_{j-1} \leqslant f(x) \leqslant t_j\} \supset B_j = \{x \in \mathbb{R} : t_{j-1} < f(x) \leqslant t_j\} \supset U_j = \{x \in \mathbb{R} : t_{j-1} < f(x) < t_j\}$. Где A_j - замкнутое,а U_j - открытое.

   Мы поняли, что $U_j \subset Int\, B_j \subset B_j \subset Cl\, B_j \subset A_j \implies Cl\, B_j \setminus Int\, B_j \subset A_j \setminus U_j$

   Тогда B_j регулярно относительно P_ξ

   Определим $g(x) = \sum_{j=1}^{m} t_{j-1} \mathbb{1}_{B_j}(x)$. Тогда $g(x) < f(x) < g(x) + \varepsilon$.

   $|g(x) - f(x)| < \varepsilon$ и тогда $\mathbb{E}|g(\xi_n) - f(\xi_n)| \leqslant \varepsilon$ и мы знаем, что $\mathbb{E}g(\xi_n) \to \mathbb{E}g(\xi)$ - видно, если расписать матожидание g по линейности.

   $\mathbb{E}|f(\xi_n) - f(\xi)| \leqslant |\mathbb{E}f(\xi_n) - \mathbb{E}g(\xi_n)| + |\mathbb{E}g(\xi_n) - \mathbb{E}g(\xi)| + |\mathbb{E}g(\xi) - \mathbb{E}f(\xi)| < 3\varepsilon$ при больших n, каждый из модулей $< \varepsilon$

7. $7 \implies 1$

   Возьмём $\eta \sim \mathcal{N}(0, \sigma^2)$, такую, что η не зависит от всех ξ_n и ξ

   $\varphi_{\xi_n+\eta}(t) = \varphi_{\xi_n}(t)\varphi_\eta(t) = \varphi_n(t) \cdot e^{-\frac{\sigma^2 t^2}{2}} \overset{\text{поточечно}}{\to} \varphi(t) e^{-\frac{\sigma^2 t^2}{2}} = \varphi_{\xi+\eta}(t)$

   $\xi_n + \eta$ и $\xi + \eta$ имеют непрерывное распределение, поэтому можем не задумываясь писать формулу обращения:

   $P(\xi_n + \eta \in (a, b]) = \frac{1}{2\pi} \int_{\mathbb{R}} \frac{e^{-iat} - e^{-ibt}}{it} \varphi_{\xi_n+\eta}(t)\, dt \overset{*}{\to} \frac{1}{2\pi} \int_{\mathbb{R}} \frac{e^{-iat} - e^{-ibt}}{it} \varphi_{\xi+\eta}\, dt = P(\xi + \eta \in (a, b])$.

   $(*)$ Нужна суммируемая мажоранта $\left|\frac{e^{-iat} - e^{-ibt}}{it} \varphi_{\xi_n+\eta}(t)\right| \leqslant e^{-\frac{\sigma^2 t^2}{2}}$ - суммируемая мажоранта.

   То есть $\underbrace{P(\xi_n + \eta \in (a, b])}_{G_n(b) - G_n(a)} \to \underbrace{P(\xi + \eta \in (a, b])}_{G(b) - G(a)}$, где $G_n(x) = F_{\xi_n+\eta}(x)$ и $G(x) = F_{\xi+\eta}(x)$

   Тогда из замечания $G_n(x) \to G(x)$

   Возьмём x - точка непрерывности F и выберем $\delta > 0$, так, что $|F(x \pm \delta) - F(x)| < \varepsilon$ - есть из непрерывности.

   $\{\xi_n + \eta \leqslant x - \delta\} \setminus \{|\eta| > \delta\} \subset \{\xi_n \leqslant x\} \subset \{\xi_n + \eta \leqslant x + \delta\} \cup \{|\eta| > \delta\}$.

   Тогда $\underbrace{G_n(x - \delta) - P(|\eta| > \delta)}_{G_n(x-\delta-\frac{\sigma^2}{\delta^2}) > G_n(x-\delta)-\varepsilon > G(x-\delta)-2\varepsilon > F(x-2\delta)-3\varepsilon > F(x)-4\varepsilon} \leqslant F_n(x) \leqslant \underbrace{G_n(x + \delta) + P(|\eta| > \delta)}_{G_n(x+\delta)+\frac{\sigma^2}{\delta^2} < G_n(x+\delta)+\varepsilon < G(x+\delta)+2\varepsilon < F(x+2\delta)}$

   Оценим вероятность: $P(|\eta| > \delta) \leqslant \frac{\mathbb{D}\eta}{\delta^2} = \frac{\sigma^2}{\delta^2}$

   $\square$

Теорема 3.6. $F_n, F\,\mathbb{R} \to [0,1]$ монотонные, $F \in C(\mathbb{R})$ и $F_n \to F$ поточечно.

Тогда $F_n \rightrightarrows F$

Доказательство. Берём $\varepsilon > 0$ и $m > \frac{1}{\varepsilon}$. Пусть t_j, такие, что $F(t_j) = \frac{j}{m}$. Если для большого j точки не нашлось, то $F < \frac{j+1}{m}$, а если для маленького не нашлось, то $F > \frac{j-1}{m}$ (потому что иначе из непрерывности такие точки найдутся).

Знаем, что $F_n(t_j) \to F(t_j)$. Берём $N\,:\,\forall\, n \geqslant N\, |F_n(t_j) - F(t_j)| < \varepsilon$.

