Теория вероятностей

Швецова Анна, Василенко Елизавета

16 мая 2018 г.

Содержание

1. Введение **1**

1.1 Графовые теоремы. 1

1.2 Предельные теоремы для схем Бернулли 2

2. Общая теория вероятностей **5**

2.1 Колмогоровское определение вероятности 5

2.2 Случайная величина, распределение, плотность 7

2.3 Примеры вероятностных распределений 9

2.4 Многомерные распределения 11

2.5 Свёртки мер 12

3. Математическое ожидание **15**

3.1 Корреляция и ковариация 21

3.2 Сходимость случайных величин и закон больших чисел 24

3.3 Закон больших чисел 26

4. Характеристическая функция **29**

5. Центральная предельная теорема **40**

6. Случайные процессы **45**

6.1 Определения 45

6.2 Условные мат. ожидания 46
6.3 Марковские цепи 49

1. Введение

27.02.18

1.1. Графовые теоремы.

Теорема 1.1 (Теорема Рамсея). $\forall k, m \exists n = n(k, m)$ т.ч. для любого графа на n вершинах существует либо пустой подграф на m вершинах, либо полный на k.

***Определение* 1.1.** $\mathcal{R}(k, m)$ – наименьшее такое n.

Нижняя оценка на $\mathcal{R}(k, m)$:

Теорема 1.2 (Теорема Эрдеша). Если $2^{1-C_k^2}C_n^k < 1$, то существует такой граф на n вершинах, что у него нет ни полного подграфа на k вершинах, ни пустого подрафа на m.

В частности $\mathcal{R}(k, k) > 2^{k/2}$

Доказательство. Ω будет состоять из графов на n вершинах, $\#\Omega = 2^{C_n^2}$.
P(событие: между a и b проведено ребро) $= 1/2$.

Такие события независимы. Значит, можно считать, что случайный граф – это граф, в котором каждое ребро с вероятностью $1/2$.

$P(\bigcup_{j=1}^{l} A_j) \leqslant \sum_{j=1}^{l} P(A_j)$.

Зафиксируем набор из k вершин. В нём C_k^2 пар вершин. С вероятностью $2^{-C_k^2}$ все пары вершин соединены ребром и подграф полный. Аналогично, с вероятностью $2^{-C_k^2}$ подграф пустой.

P(выбранный набор пустой или полный подграф) $= 2 \cdot 2^{-C_k^2} = 2^{1-C_k^2}$

P(хоть какой-то набор полный или пустой) $\leqslant 2^{1-C_k^2}C_n^k < 1$

$\Rightarrow$ существует граф на n вершинах, у которого никакие k вершин не походят.

В частности, при $n \leqslant 2^{k/2}$

$$2^{1-C_k^2}C_n^k = \frac{n(n-1)\dots(n-k+1)}{k!}\frac{1}{2^{C_k^2-1}} < \frac{n^k}{k!}\frac{2}{2^{k/2(k-1)}} = \frac{n^k}{2^{k^2/2}}\frac{2^{k/2+1}}{k!} < \frac{n^k}{2^{k^2/2}} = \left(\frac{n}{2^{k/2}}\right)^k \leqslant 1$$

□

***Определение* 1.2.** $\mathcal{R}(k, m)$ – наименьшее такое n.

1.2. Предельные теоремы для схем Бернулли

***Определение* 1.3** (Схема Бернулли)**.** Монетку кидают n раз.

С вероятностью $p \in (0, 1)$ выпадает орёл (1), с вероятностью $q = 1 - p$ – решка (0)

S_n – число выпавших орлов. $P(S_n = k) = C_n^k p^k q^{n-k}$.

Пример.

- $p = 1/5$ $P(S)_{1000} = 220 \approx 0,008984\ldots$
- $p = 1/6$ $P(S)_{2000} = 360 \approx 0,006625\ldots$

Хочется приближённые оценки

Теорема 1.3 (Пуассона)**.** Последовательность схем Бернулли.
$np_n \underset{n\to\infty}{\to} \lambda > 0$
Тогда $P(S_n = k) \underset{n\to\infty}{\to} \frac{\lambda^k e^{-\lambda}}{k!}$.

Замечание. При этом сходимость равномерная по k. Утверждение верно, если k не фиксировано, а $o(\sqrt{n})$. (Доказывается аналогично, но чуть труднее.)

Доказательство.

$$P(S_n = k) = C_n^k p_n^k (1 - p_n)^{n-k} = \frac{n(n-1)\ldots(n-k+1)}{k!} p_n^k (1-p_n)^{n-k}$$

$$\sim \frac{1}{k!} n^k p_n^k (1-p_n)^n \frac{1}{(1-p_n)^k} \sim \frac{\lambda^k}{k!}(1-p_n)^n$$

Надо доказать, что $(1 - p_n)^n \to e^{-\lambda}$, т.е. $n \log(1 - p_n) \to -\lambda$ □

Теорема 1.4 (Прохорова)**.** $\sum_{k=0}^{\infty} \left| P(S_n = k)\ \frac{\lambda^k e^{-\lambda}}{k!} \right| \leqslant \frac{\min\{2,\lambda\}}{n} 2\lambda$

Доказательство.

$$P(S_n = k) = C_n^k p_n^k (1 - p_n)^{n-k} = \frac{n(n-1)\ldots(n-k+1)}{k!} p_n^k (1-p_n)^{n-k}$$

$$\sim \frac{1}{k!} n^k p_n^k (1-p_n)^n \frac{1}{(1-p_n)^k} \sim \frac{\lambda^k}{k!}(1-p_n)^n$$

Надо доказать, что $(1 - p_n)^n \to e^{-\lambda}$, т.е. $n \log(1 - p_n) \to -\lambda$

□

Пример. Рулетка $p = 1/37$ – вероятность выигрыша, $n = 111$ – число игр, $\lambda = np = 3$. При выигрыше $+37$, за игру -1.