Теперь смотрим на произвольную точку: $t_j < t < t_{j+1}$. Тогда $F_n(t) \leqslant F_n(t_{j+1}) < F(t_{j+1}) + \varepsilon = \frac{j+1}{m} + \varepsilon = F(t_j) + \frac{1}{m} + \varepsilon \leqslant F(t) + \frac{1}{m} + \varepsilon < F(t) + 2\varepsilon$.

Аналогично в другую сторону:

$F_n(t_j) > F(t_j) - \varepsilon = \frac{j}{m} - \varepsilon = F(t_{j+1}) - \frac{1}{m} - \varepsilon \geqslant F(t) - 2\varepsilon$.

□

3.3. Центральная предельная теорема

Теорема 3.7. ЦПТ в форме Леви

$\xi_1, \xi_2, \ldots$ независимые, одинаково распределённые случайные величины. $a = \mathbb{E}\xi_1, \sigma^2 = \mathbb{D}\xi_1, S_n = \xi_1 + \xi_2 + \ldots + \xi_n$.

Тогда $P\left(\frac{S_n - nq}{\sigma\sqrt{n}} \leqslant x\right) = P\left(\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}} \leqslant x\right) \rightrightarrows \Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt$

Доказательство. Достаточно проверять поточечную сходимость характеристических функций.

$\varphi_{\xi_k - a}(t) = 1 - \frac{\sigma^2 t^2}{2} + o(t^2)$, потому что мы знаем, что $\mathbb{E}(\xi_k - a) = 0$ и $\mathbb{D}(\xi_k - a)\sigma^2$

Тогда $\varphi_{S_n - na}(t) = \prod\limits_{k=1}^{n} \varphi_{\xi_k - a}(t) = \left(1 - \frac{\sigma^2 t^2}{2} + o(t^2)\right)^n$

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

Чтобы получить последний переход - логарифмируем.

То есть $\frac{S_n - na}{\sqrt{n}\sigma}$ сходится по распределению к $\mathcal{N}(0,1) \implies$ функция распределения сходится равномерно. □

Следствие. **Интегральная теорема Муавра-Лапласа**

S_n - количество успехов в схеме Бернулли с вероятностью успеха $p \in (0,1)$.

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

Доказательство. $\mathbb{E}\xi_k = p, \mathbb{D}\xi_k = pq$ □

Пример. Посчитаем характеристическую функцию для Poisson(λ)

$\varphi_\xi(t) = \mathbb{E}e^{it\xi} = \sum_{n=0}^{\infty} e^{itn}\frac{e^{-\lambda}\lambda^n}{n!} = \sum_{n=0}^{\infty}\frac{(\lambda e^{it})^n}{n!}\cdot e^{-\lambda} = exp(\lambda(e^{it} - 1))$

Теорема 3.8. $P(\xi_{nk} = 1) = p_{nk}$ и $P(\xi_{nk} = 0) = 1 - p_{nk}$ и пусть $S_n = \xi_{n1} + \ldots + \xi_{nn}$.

$\max\{p_{n1}, \ldots, p_{nn}\} \to 0$ и $p_{n1} + \ldots p_{nn} \to \lambda$.

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

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

Тогда $\varphi_{S_n} = \prod_{k=1}^{n}(1 + p_{nk}(e^{it} - 1)) \stackrel{?}{\to} exp(\lambda(e^{it} - 1))$.

Пусть $z = e^{it} - 1$

Значит, нужно проверить $\sum_{k=1}^n \ln(1 + p_{nk}z) \to \lambda z$. Раскладываем логарифм:

$\underbrace{\sum_{k=1}^n p_{nk}z}_{\to\lambda z} + \sum_{k=1}^n \mathcal{O}(p_{nk}^2)$. Значит осталось показать, что вторая сумма стремится к нулю.

Действительно, $\sum_{k=1}^n \mathcal{O}(p_{nk}^2) \leqslant (p_{n1} + \dots p_{nn}) \cdot \max\{p_{n1}, \dots, p_{nn}\} \to 0$ □

Теорема 3.9. ЦПТ в форме Линденберга

$\xi_1, \dots$ - независимые случайные величины, $a_k = \mathbb{E}\xi_k, \sigma_k^2 = \mathbb{D}\xi_k > 0, S_n = \xi_1 + \ldots + \xi_n, \mathbb{D}_n^2 = \sum_{k=1}^n \sigma_k^2$

$f(x) = x^2 \mathbb{1}_{\{|x| \geqslant \varepsilon \mathbb{D}_n\}}(x) \,\forall\, \varepsilon > 0$ и $Lind(\varepsilon, n) = \frac{1}{\mathbb{D}_n^2} \sum_{k=1}^n \mathbb{E}f(\xi_k - a_k) \to 0$

Тогда $P\left(\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}} \leqslant x\right) \rightrightarrows \Phi(x)$

Теорема 3.10. ЦПТ в форме Ляпунова

$\xi_1, \dots$ независимые случайные величины

$L(\delta, n) = \frac{1}{\mathbb{D}_n^{2+\delta}} \sum_{k=1}^n \mathbb{E}|\xi_k - a_k|^{2+\delta} \to 0$ при некотором δ

Тогда $P\left(\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}} \leqslant x\right) \rightrightarrows \Phi(x)$

Доказательство. □

Теорема 3.11. Пусть $\delta \in [0, 1]$ и $\xi_1, \dots$ независимые случаные величины.