Вероятность остаться с суммой 0:

$P(S_{111} = 3) = C_{111}^3 \left((\frac{1}{37}) \right)^3 (1 - 1/37)^{111-3} \approx 0,227127\dots$

По теореме: $\approx \frac{\lambda^3 e^{-\lambda}}{3!} = \frac{3^3 e^{-3}}{6} = 0,224041\dots$

Вероятность не проиграть:

$P(S_n > 3) = 1 - P(S_n = 0) - P(S_n = 1) - P(S_n = 2) - P(S_n = 3) \approx 1 - \frac{3^0 e^{-3}}{0!} - \frac{3^1 e^{-3}}{1!} - \frac{3^2 e^{-3}}{2!} - \frac{3^3 e^{-3}}{3!} = 0,352768\dots$

Точное значение:

$1 - \frac{13}{e^3} = 0,352754\dots$

Теорема 1.5 (Локальная теорема Муавра-Лапласа)**.** $0 < p < 1$, $x \coloneqq \frac{k-np}{\sqrt{npq}}$, T – фиксированное число. Если $n \to \infty$, а k меняется таким образом, что при $|x| \leqslant T$, $P(S_n = k) \sim \frac{1}{\sqrt{2\pi npq}} e^{-x^2/2}$ равномерно по x.

Замечание. Более того, если $\phi(n) = o((npq)^{1/6})$ и $|x| \leqslant \phi(n)$, то заключение верно и эквивалентность равномерна по x (то есть, предел, равный 1, равномерный).

Доказательство. $k = np + x\sqrt{npq} \geqslant np - T\sqrt{npq} \to +\infty$

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

$\alpha \coloneqq \frac{k}{n} = \frac{np + x\sqrt{npq}}{n} = p + x\sqrt{\frac{pq}{n}} \to p$

$1 - \alpha = 1 - p - x\sqrt{\frac{pq}{n}} = q - x\sqrt{\frac{pq}{n}} \to q$

$P(S_n = k) = C_n^k p^k q^{n-k} \sim \frac{n^n e^{-n} \sqrt{2\pi n} p^k q^{n-k}}{} k^k e^{-k} \sqrt{2\pi k} (n-k)^{n-k} e^{-(n-k)} \sqrt{2\pi(n-k)} =$

$= \frac{p^k (1-p)^{n-k}}{\alpha^k (1-\alpha)^{n-k}} \frac{1}{\sqrt{2\pi n\alpha(1-\alpha)}} \sim \frac{p^k (1-p)^{n-k}}{\alpha^k (1-\alpha)^{n-k}} \frac{1}{\sqrt{2\pi npq}}$

Надо понять, что $\left(\frac{p}{\alpha}\right)^k \left(\frac{1-p}{1-\alpha}\right)^{n-k} \to e^{-x^2/2}$

$A \coloneqq k \log(\frac{\alpha}{p}) + (n-k)\log(\frac{1-\alpha}{1-p}) \to^? \frac{x^2}{2}$

$\ln(1+t) = t - \frac{t^2}{2} + \mathcal{O}(t^3)$

$\frac{\alpha}{p} = 1 + x\sqrt{\frac{q}{pn}}$

$\frac{1-\alpha}{1-p} = 1 - x\sqrt{\frac{p}{qn}}$

$\ln(\frac{\alpha}{p}) = \ln(1 + x\sqrt{\frac{q}{np}}) = x\sqrt{\frac{q}{np}} - 1/2 x^2 \frac{q}{np} + \mathcal{O}(n^{-3/2})$

$\ln(\frac{1-\alpha}{1-p}) = \ln(1 - x\sqrt{\frac{p}{nq}}) = -x\sqrt{\frac{p}{nq}} - 1/2 x^2 \frac{p}{nq} + \mathcal{O}(n^{-3/2})$

$k = n\alpha = np + x\sqrt{npq}$

$n - k = n(1-\alpha) = nq - x\sqrt{npq}$

$A = x\sqrt{qpn} + x^2 q - 1/2 x^2 q - 1/2 x^3 \frac{q^{2/3}}{\sqrt{np}} + \mathcal{O}(n^{-1/2}) -$

$-x\sqrt{qpn} + x^2 p - 1/2 x^2 p + 1/2 x^3 \frac{p^{2/3}}{\sqrt{nq}} + \mathcal{O}(n^{-1/2}) = \frac{x^2}{2} + \mathcal{O}(n^{-1/2})$

□

Пример Рулетка. Ставим только на красные. $n = 222, k = 111$(сколько раз выиграли)$, p = 18/37, q = 19/37$

$P(S_{222} = 111) \approx 0,0493228\dots$

$\frac{1}{\sqrt{2\pi npq e^?}} \approx 0,0493228\dots$

Теорема 1.6 (Интегральная теорема)**.** $0 < p < 1$

$P(a < \frac{S_n - np}{\sqrt{npq}} \leqslant b) \underset{n\to\infty}{\rightarrow} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt$ равномерно по a и b.

Замечание. Частный случай теоремы Берри-Эссена

$\sup \left| P(\frac{S_n - np}{\sqrt{npq}} \leqslant x) - \frac{1}{\sqrt{1\pi} \int_{-\infty}^x} e^{-t^2/2} dt \right| \leqslant \frac{p^2 + q^2}{\sqrt{npq}}$

Замечание. Лучше $\sqrt{n}$ в знаменателе быть не может.