Тогда $\left|P\left(\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}} \leqslant x\right) - \Phi(x)\right| \leqslant C_\delta L(\delta, n)$

Теорема 3.12. Берри-Эссенна

$\xi_1, \dots$ независимые, одинаково распределённые случаные величины.

Тогда$\left|P\left(\frac{S_n - na}{\sqrt{\sqrt{n}\sigma}} \leqslant x\right) - \Phi(x)\right| \leqslant C_\delta \cdot \frac{\mathbb{E}|\xi_1 - a|^{2+\delta}}{n^{\frac{\delta}{2}} \sigma^{2+\delta}}$

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

В частности при $\delta = 1$ $|P - \Phi| \leqslant C \frac{\mathbb{E}|\xi_1 - a|^3}{\sqrt{n}\sigma^3}$ □

Замечание. **Про константы**

1. Эссен (1956) $C \geqslant \frac{3+\sqrt{10}}{6\sqrt{2\pi}} \approx 0,4097$
2. Шевцова (2014) $C \leqslant 0,469$
3. Для общего случая: $C_1 \leqslant 0,5583$
4. Для схемы Бернулли (2018) $C_1 \leqslant 0,4099$
5. Для схемы Бернулли с $p = \frac{1}{2}$ $C = \frac{1}{\sqrt{2\pi}}$

Теорема 3.13. Хартмана-Винтнера (закон повторного логарифма)

$\xi_1, \xi_2, \dots$ - независимые, одинаково распределённые. $\mathbb{E}\xi_1 = 0$, $\mathbb{D}\xi_1 = \sigma^2 > 0$

Тогда $\overline{\lim} \frac{S_n}{\sqrt{2n \ln\ln n}} = \sigma$, а $\underline{\lim} \frac{S_n}{\sqrt{2n \ln\ln n}} = -\sigma$

Теорема 3.14. Штрассена

$\xi_1, \xi_2 \ldots$ независимые одинаково распределённые, $\mathbb{E}\xi_1 = 0$ и $\mathbb{D}\xi_1 = \sigma^2 > 0$

Тогда любое число из $[-\sigma, \sigma]$ - пред. точка послед. $\frac{S_n}{\sqrt{2n \ln \ln n}}$

3.4. Большие уклонения

ЗБЧ в форме Чебышёва

$\xi_1, \xi_2, \ldots$ - незаисимые, одинаково распределённые случайные величины. $\mathbb{E}\xi_1 = a, \mathbb{D} = \sigma^2 > 0$

Тогда $P\left(\frac{S_n}{n} \geqslant r\right) \to 0$, если $r > a$

Более того, $P\left(\frac{S_n}{n} \geqslant r\right) \leqslant \frac{\sigma^2}{(r-a)n}$ - это оценка из доказательства. Но оценка довольно плохая.

***Определение* 3.7.** ξ удовлетворяет условию Крамера, если при $\lambda \in (0, \lambda_0)$: $\mathbb{E}e^{\lambda\xi} < +\infty$

Теорема 3.15. Оценка Чернова

$\xi_1, \xi_2, \ldots$ - независимые, одинаково распределённые, удовлетворяющие условию Крамера и $r > a = \mathbb{E}\xi_1$

Хотим оценить $P\left(\frac{S_n}{n} \geqslant r\right) = P\left(S_n \geqslant nr\right) = P\left(\lambda S_n \geqslant \lambda nr\right) = P\left(e^{\lambda S_n} \geqslant e^{\lambda nr}\right) \overset{\text{Марков}}{\leqslant} \frac{\mathbb{E}e^{\lambda S_n}}{e^{\lambda nr}} \overset{*}{=} \left(\frac{\mathbb{E}e^{\lambda\xi_1}}{e^{\lambda r}}\right)^n$

$(*)\mathbb{E}e^{\lambda S_n} = \mathbb{E}(\prod_{k=1}^n e^{\lambda\xi_k}) = (\mathbb{E}e^{\lambda\xi_1})^n$

А теперь, то что получилось - будем минимизировать по λ

$\varphi(\lambda) = \ln \mathbb{E}e^{\lambda\xi_1} - \lambda r \to \inf$, где λ допустимое

$I(r) = \sup_\lambda\{\lambda r - \ln(\mathbb{E}e^{\lambda\xi_1})\} \implies P\left(\frac{S_n}{n} \geqslant r\right) \leqslant e^{-nI(r)}$

Пример. 1. $\xi \sim \mathcal{N}(0,1)$

$\mathbb{E}e^{\lambda\xi} = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{\lambda x}e^{\frac{x^2}{2}}\,dx = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{(x-\lambda)^2}{2}}\,dx \cdot e^{\frac{\lambda^2}{2}} = e^{\frac{\lambda^2}{2}}$

$\lambda r - \frac{\lambda^2}{2} \to \max\limits_{\lambda > 0}$, максимум при $\lambda = r$. И тогда $I(r) = \frac{r^2}{2}$

Значит $P\left(\frac{S_n}{n} \geqslant r\right) \leqslant e^{-\frac{nr^2}{2}}$ при $r > 0$

2. $\xi \sim Exp(1)$

$\mathbb{E}e^{\lambda\xi} = \int_0^{+\infty} e^{\lambda x}e^{-x}\,dx = \frac{e^{(\lambda-1)x}}{\lambda-1}\Bigg|_0^{+\infty} = \frac{1}{1-\lambda}$ сходимость есть при $\lambda \in (0,1)$