$P(S_{2n} = n) = C_{2n}^n (1/2)^n (1/2)^n = \frac{C_{2n}^n}{4^n} \sim \frac{1}{\sqrt{\pi n}}$

$p = q = 1/2$

$P(S_{2n} < n) + P(S_{2n} = n) + p(S_{2n} > n) = 1$

$P(S_{2n} < n) \approx \frac{1 + \sqrt{\pi n}}{2} = 1/2 + \frac{1}{2\sqrt{\pi n}}$

$\left| P(S_{2n} < n) - \frac{1}{\sqrt{2\pi n}} \int_{-\infty}^0 e^{-t^2/2} dt \right| \approx \frac{1}{2\sqrt{\pi n}}$

Пример. Задача о театре

1600 мест. Два гардероба у двух разных входов. Люди случайно выбирают один из входов. Сколько сделать мест в гардеробе, чтобы он переполнялся не чаще раза в месяц?

В гардеробах по C мест.

$P((S_n < C) \cap (S_n \geqslant n - C)) \leqslant 29/30$

$= P(n - C \leqslant S_n \leqslant C) = P(\frac{n - C - np}{\sqrt{npq}} \leqslant \frac{S_n - np}{\sqrt{npq}} \leqslant \frac{C - np}{\sqrt{npq}}) \approx \frac{1}{2\pi} \int_{\frac{nq - C}{\sqrt{npq}}}^{\frac{C - np}{\sqrt{npq}}} e^{-t^2/2} dt \approx 2 \cdot 29/60$

$\frac{C - 800}{20} \approx 2,13$, $C \approx 800 + 20 \cdot 2 \cdot 13 = 843$

2. Общая теория вероятностей

07.03.2018

2.1. Колмогоровское определение вероятности

***Определение* 2.1** (Вероятностное пространство)**.** (Ω, F, P), где

Ω – пространство элементарных событий (простыми словами множество).

F – совокупность случайных событий $\subset 2^{\Omega}$. σ-алгебра

P – мера на F, такая что $P(\Omega) = 1$ (вероятностная мера)

Замечание. На самом деле от меры требуется только конечность т.к. её всегда можно отнормировать.

Замечание. Если Ω не более чем счётна, то в качестве F всегда можно взять не более чем счётные подмножества. $F = 2^{\Omega}$. В этом случае можно воспользоваться счётной аддитивностью меры и тогда мера определится везде.

***Определение* 2.2** (Условная вероятность)**.** Пусть $B \in F$, $P(B) > 0$, тогда $P(A|B) := \frac{P(A \cap B)}{P(B)}$

Формула полной вероятности, формула и теорема Байеса верны и тут т.к. основывались в дискретном пространстве только на этой формуле.

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

$A_1, A_2 \ldots, A_n$, $P(A_1)P(A_2)\ldots P(A_n) = P(A_1 \cap A_2 \ldots \cap A_n)$

Для бесконечных последовательностей событий скажем, что они независимы, если любой конечный поднабор независим.

Лемма (Бореля-Кантелли)**.** $A_1, A_2, \ldots$ – события, B – событие "наступило бесконечное число событий из $A_1, A_2 \ldots$".

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

Доказательство. Начнём с понимания того, что такое B.

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

Пояснение: что значит произошло бесконечное число A_i? Значит, мы взяли какой-то элемент, который лежит в бесконечном числе A_i. Значит, в любом $\bigcup_{k=n}^{\infty} A_k$ он тоже лежит. Но тогда такой элемент лежит и в пересечении.

Если некоторый элемент лежал лишь в конечном числе A_i, то, начиная с некоторого i в объединение он не попадёт. Тогда и в пересечение тоже.

$P(B) = P(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k)$

$\bigcup_{k=n}^{\infty} A_k$ вложены друг в друга как множества, поэтому

$P(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k) = \lim\limits_{n\to\infty} P(\bigcup_{k=n}^{\infty} A_k) \leqslant \lim\limits_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$ (в первых двух переходах вспоминаем про то, что P – это мера, в последнем получаем ноль т.к. хвост сходящегося ряда).

Первый пункт доказан. Второй.

Если $A_1, A_2 \ldots$ независимы, тогда и $\overline{A_1}, \overline{A_2}, \ldots$ независимы. Тогда

$$P\left(\bigcap_{k=1}^{\infty} \overline{A_k}\right) \leftarrow P\left(\bigcap_{k=1}^{m} \overline{A_k}\right) = \Pi_{k=1}^{m} P(\overline{A_k}) \to \Pi_{k=1}^{\infty} P(\overline{A_k})$$

Заметим, что первая стрелочка требует конечности меры пространства, которая у нас конечно есть.

$$P\left(\bigcap_{k=1}^{\infty} \overline{A_k}\right) = \Pi_{k=1}^{\infty} P(\overline{A_k}) = \Pi_{k=1}^{\infty}(1 - P(A_k))$$

$$P\left(\bigcap_{k=1}^{\infty} \overline{A_k}\right) = 1 - P\left(\overline{\bigcap_{k=1}^{\infty} \overline{A_k}}\right) = 1 - P\left(\bigcup_{k=1}^{\infty} A_k\right)$$

А теперь соберём равенство и будем считать всё не с 1, а с n и получим, что:

$$\Pi_{k=n}^{\infty}(1 - P(A_k)) = 1 - P\left(\bigcup_{k=n}^{\infty} A_k\right) \underset{n\to\infty}{\to} 1 - P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right) = 1 - P(B)$$

Осталось понять, что $\Pi_{k=n}^{\infty} P(1 - P(A_k)) \underset{n\to\infty}{\to} 0$

Прологарифмируем, и вспомним, что $\ln(1-t) \leqslant t$.

$$\ln(\Pi_{k=n}^{\infty}(1 - P(A_k))) = \sum_{k=n}^{\infty} \ln(1 - P(A_k)) \leqslant -\sum_{k=n}^{\infty} P(A_k) = -\infty$$

$\square$

Следствие Закон 0 и 1. Если $A_1, A_2 \ldots$ независимые в совокупности события, то или $P(B) = 0$ или $P(B) = 1$.