$\lambda r - \ln\frac{1}{1-\lambda} = \lambda r + \ln(1-\lambda)$ - считаем производную по λ

$(\lambda r + \ln(1-\lambda))'_\lambda = r - \frac{1}{1-\lambda} \implies 1 - \lambda = \frac{1}{r} \implies \lambda = 1 - \frac{1}{r}, r > 1$

Тогда $I(r) = r(1 - \frac{1}{r}) - \ln r = r - 1 - \ln r$

$P\left(\frac{S_n}{n} \geqslant r\right) \leqslant e^{-n(r-1-\ln r)} = r^n e^{-n(r-1)}$

4. Дискретные случайные процессы

4.1. Условные математические ожидания

***Определение* 4.1.** $(\Omega, \mathcal{F}, P)$ - вероятностное пространство, $\xi : \Omega \to \mathbb{R}$ и $\mathbb{E}|\xi| < +\infty$. Пусть $\mathcal{A} \subset \mathcal{F}$ и $\mathcal{A}$ - $\sigma-$алгебра.

$\mathbb{E}(\xi|\mathcal{A}) = \eta : \Omega \to \mathbb{R}$ - случайная величина, которая:

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

Теорема 4.1. $\mathcal{A} \subset \mathcal{F}, \mathbb{E}|\xi| < +\infty$, тогда $\mathbb{E}(\xi|\mathcal{A})$ существует и единственно, с точностью до почти наверное

Доказательство. Существование:

$\xi = \xi_+ - \xi_-$. Пусть $A \in \mathcal{A}$, определим $\mu_\pm A = \int_A \xi_\pm\, dP$ - это конечные меры на $\mathcal{A}$, так как интеграл от измеримой неотрицательной функции. А ещё эти меры абсолютно непрерывны относительно $P \overset{\text{т. Радона-Никодима}}{\Longrightarrow} \exists\, \eta_\pm > 0$ измеримые относительно $\mathcal{A}$, т.ч. $\mu_\pm A = \int_A \eta_\pm\, dP$

$\eta = \eta_+ - \eta_-$, надо проверить, что $\forall\, A \in \mathcal{A} : \underbrace{\mathbb{E}(\xi \mathbb{1}_A)}_{\mathbb{E}(\xi_+ \mathbb{1}_A) - \mathbb{E}(\xi_- \mathbb{1}_A)} = \mathbb{E}(\eta \mathbb{1}_A)$

А ещё $\mathbb{E}(\xi_+ \mathbb{1}_A) = \mu_+ A = \int_A \xi_+\, dP$ и для остальных точно также

Единственность:

Пусть η_1 и η_2 - условные матожидания. Тогда $\{\eta_1 > \eta_2\} \in \mathcal{A}$

$\mathbb{E}(\eta_1 \mathbb{1}_A) = \mathbb{E}(\xi \mathbb{1}_A) = \mathbb{E}(\eta_2 \mathbb{1}_A) \implies \underbrace{\mathbb{E}((\eta_1 - \eta_2)\mathbb{1}_A)}_{=\int_A (\eta_1 - \eta_2)\, dP} = 0 \implies P(A) = P(\eta_1 > \eta_2) = 0$. Аналогично $P(\eta_1 < \eta_2) = 0$ □

Свойства.

1. $\mathbb{E}(c|\mathcal{A}) = c$
2. $\mathbb{E}(\xi|\mathcal{A})$ линейно по ξ
3. $\xi \leqslant \eta$, то $\mathbb{E}(\xi|\mathcal{A}) \leqslant \mathbb{E}(\eta|\mathcal{A})$

   Доказательство. Достаточно проверить, что если $\xi \geqslant 0$, то $\mathbb{E}(\xi|\mathcal{A}) \geqslant 0$ □

4. $\mathbb{E}(\xi|\{\varnothing, \Omega\}) = \mathbb{E}\xi$

   Доказательство. Измеримы относительно такой σ-алгебры только константы. Надо проверить, что $\mathbb{E}(\mathbb{E}\xi \mathbb{1}_A) = \mathbb{E}(\xi \mathbb{1}_A)$ для $A = \varnothing$ и $A = \Omega$ □

5. $\mathcal{F} \supset \mathcal{A}_1 \supset \mathcal{A}_2$ - σ-алгебры

   Тогда $\mathbb{E}(\mathbb{E}(\xi|\mathcal{A}_1)|\mathcal{A}_2) = \mathbb{E}(\xi|\mathcal{A}_2)$

   Доказательство. $\eta = \mathbb{E}(\xi|\mathcal{A}_2)$ и $\zeta = \mathbb{E}(\xi|\mathcal{A}_1)$

   Надо доказать, что $\eta = \mathbb{E}(\zeta|\mathcal{A}_2)$. η измерима относительно $\mathcal{A}_\in$. Надо проверить, что $\forall\, A \in \mathcal{A}_2 : \mathbb{E}(\eta \mathbb{1}_A) = \mathbb{E}(\zeta \mathbb{1}_A) = \mathbb{E}(\xi \mathbb{1}_A)$, т.к. $A \in \mathcal{A}_1$ по определению ζ. А ещё $\mathbb{E}(\xi \mathbb{1}_A) = \mathbb{E}(\eta \mathbb{1}_A)$ □

6. $\mathbb{E}(\mathbb{E}(\xi|\mathcal{A})) = \mathbb{E}\xi$ - из 4 и 5
7. Если ξ измерима относительно $\mathcal{A}$, о $\mathbb{E}(\xi|\mathcal{A}) = \xi$