2.2. Случайная величина, распределение, плотность

***Определение* 2.4** (Случайная величина)**.** Обозначают как $\xi, \eta, \zeta \ldots$

Это просто измеримая функция : $\Omega \to \mathbb{R}$

***Определение* 2.5** (Распределение случайной величины ξ)**.** $A \subset \mathbb{R}$ – борелевское.

$P_\xi(A) = P(\xi \in A) = P(\omega \in \Omega : \xi(\omega) \in A)$

$P(\xi)$ – вероятностная мера на борелевских подмножествах $\mathbb{R}$. (упражнение – понять, что это одно и то же)

Замечание. Для определения меры на борелевской оболочке достаточно определить её на ячейках. На ячейках, в свою очередь, достаточно определить меру на лучах

***Определение* 2.6** (Функция распределения с.в. ξ)**.** $F_\xi(x) := P(\xi \leqslant x)$

Замечание. $P_\xi((a, b]) = P(a < \xi \leqslant b) = F_\xi(b) - F_\xi(a) \Rightarrow$ Функция распределения однозначно задаёт распределение с.в.

***Определение* 2.7.** Случайные величины ξ и η одинаковораспределены, если $P_\xi = P_\eta$ (распределения одинаковы).

Замечание. Т.е. одинаковы их функции распределения.

Свойства.

1. $0 \leqslant F_\xi(x) \leqslant 1$

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

2. F_ξ не убывает

   Очевидно из строчки выше

3. $\lim\limits_{x \to -\infty} F_\xi(x) = 0$ и $\lim\limits_{x \to +\infty} F_\xi(x) = 1$

   $\lim\limits_{x \to -\infty} F_\xi(x) = P(\bigcap_{n=1}^{\infty}(\xi \leqslant x_n)) = P(\varnothing) = 0$

   $\lim\limits_{x \to +\infty} F_\xi(x) = P(\bigcup_{n=1}^{\infty}(\xi \leqslant x_n)) = P(\Omega) = 1$

4. F_ξ непрерывна справа

   $\lim\limits_{x_n \searrow x} P(\xi \leqslant x_n) = P(\bigcap_{n=1}^{\infty}(\xi \leqslant x_n)) = P(\xi \leqslant x)$

5. $F_{\xi+c}(x) = F_\xi(x - c)$

   $F_{\xi+c}(x) = P(\xi + c \leqslant x) = P(\xi \leqslant x - c) = F_\xi(x - c)$

6. $F_{x\xi}(x) = F_\xi(\frac{x}{c})$, $c > 0$

   $F_{c\xi}(x) = P(c\xi \leqslant x) = P(\xi \leqslant \frac{x}{c}) = F_\xi(\frac{x}{c})$

Замечание. Любая функция $F : \mathbb{R} \to \mathbb{R}$, обладающая свойствами с 2 по 4 – это функция распределения некоторой случайной величины.

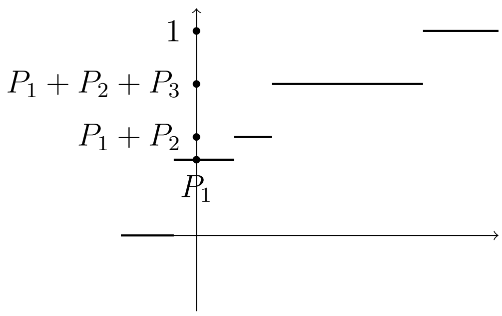
***Определение* 2.8** (Дискретное распределение)**.** $\xi : \Omega \to \{y_1, y_2, \ldots\}$ – не более чем счётно. – дискретная случайная величина.

Замечание. Как устроена вероятность дискретного распредлеения?

$P_\xi(\{x\}) = 0$, если $x \neq y_k$.

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

А функция распределения? Пусть $y_1 < y_2 < y_3 < \ldots$, тогда функция устроена ступенечками, высоты которых – это вероятности $P(\xi = y_i)$. Например:



Если не так, то там будут какие-то хаотические ступеньки.

Вывод: распределение полностью определяется величинами $P(\xi = y_k)$. $\sum_{k=1}^{\infty} P(\xi = y_k) = 1$.

***Определение* 2.9** (Непрерывное распределение)**.** $\forall x \in \mathbb{R}$ $P_\xi(\{x\}) = 0$ т.е. $\forall x \in \mathbb{R}$ $P(\xi = x) = 0$.

Замечание. $P(\xi = x) = P(\xi \leqslant x) - P(\xi < x) = F_\xi(x) - \lim\limits_{y \to x-} F_\xi(y)$

$\lim\limits_{y \to x-} F_\xi(y) = F_\xi(x)$

т.е. непрерывное распределение = распределение, у которых функция распределения непрерывна

Пример Мерзкий пример – Канторова лестница. Иногда, несмотря на то, что распределение непрерывно, могут получаться весьма непривлекательные картинки. gif с построением Это некоторая функция распределения.

***Определение* 2.10** (Абсолютно непрерывное распределение)**.** Если существует $p_\xi(t)$ – измеримая функция т.ч.

$F_\xi(x) = \int_{-\infty}^{x} p_\xi(t), p_\xi : \mathbb{R} \to \mathbb{R}$

$p_\xi(t)$ – плотность распределения.

Свойства.

1. $P_\xi(A) = P(\xi \in A) = \int_A p_\xi(t)dt$

   $P_\xi(A) = \int_A p_\xi(t)dt$ верна на лучах $(-\infty, x]$. Значит верна на полуинтервалах (как на разнице лучей).

   Тогда по единственности продолжения меры на полуинтервалах, она верна на борелевских множествах.

2. $p_\xi \geqslant 0$ почти везде.

   $A := \{t : p_\xi(t) < 0\}$. Тогда из предыдущего свойства

   $P_\xi(A) = \int_{\{t: p_\xi(t)<0\}} p_\xi(t)dt < 0$, если

   $P(p_\xi < 0) > 0$, но в то же время $P_\xi(A) \geqslant 0$ как вероятность. Противоречие.

3. $\int_\mathbb{R} p_\xi(t)dt = 1$.

   $\int_\mathbb{R} p_\xi(t)dt = \lim\limits_{x\to+\infty} F_\xi(x) - \lim\limits_{x\to-\infty} F_\xi(x) = 1 - 0 = 1$

   Или ещё можно доказать так: $\int_\mathbb{R} p_\xi(t)dt = P_\xi(\mathbb{R}) = 1$

4. $p_\xi(t) = F'_\xi(t)$ при почти всех t.

   Если доказывали в теории меры, то см. туда, если нет, то я её доказывать не буду.

14.03.2018

2.3. Примеры вероятностных распределений

Пример.

1. Биномиальные распределения

   $\xi \sim Binom(n, p)$

   ($\sim$ – ‘случайная величина с таким распределением‘)

   $0 < p < 1,\ n \in \mathbb{N}$

   $P(\xi = k) = C_n^k p^k (1-p)^{n-k}$ – число успехов в схеме Бернулли

2. Распределение Пуассона

   $\xi \sim Poisson(\lambda)$

   $\lambda > 0$

   $P(\xi = k) = \frac{\lambda^k}{k!}e^{-\lambda}$

3. Геометрическое

   $\xi \sim Geom(p)$

   $0 < p < 1$

   $P(\xi = k) = p(1-p)^{k-1}$, $k = 1, 2, \dots, \infty$

4. Дискретное равномерное распределение

   $\xi : \Omega \to \{y_1, y_2, \dots, y_n\}$

   $P(\xi = y_k) = 1/n$

5. Непрерывное равномерное распределение

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

   $P_\xi(t) = \frac{1}{b-a}\mathbb{1}_{[a,b]}(t)$

6. Нормальное распределение

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

   $a \in \mathbb{R},\ \sigma > 0$

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

6'. Стандартное нормальное распределение

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

   $p_\xi(t) = \frac{1}{\sqrt{2\pi}}e^{-t^2/2}$

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

***Определение* 2.11.**

$\Phi(x) \coloneqq \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt$

$\Phi(0) = 1/2\, \Phi(\infty) = 1/2$ т.к. интеграл симметричен относительно 0

$\Phi_0(x) \coloneqq \frac{1}{\sqrt{2\pi}}\int_{0}^{x} e^{-t^2/2}dt$

при $x > 0$ $\Phi(x) = \Phi_0(x) + 1/2$

при $x < 0$ $\Phi(x) = 1/2 - \Phi_0(-x)$

Пример (продолжение).

7. Экспоненциальное распределение

   $\xi \sim Exp(\lambda)$

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

2.4. Многомерные распределения

***Определение* 2.12** (Совместное (многомерное) распределение)**.**

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

$A \in \mathbb{R}^n$, $P_{\overrightarrow{\xi}}(A) = P(\overrightarrow{\xi} \in A)$

Будем считать, что вероятностная мера P задана на $\mathcal{B}^n$ – борелевских подмножествах $\mathbb{R}^n$

Замечание.

$P_{\overrightarrow{\xi}}$ определяет меры $P_{\xi_1}, P_{\xi_2}, \ldots, P_{\xi_n}$

$B \subset \mathbb{R}$, $P_{\xi_1}(B) = P_{\overrightarrow{\xi}}(B \times \mathbb{R}^{n-1})$

Обратное неверно

Пример.

$\xi_1, \xi_2 : \Omega \to \{0, 1\}$ с равными вероятностями.

Если подбрасывания независимы, то $\overrightarrow{\xi} = (\xi_1, \xi_2)$, то четыре равновероятных исхода.

Если $\xi_1 = \xi_2$, то два равновероятных исхода $(0, 0)$ и $(1, 1)$.

***Определение* 2.13.**

Случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ независимы, если

$\forall A_1, A_2, \ldots, A_n \in \mathbb{R}$ борелевских события $\{\xi_1 \in A_1\}, \{\xi_2 \in A_2\}, \ldots, \{\xi_n \in A_n\}$ – независимы

Теорема 2.1.

$\xi_1, \xi_2, \ldots, \xi_n$ – независимы $\Leftrightarrow P_{\overrightarrow{\xi}} = \bigotimes_{k=1}^n P_{\xi_k}$ (произведение мер)

Доказательство.

$P_{\overrightarrow{\xi}}(A_1 \otimes A_2 \otimes \cdots \otimes A_n) = P(\overrightarrow{\xi} \in A_1 \otimes A_2 \otimes \cdots \otimes A_n) =$

$= P(\xi_1 \in A_1, \xi_2 \in A_2, \ldots, \xi_n \in A_n) \underset{\text{так как независ.}}{=} \prod_{k=1}^n P(\xi_k \in A_k) = \prod_{k=1}^n P_{\xi_k}(A_k)$ □

***Определение* 2.14** (Совместная (многомерная) функция распределения).
$F_{\overrightarrow{\xi}}(x_1, x_2, \ldots, x_n) = P(\xi_1 \leqslant x_1, \xi_2 \leqslant x_2, \ldots, \xi_n \leqslant x_n)$

***Определение* 2.15** (Совместная плотность распределения).
$p_{\overrightarrow{\xi}}(t_1, t_2, \ldots, t_n)$, т.ч.
$F_{\overrightarrow{\xi}}(x_1, x_2, \ldots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \int_{-\infty}^{x_3} \ldots \int_{-\infty}^{x_n} p_\xi(t_1, t_2, t_3, \ldots, t_n) dt_n \ldots dt_3 dt_2 dt_1$

(интеграл по n-мерной мере Лебега)

Следствие.
$\xi_1, \xi_2, \ldots \xi_n$ – независимы $\Leftrightarrow F_{\overrightarrow{\xi}}(x_1, x_2, \ldots, x_n) = F_{\xi_1}(x_1) F_{\xi_2}(x_2) \ldots F_{\xi_n}(x_n)$

Доказательство.