Пример. Пусть $\Omega = \bigcup A_k$ не более чем счётное объединение

$\mathcal{A}$ - натянутая на $A_1, A_2, \ldots$ $\sigma-$алгебра

$\mathbb{E}(\xi|\mathcal{A}) =?$

Если η измерима относительно $\mathcal{A} \implies \eta = \sum c_k \mathbb{1}_{A_k}$

Нужно чтобы $\mathbb{E}(\xi \mathbb{1}_{A_n}) = \mathbb{E}(\underbrace{\eta \mathbb{1}_{A_n}}_{c_n \mathbb{1}_{A_n}}) = c_n P(A_n)$

То есть $c_n = \frac{\mathbb{E}(\xi \mathbb{1}_{A_n})}{P(A_n)}$

Замечание. Из свойства 6: $\mathbb{E}\xi = \mathbb{E}\eta = \sum \frac{\mathbb{E}(\xi \mathbb{1}_{A_k})}{P(A_k)} \cdot P(A_k)$

***Определение* 4.2.** Условная вероятность относительно σ-алгебры

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

Пример. $\xi_1, \xi_2, \ldots$ - независимые, одинаково распределённые случайные величины, N - случайная величина с неотрицательными целыми значениями, не зависящая от $\xi_1, \ldots$

$S = \xi_1 + \ldots + \xi_N$

Пусть $A_n = \{N = n\}$

$\mathbb{E}S = \sum\limits_{n=0}^{+\infty} \frac{\mathbb{E}(S\mathbb{1}_{A_n})}{P(A_n)} \cdot P(A_n) = \sum\limits_{n=0}^{\infty} naP(N=n) = \mathbb{E}\xi_1 \cdot \mathbb{E}N$

$\frac{\mathbb{E}(S\mathbb{1}_{A_n})}{P(A_n)} = \mathbb{E}(S|N=n) = \mathbb{E}(S_n|N=n) = \mathbb{E}(\xi_1 + \ldots + \xi_n|N=n) \overset{\text{незаисимость}}{=} = \mathbb{E}(\xi_1 + \ldots + \xi_n) = na$, где $a = \mathbb{E}\xi_1$

Пример. Пусть ξ_k тоже принимают неотрицательные целые значения

Тогда $G_{\xi_1}(t) = G(t)$ - производящая функция ξ_k

$G_S(t) = \mathbb{E}t^S = \sum_{n=0}^{\infty} \frac{\mathbb{E}(t^S \mathbb{1}_{A_n})}{P(A_n)} \cdot P(A_n) = \sum\limits_{n=0}^{\infty} G^n(t) P(N=n) = G_n(G_\xi(t))$

$\frac{\mathbb{E}(t^S \mathbb{1}_{A_n})}{P(A_n)} = \mathbb{E}(t^S|N=n) = \mathbb{E}(t^{S_n}|N=n) = \mathbb{E}(t^{\xi_1} \cdot \ldots \cdot t^{\xi_n}|N=n) = \mathbb{E}(t^{\xi_1} \cdot \ldots \cdot e^{\xi_n}) = (\mathbb{E}t^{\xi_1})^n = (G(t))^n$

Замечание. **Геометрическая интерпретация**

Пусть $\mathbb{E}\xi^2 < +\infty$

Тогда $\xi \in L^2(\Omega, \mathcal{F}, P)$. $\mathcal{A} \subset \mathcal{F}$ - σ - алгебра. Поэтому $\underbrace{L^2(\Omega, \mathcal{A}, P)}_{\text{замкнутое подпространство}} \subset L^2(\Omega, \mathcal{F}, P)$

$\eta = \mathbb{E}(\xi|\mathcal{A})$ - тогда проекция ξ на $L^2(\Omega, \mathcal{A}, P)$

Нужно проверить, что $\xi - \eta \perp L^2(\Omega, \mathcal{A}, P)$

В L^2 плотны ступенчатые, давайте проверим для них, а потом сделаем предельный переход. Достаточно даже понять только для 1 ступеньки.

Достаточно понять, что $\forall A \in \mathcal{A} : \xi - \eta \perp \mathbb{1}_A$

$0 \overset{?}{=} \langle \xi - \eta, \mathbb{1}_A \rangle = \mathbb{E}((\xi - \eta)\mathbb{1}_A) = \mathbb{E}(\xi \mathbb{1}_A) - \mathbb{E}(\eta \mathbb{1}_A)$

***Определение* 4.3.** η - случайная величина. Пусть $\sigma(\eta)$ - наименьшая σ-алгебра, относительно которой η измерима

Замечание. Чтобы её получить, нужно взять все Лебеговы множества и натянуть на них σ-алгебру

***Определение* 4.4.** $\mathbb{E}(\xi|\eta) = \mathbb{E}(\xi|\sigma(\eta))$ - условное матожидание ξ относительно η

Пример. η - дискретная, $\{y_1, y_2, \ldots\}$ - множество её значений

Все $\{\eta = y_k\}$ - измеримы, $\Omega = \bigcup\{\eta = y_k\}$, $\sigma(\eta)$ - всевозможные объединения $\{\eta = y_k\}$

Теорема 4.2. 1. Если ξ и 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. Надо доказать, что $\forall\, A \in \sigma(\eta)\ :\ \mathbb{E}(\xi\mathbb{1}_A) \stackrel{?}{=} \mathbb{E}(\mathbb{E}\xi \cdot \mathbb{1}_A) = \mathbb{E}\xi\mathbb{E}\mathbb{1}_A$