"$\Rightarrow$" $P_{\overrightarrow{\xi}}((-\infty, x_1] \times \cdots \times (-\infty, x_n]) = \prod_{k=1}^n P_{\xi_k}(-\infty, x_k] = \prod_{k=1}^n F_{\xi_k}(x_k)$

"$\Leftarrow$" Функция распределения однозначно определяет меру на ячейках $\Rightarrow$ везде. □

Следствие. $\xi_1, \xi_2, \ldots \xi_n$ – абсолютно непрерывные с.в.
$\xi_1, \xi_2, \ldots \xi_n$ – независимы $\Leftrightarrow p_{\overrightarrow{\xi}}(t_1, t_2, \ldots, t_n) = p_{\xi_1}(t_1) p_{\xi_2}(x_2) \ldots p_{\xi_n}(x_n)$.

Доказательство. По абсолютной непрерывности мы знаем, что у функций распределения есть производная, она же плотность. Значит, продифференцируем их и получим равенство выше. □

2.5. Свёртки мер

***Определение* 2.16** (Свёртка мер). μ и ν – конечные меры на борелевских подмножествах $\mathbb{R}$.
Свёртка мер $(\mu * \nu)(A) \coloneqq \int_{\mathbb{R}} \mu(A - x) d\nu(x)$
$(A - x)$ – это сдвиг множества A на x влево на прямой.

Свойства.

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

   Доказательство.
   $(\mu * \nu)(A) = \int_{\mathbb{R}} \mu(A - x) d\nu(x) = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_{A-x}(y) d\mu(y) d\nu(x) = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_A(y + x) d\mu(y) d\nu(x)$
   Первый переход – определение интеграла по мере, второй – $y \in A - x \iff y + x \in A$ □

2. $(\mu_1 * \mu_2 * \cdots * \mu_n)(A) = \int_{\mathbb{R}^n} \mathbb{1}_A(x_1 + \cdots + x_n) d\mu_1(x_1) \ldots \mu_n(x_n)$

3. $\mu * \nu = \nu * \mu$

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

5. $c\mu * \nu = c(\mu * \nu)$

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

7. δ_x – мера, состоящая из единичной нагрузки в точке x.

   $\mu * \delta_0 = \mu$

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

Теорема 2.2.

Важный частный случай – мера с плотностями

μ имеет плотность p_μ

ν имеет плотность p_ν

$\Rightarrow \mu * \nu$ имеет плотность $p(t) = \int_{\mathbb{R}} p_\mu(t-s) p_n u(s) ds$ – свёртка функций p_μ и p_ν

Доказательство.

$(\mu * \nu)(A) \overset{?}{=} \int_A \int_{\mathbb{R}} p_\mu(t-s) p_\nu(s) ds dt$

$\int_A \int_{\mathbb{R}} p_\mu(t-s) p_\nu(s) ds dt = \int_{\mathbb{R}^2} \mathbb{1}_A(t) p\mu(t-s) p_\nu(s) ds dt =$

$u := t - s,\ du = dt$

$= \int_{\mathbb{R}^2} \mathbb{1}_A(u+s) p_\mu(u) p_\nu(s) du ds = \int_{\mathbb{R}} \mathbb{1}_A(x+y) d\mu(x) d\nu(y) \underset{\text{по св-ву 1}}{=} (\mu * \nu)(A)$

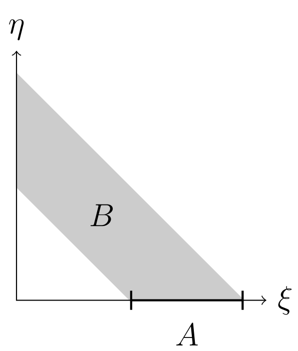
Предпоследний переход: потому что интеграл по мере – это домножить на плотность и проинтегрировать обычным образом. (TODO: Понять это) □

Теорема 2.3 (о распределении суммы независимых с.в.).

ξ и η – независимые с.в. $\Rightarrow P_{\xi+\eta} = P_\xi * P_\eta$

Доказательство. $A \subset \mathbb{R}$

$(\xi, \eta) \in B \subset \mathbb{R}^2 \Leftrightarrow \xi + \eta \in A$



$P_{\xi+\eta}(A) = P(\xi+\eta \in A) = P((\xi,\eta) \in B) = \int_{\mathbb{R}^2} \mathbb{1}_B(x,y) dP_{\xi,\eta}(x,y) =$
$= \int_{\mathbb{R}^2} \mathbb{1}_A(x+y) dP_{\xi,\eta}(x,y) \underset{\text{по т. Фубини}}{=} \int_{\mathbb{R}^2} \mathbb{1}_A(x+y) dP_\xi(x) dP_\eta(y)$

□

Пример.

1. Свёртка с дискретным распределением
   $\nu = \sum p_k \delta_{x_k}$, $p_k > 0$
   $(\mu * \nu)(A) = \int_{\mathbb{R}} \mu(A-x) d\nu(x) = \sum p_k \mu(A - x_k)$

2. $\xi_1 \sim Poisson(\lambda_1)$, $\xi_2 \sim Poisson(\lambda_2)$ – независимы
   $P_{\xi_2} \sum_{k=0}^{\infty} \frac{\lambda_2^k e^{-\lambda_2}}{k!} \delta_k$
   $P_{\xi_1+\xi_2}(\{n\}) = P_{\xi_1} * P_{\xi_2}(\{n\}) = \sum_{k=0}^{\infty} \frac{\lambda_2^k e^{-\lambda_2}}{k!} \frac{\lambda_1^{n-k} e^{-\lambda_1}}{(n-k)!} =$
   $= \sum_{k=0}^{n} \frac{\lambda_2^k e^{-\lambda_2}}{k!} \frac{\lambda_1^{n-k} e^{-\lambda_1}}{(n-k)!} = e^{-\lambda_1+\lambda_2} \frac{1}{n!} \sum_{k=0}^{n} C_n^k \lambda_2^k \lambda_1^{n-k} = \frac{1}{n!} (\lambda_1+\lambda_2)^n e^{-(\lambda_1+\lambda_2)}$

   Замечание. $P(\xi_1+\xi_2 = n) = \sum_{k=0}^{n} P(\xi_1 = k, \xi_2 = n-k) = \sum_{k=0}^{n} P(\xi_1 = k) P(\xi_2 = n-k)$ и подставить

Упражнение. 1. ξ_1 и $\xi_2 \sim Exp(1)$ найти распределение $\xi_1 + \xi_2$

2. $\xi_1 \sim \mathcal{N}(a_1, \sigma_1^2),\ \xi_2 \sim \mathcal{N}(a_2, \sigma_2^2) \Rightarrow \xi_1 + \xi_2 \sim \mathcal{N}(a_1 + a_2, \sigma_1^2 + \sigma_2^2)$

21.03.2018

3. Математическое ожидание

***Определение* 3.1.** (Математическое ожидание) $\xi : \Omega \to \mathbb{R}$ – случайная величина
$E_\xi \coloneqq \int_\Omega \xi(w)dP(w)$ – математическое ожидание (если такой интеграл существует)

Замечание. Если случайная величина неотрицательна, то такой интеграл всегда существует (т.к. с.в. – это измеримая функция). Если же она меняет знак, необходимо, чтобы отрицательная и положительная составляющие не были бесконечностями одновременно.

Замечание. Т.к. вероятностная мера на всём пространстве единичка, то такой интеграл – это в точности среднее значение случайной величины. (В помощь к осознанию: представьте себе для начала дискретную вероятность. Тогда матожидание, как и ожидается, будет выглядеть ровно как сумма произведений значений на их вероятности).

Свойства.

1. $E_\xi < +\infty \iff E_{|\xi|} < +\infty$
   Значем из теории меры TODO: Откуда именно
2. $E_{\alpha\xi+\beta\eta} = \alpha E_\xi + \beta E_\eta$ (Линейность интеграла по мере)
3. $\xi \geqslant 0$ с вероятностью 1 $\Rightarrow E_\xi \geqslant 0$
   Измеримая функция больше нуля почти везде. Тогда и интеграл по ней неотрицательный
4. $\xi \geqslant \eta$ с вероятностью 1 $\Rightarrow E_\xi \geqslant E_\eta$
   Аналогично, одна измеримая функция почти везде не меньше другой
5. $E_\xi = \int_\mathbb{R} x dP_\xi(x)$ (Следует из 6)
6. Если $\xi_1, \xi_2, \ldots, \xi_n$ – с.в
   $f : \mathbb{R}^n \to \mathbb{R}$ измеримо относительно борелевской σ-алгебры
   Тогда $E_{f(\xi_1,\xi_2,\ldots,\xi_n)} = \int_{\mathbb{R}^n} f(x_1, x_2, \ldots, x_n) dP_{\xi_1,\xi_2,\ldots,\xi_n}(x_1, x_2, \ldots, x_n)$
   В частности, $E_{f(\xi)} = \int_\mathbb{R} f(x)dP_\xi(x)$.

   Доказательство.

   Шаг 1

   $f = \mathbb{1}_A$
   $E_{\mathbb{1}_A(\xi_1,\xi_2,\ldots,\xi_n)} = \int_\Omega \xi_1(w), \xi_2(w), \ldots, \xi_n(w)dP(w) =$

Это просто мера точек, в которых характеристическая функция единица.

$= P((\xi_1,\xi_2,\dots,\xi_n) \in A) = P_{\xi_1,\xi_2,\dots,\xi_n}(A) = \int_{\mathbb{R}^n} \mathbb{1} dP_{\xi_1,\xi_2,\dots,\xi_n}$

Шаг 2

Для простых функций сразу получили всё по линейности

Шаг 3

$f \geqslant 0$. Берём $\{\phi_k\}$ – простые, которые $\nearrow f$

$\int_{\mathbb{R}^n} \phi_k(x_1,x_2,\dots,x_n) dP_{\xi_1,\xi_2,\dots,\xi_n}(x_1,x_2,\dots,x_n) = E_{\phi_k(\xi_1,\xi_2,\dots,\xi_n)} =$

$= \int_\Omega \phi_k(\xi_1(w),\xi_2(w),\dots,\xi_n(w)) dP(w)$

Перейдём к пределу по монотонной последовательности $\phi_k \underset{k\to+\infty}{\to} f$ (так можно по теореме Беппо-Леви)

Тогда

$\int_{\mathbb{R}^n} \phi_k(x_1,x_2,\dots,x_n) dP_{\xi_1,\xi_2,\dots,\xi_n}(x_1,x_2,\dots,x_n) \to \int_{\mathbb{R}^n} f(x_1,x_2,\dots,x_n) dP_{\xi_1,\xi_2,\dots,\xi_n}(x_1,x_2,\dots,x_n)$

$\int_\Omega \phi_k(\xi_1(w),\xi_2(w),\dots,\xi_n(w)) dP(w) \to \int_\Omega f(\xi_1(w),\xi_2(w),\dots,\xi_n(w)) dP(w)$

Шаг 4

$f = f_+ - f_-$

□

7. Если ξ и η независимы, то $E_{\xi\eta} = E_\xi E_\eta$.