То есть достаоточно проверить, что ξ и $\mathbb{1}_A$ независимы

Пусть $A = \{\eta \leqslant a\}$. $P(\xi \in B, \mathbb{1}_A \in C) = P(\xi \in B) \cdot P(\mathbb{1}_A \in C)$

Достаточно рассмотреть только $C = \{0\}$ и $C = \{1\}$

$P(\xi \in B, \mathbb{1}_A = 1) = P(\xi \in B, \eta \leqslant a) \stackrel{\text{независимость } \xi \text{ и } \eta}{=} P(\xi \in B)P(\eta \leqslant a) = P(\xi \in B)P(\mathbb{1}_A = 1)$. Для $\mathbb{1}_A = 0$ аналогично

Для Лебеговых множеств мы это получили, поэтому есть и для любых праобразов ячеек

2. Проверяем для $\eta = \mathbb{1}_A$, где $A \in \mathcal{A}$

$\mathbb{1}_A\mathbb{E}(\xi|\mathcal{A})$ - условное матожидание $\mathbb{E}(\xi\mathbb{1}_A|\mathcal{A})$

Измеримость есть, поэтому достаточно проверить только второе условие

$\forall\, B \in \mathcal{A}\ :\ \underbrace{\mathbb{E}(\xi\mathbb{1}_A \cdot \mathbb{1}_B)}_{=\mathbb{E}(\xi\mathbb{1}_{A\cap B})} \stackrel{?}{=} \underbrace{\mathbb{E}(\mathbb{1}_A\mathbb{E}(\xi|\mathcal{A})\mathbb{1}_B)}_{=\mathbb{E}(\mathbb{1}_{A\cap B}\mathbb{E}(\xi|\mathcal{A}))}$

Тогда по линейности верно для простых η, по теореме Леви предельный переход $\eta_n \to \eta$ поточечно.

Мы знаем, что есть равенство $\underbrace{\mathbb{E}(\xi\eta_n\mathbb{1}_B)}_{=\int_\Omega \xi\eta_n\mathbb{1}_B\, dP} = \underbrace{\mathbb{E}(\eta_n\mathbb{E}(\xi|\mathcal{A})\mathbb{1}_B)}_{=\int_\Omega \eta_n\mathbb{E}(\xi|\mathcal{A}\mathbb{1}_B\, dP)}$

Предельный переход можно делать для η_+ и η_-, сделаем, потом перейдём к η

$\square$

4.2. Ветвящиеся процессы

ξ_{nk} - независимые случаные величины с неотрицательными целыми значениями

Интерпретация - есть много частиц, которые размножаются/умирают. Тогда n - момент времени, k - номер частицы, ξ_{nk} - количество её потомков

$\eta_n = \xi_{n1} + \xi_{n2} + \ldots + \xi_{n\eta_{n-1}}$ - количество частиц в момент n

$\eta_0 = 1$ - изначально у нас есть только 1 частица

Считаем, что все ξ_{nk} одинаково распределены и $P(\xi_{nk} = m) = f_m$.

$F(t) = \sum\limits_{m=0}^{\infty} f_m t^m$ - производящая функция

Пусть $G_n(t)$ - производящая функция для η_n. Тогда $G_n(t) = G_{n-1}(F(t)) = F \circ F \circ F \ldots \circ F(t)$ - результат был получен в примере выше.

$\mathbb{E}\eta_n = G_n'(1) = G_{n-1}'(\underbrace{F(1)}_{=1}) \cdot \underbrace{F'(1)}_{=\mathbb{E}\xi} = \mathbb{E}\xi \cdot \mathbb{E}\eta_{n-1} = (\mathbb{E}\xi)^n$

Теорема 4.3. Вероятность вырождения процесса - наименьший неотрицательный корень уравнения $F(x) = x$

Доказательство. $A_n = \{\eta_n = 0\}$ - на n-ном шаге не осталось частиц

$P(A_n) = G_n(0) \leqslant 1$, а ещё $A_1 \subset A_2 \subset \ldots$ - если процесс выродился, то он и останется вырожденным.

Поэтому у нас существует предел $q = \lim P(A_n) \leqslant 1$

$\underbrace{G_{n+1}(0)}_{\to q} = \underbrace{F(G_n(0))}_{F(q)}$, а F непрерывная, поэтому $q = F(q)$, поэтому веротяность - корень уравнения. Осталось понять, что это наименьший корень

Пусть r другой корень уравнения $r = F(r)$. Ещё мы знаем, что F монотонна, потому что производная неотрицательная (просто коэффициенты неотрицательны).

$P(A_1) = G_1(0) = F(0) \overset{\text{монотонность}}{\leqslant} F(r) = r$ - верно в стартовый момент времени.

Пусть $P(A_n) \leqslant r$, тогда $P(A_{n+1}) = G_{n+1}(0) = F(G_n(0)) = F(P(A_n)) \leqslant F(r) = r$

Переходим к пределу и получаем, что $q \leqslant r$ $\square$

`Замечание.` F непрерывная, монотонная, выпуклая на $[0, 1]$, а ещё $F(1) = 1$ и $F(0) \geqslant 0$

Если $m = \mathbb{E}\xi = F'(1) > 1$, то есть вероятность вырождения < 1, если же $m \leqslant 1$, то вероятность вырождения $= 1$

Теорема 4.4. Пусть $m = \mathbb{E}\xi = 1$, $b = \mathbb{D}\xi > 0$, q_n - вероятность вырождения к n-ному шагу, $\gamma_n = q_n - q_{n-1}$ - вероятность вырождения ровно на n-ном шаге.

Тогда

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

Доказательство. Пусть $p_n = 1 - q_n$ и $H(x) = 1 - F(1 - x)$

Тогда $H(p_n) = 1 - F(q_n) = 1 - q_{n+1} = p_{n+1}$, $H(0) = 1 - F(1) = 0$, $H'(0) = F'(1) = 1$, $H''(x) = F''(1 - x)$ и тогда $H''(0) = -F''(1) = -b$

В итоге $H(x) = x - \frac{bx^2}{2} + o(x^2)$

Пусть $a_n = \frac{1}{p_n}$, $a_n - a_{n-1} = \frac{1}{p_n} - \frac{1}{p_{n-1}} = \frac{p_{n-1} - p_n}{p_n p_{n-1}} = \frac{p_{n-1} - H(p_{n-1})}{p_{n-1} H(p_{n-1})} = \frac{\frac{bp_{n-1}^2}{2} + o(p_{n-1}^2)}{p_{n-1}(p_{n-1} + o(p_{n-1}))} = \frac{b}{2} + o(1) \implies a_n \sim \frac{bn}{2}$

Тогда $p_n \sim \frac{2}{bn}$

$\gamma_n = q_n - q_{n-1} = p_{n-1} - p_n = p_{n-1} - H(p_{n-1}) = \frac{bp_{n-1}^2}{2} + o(p_{n-1}^2) \sim \frac{bp_{n-1}^2}{2} \sim \frac{b}{2}\left(\frac{2}{bn}\right)^2 = \frac{2}{bn^2}$ $\square$

4.3. Цепи Маркова

***Определение* 4.5.** Y, не более чем счётное множество - фазовое пространство

$(\Omega, \mathcal{F}, P)$ - вероятностное пространство, $\xi_n : \Omega \to Y$ - случайная величина, такая, что $P(\xi_n = a_n | \xi_{n-1} = a_{n-1}, \ldots, \xi_0 = a_0) = P(\xi_n = a_n | \xi_{n-1} = a_{n-1}) \, \forall a_0, a_1, \ldots, a_n \in Y$

Такая последовательность ξ_n - цепь Маркова

`Замечание.` То есть ξ_n зависит только от ξ_{n-1}

Пример. 1. Случайное блуждание по $\mathbb{Z}$

$P(\xi_n = \xi_{n-1} + 1) = p$

$P(\xi_n = \xi_{n-1} - 1) = 1 - p$

2. Прибор, который бывает в двух состояниях - работает и не работает.

`Замечание.` $\pi_0 = P_{\xi_0}$ - начальное распределение

$p_n(a, b) = P(\xi_n = b | \xi_{n-1} = a)$ - вероятностости переходов. Этот набор данных однозначно определяет все распределения

***Определение* 4.6.** Цепь Маркова называется однородной, если $p_n(a, b)$ не зависят от n.

То есть вероятности переходов не зависят от времени

Обозначение. $p_n(a, b) = p_{ab}$

`Замечание.` Интерпретация: есть частица, которая бегает по фазовому пространству. И мы в каждый момент фиксируем место, где находится частица.

***Определение* 4.7.** Траектория: $\xi_0 = a_0, \xi_1 = a_1, \dots, \xi_n = a_n$

Теорема 4.5. $P(\xi_0 = a_0, \dots, \xi_n = a_n) = \pi_0(a_0) p_{a_0 a_1} p_{a_1 a_2} \dots p_{a_{n-1} a_n}$

Доказательство. Индукция по n

1. База индукция - определение π_0

2. Переход: $n - 1 \to n$

   $P(\xi_0 = a_0, \dots, \xi_n = a_n) = P(\xi_n = a_n | \xi_{n-1} = a_{n-1}, \dots, \xi_0 = a_0) \cdot P(\xi_{n-1} = a_{n-1}, \dots, \xi_0 = a_0) = \underbrace{P(\xi_n = a_n | \xi_{n-1} = a_{n-1})}_{= p_{a_{n-1} a_n}} \cdot \pi_0(a_0) p_{a_0 a_1} \dots p_{a_{n-2} a_{n-1}}$

□

Теорема 4.6. Пусть $\pi_0 : Y \to [0, 1]$, т.ч. $\sum\limits_{y \in Y} \pi_0(y) = 1$, $p : Y \times Y \to [0, 1]$, т.ч. $\sum\limits_{y \in Y} p_{ay} = 1 \, \forall \, a \in Y$

Тогда существует такое пространство $(\Omega, \mathcal{F}, P)$ и последовательность $\xi_n : \Omega \to Y$, такая, что ξ_n цепь Маркова с начальным распределением π_0 и вероятностью перехода p_{ab}

Обозначение. $\pi_n = P_{\xi_n}$ и P - матрица $(p_{ab})_{a,b \in Y}$

Теорема 4.7. $\pi_n = \pi_0 P^n$

Доказательство. Индукция по n. Переход $n - 1 \to n$

$\pi_n(b) = P(\xi_n = b) = \sum\limits_{y \in Y} P(\xi_n = b | \xi_{n-1} = y) \cdot P(\xi_{n-1} = y) = \sum\limits_{y \in Y} p_{yb} \pi_{n-1}(y)$

То есть $\pi_n = \pi_{n-1} P$ □

Обозначение. $p_{ab}(n) = P(\xi_{n+k} = b | \xi_k = a)$ - вероятность перехода за n шагов

***Определение* 4.8.** $\pi : Y \to [0, 1]$ - распределение на Y, если $\sum\limits_{y \in Y} \pi(y) = 1$

***Определение* 4.9.** π - стационарное распределение для цепи Маркова, если $\pi = \pi P$

Пример. Симметричное случайное блуждание на $\mathbb{Z}$, то есть $p = \frac{1}{2}$

Пусть π - стационарное распределение для этого блуждания

Тогда $\frac{1}{2}\pi(n-1) + \frac{1}{2}\pi(n+1) = \pi(n) \iff \pi(n) - \pi(n-1) = \pi(n+1) - \pi(n)$, то есть разность $\alpha = \pi(n) - \pi(n-1)$ не зависит от n

1. $\alpha > 0$, то $\pi(n) = n\alpha + \pi(0) \to +\infty$, так не бывает
2. $\alpha < 0$, то $\pi(n) = n\alpha + \pi(0) \to -\infty$, так тоже не бывает
3. $\alpha = 0$ и $\pi = const$, но так тоже не бывает

Теорема 4.8. Эргодическая теорема Маркова

ξ_n - конечная цепь Маркова и $p_{ab} > 0 \, \forall \, a, b \in Y$

Тогда существует единственное стационарное распределение и $\pi(b) = \lim_{n\to\infty} p_{ab}(n)$

Более того, $|\pi(b) - p_{ab}(n)| \leqslant cq^n$, где $q \in (0,1)$

Доказательство. Доказательство с использованием теоремы Банаха о сжатии из матанализа

d - количество элементов в Y. Рассмотрим $\mathbb{R}^d$ с нормой $||x|| = |x_1| + \ldots + |x_d|$ - полное пространство

$S = \{x \in \mathbb{R}^d \, : \, ||x|| = 1, x_1, x_2, \ldots, x_d \geqslant 0\}$ - замкнутое подмножество $\mathbb{R}^d$ - полное

$T \, : \, S \to S$ и $T(x) = x^T P$, $\delta = \min\limits_{a,b\in Y} p_{ab} > 0$

Проверяем, что T - сжатие с $\lambda = 1 - d\delta$

$$||T_x - T_y|| \overset{z=x-y}{=} ||T_z|| = \sum_{j=1}^{d} |(T_z)_j| = \sum_{j=1}^{d} \left| \sum_{k=1}^{d} z_k p_{kj} \right| = \sum_{j=1}^{d} \left| \sum_{k=1}^{d} z_k (p_{kj} - \delta) + \delta \underbrace{\sum_{k=1}^{d} z_k}_{=0} \right| \leqslant \sum_{j=1}^{d} \sum_{k=1}^{d} |z_k|(p_{kj} -$$
$$\delta) = \sum_{k=1}^{d} |z_k| \sum_{j=1}^{d} \underbrace{(p_{kj} - \delta)}_{=1-\delta d = \lambda} = \lambda \sum_{k=1}^{d} |z_k| = \lambda ||x - y|| \qquad \square$$

`Замечание.` Пусть ξ_n - конечная цепь Маркова, $m \in \mathbb{N}$, т.ч. $p_{ab}(m) > 0 \, \forall a, b \in Y$

Тогда существует единственное стационарное распределение

***Определение* 4.10.** Состояние b достижимо из a, если $\exists \, n \in \mathbb{N}$, т.ч. $p_{ab}(n) > 0$

***Определение* 4.11.** Состояния a и b сообщающиеся, если a достижимо из b, а b достижимо из a

***Определение* 4.12.** Состояние a существенное, если $\forall \, b$, достижимого из a - состояния a и b сообщающиеся

Обозначение. $f_a(n) = P(\xi_n = a | x_{n-1} \neq a, \xi_{n-2} \neq a, \ldots, \xi_1 \neq a, \xi_0 = a)$ - вероятность, стартовав из a, впервые вернуться назад на n-ном шаге.

$F_a = \sum\limits_{n=1}^{\infty} f_a(n)$ - вероятность возврата назад в a

***Определение* 4.13.** a - возвратное состояние, если $F_a = 1$

***Определение* 4.14.** a - нулевое состояние, если $p_{aa}(n) \to 0$

Теорема 4.9. Критерий возвратности

a - возвратное $\iff \mathcal{P}(a) = \sum\limits_{n=1}^{\infty} p_{aa}(n)$ расходится

И если a не возвратное, то $F_a = \frac{\mathcal{P}(a)}{1+\mathcal{P}(a)}$

Следствие. Если a не возвратное $\implies$ a - нулевое

Теорема 4.10. Теорема солидарности

a и b сообщающиеся состояния

Тогда они возвратны/не возвратны (нулевые/не нулевые) одновременно

Доказательство. a и b сообщающиеся, значит $\exists\, j, k \in \mathbb{N}\, :\, p_{ab}(j) > 0$ и $p_{ba}(k) > 0$

$p_{aa}(n+j+k) \geqslant p_{ab}(j)p_{bb}(n)p_{ba}(k)$ и $\sum\limits_{n=1}^{\infty} p_{aa}(n+j+k) \geqslant p_{ab}(j)p_{ba}(k) \sum\limits_{n=1}^{\infty} p_{bb}(n)$

Отсюда всё следует, потому что:

Если $p_{aa}(n+j+k) \to 0$, то $p_{bb}(n) \to 0$

$\square$