   Если $\xi_1,\xi_2,\dots,\xi_n$ независимы, то $E_{\xi_1\xi_2\dots\xi_n} = E_{\xi_1}E_{\xi_2}\dots E_{\xi_n}$

   (Если все эти интегралы существуют, конечно)

Доказательство.

$f(x_1,x_2,\dots,x_n) = x_1x_2\dots x_n$

$E(\xi_1,\xi_2,\dots,\xi_n) = \int_{\mathbb{R}^n} x_1x_2\dots x_n dP_{\xi_1,\xi_2,\dots,\xi_n}(x_1,x_2,\dots,x_n) =$

Мы знаем, что мера независимых с.в. – произведение мер по каждой координате

$= \int_{\mathbb{R}^n} x_1x_2\dots x_n dP_{\xi_1}(x_1)P_{\xi_2}(x_2)\dots P_{\xi_n} \underset{\text{по т. Тонелли}}{=} \int_{\mathbb{R}}\dots\int_{\mathbb{R}} x_1 dP_{\xi_1}(x_1) x_2 P_{\xi_2}(x_2)\dots x_n dP_{\xi_n}(x_n) =$
$\int_{\mathbb{R}}\dots\int_{\mathbb{R}} E_{\xi_1} x_2 P_{\xi_2}(x_2)\dots x_n dP_{\xi_n}(x_n) = E_{\xi_1}\int_{\mathbb{R}}\dots\int_{\mathbb{R}} x_2 P_{\xi_2}(x_2)\dots x_n dP_{\xi_n}(x_n) = E_{\xi_1}E_{\xi_2}\dots E_{\xi_n}$

(TODO: 180321: 19:25 - 20:30, 21:40 - 22:00 даётся аккуратное объяснение тому, почему последние переходы делать корректно, я нахожусь в процессе формулировки) □

8. Если $\xi \geqslant 0$, то $E_\xi = \int_0^\infty P(\xi \geqslant t) dt$

   Была в теории меры. Называлась интегралом по функции распределения

9. $E_{|\xi\eta|} \leqslant (E_{|\xi|^p})^{\frac{1}{p}} (E_{|\eta|^q})^{\frac{1}{q}}$, $\frac{1}{p} + \frac{1}{q} = 1$, $p,q > 0$

   Неравенство Гёльдера как оно есть для интегралов по мере.

10. Если $0 < r < s$, то $(E_{|\xi|^r})^{\frac{1}{r}} \leqslant (E_{|\xi|^s})^{\frac{1}{s}}$ (неравенство Ляпунова)

    Доказательство.

    Заметим, что достаточно проверить неравенство для $r = 1$:

    $\overline{\xi} \coloneqq |\xi|^r$, $E_{\overline{\xi}} \leqslant (E_{\overline{\xi}^{\frac{s}{r}}})^{\frac{r}{s}}$

    Проверим для $r = 1$.

    Берём $p = s > 1$ и соответствующее q. $\eta \equiv 1$. Тогда по неравенству Гёльдера:

    $E_\xi \leqslant (E_{|\xi|^s})^{\frac{1}{s}}(E_{|\eta|^q})^{\frac{1}{q}} = (E_{|\xi|^s})^{\frac{1}{s}}$ т.к. матожидание единицы – это единица. □

Замечание. $E_{\xi\eta} = E_\xi E_\eta$ без независимости неверно

Пример.

$\xi : \Omega \to \{\pm 1\}$, $P(\xi = -1) = P(\xi = +1) = \frac{1}{2}$

$E_\xi = 1 \cdot \frac{1}{2} + (-1) \cdot \frac{1}{2} = 0$

$\eta \coloneqq \xi$, $\xi\eta = \xi^2 \equiv 1$.

$E_{\xi\eta} = 1 \neq 0 = E_\xi E_\eta$

***Определение* 3.2** (Медиана)**.**

$\xi : \Omega \to \mathbb{R}$

Медиана ξ – это такое число $a \in \mathbb{R}$, что вероятность того, что $P(\xi \geqslant a) \geqslant \frac{1}{2}$ и $P(\xi \leqslant a) \geqslant \frac{1}{2}$

Замечание. Такое a может быть не единственно. Например, для примера выше любое число из $[-1, 1]$ является медианой.

Пример.

Есть 1000 человек.

999 из них получает 1000$, а их начальник 1000000$

Какая средняя зарплата с точки зрения матожидания?

$E_\xi = \frac{999}{1000} \cdot 1000 + \frac{1}{1000} \cdot 1000000 = 1999$

А если посчитать медиану? Медианой будет 1000 т.к. если взять меньше, то все зарплаты будут больше, если взять больше, то вероятность получить зарплату больше медианы слишком маленькая.

Эта задача иллюстрирует, что если нас интересует корректная средняя заплата, то правильнее считать медиану. В целом, во многих странах так и делают (не в России)...

***Определение* 3.3** (Дисперсия)**.**

$D_\xi \coloneqq E(\xi - E_\xi)^2$

Для того, чтобы существовала дисперсия обязательно нужно, чтобы существовало конечное математическое ожидание.

Свойства.

1. $D_\xi = E_{\xi^2} - (E_\xi)^2$

   Доказательство. $D_\xi = E_{(\xi - E_\xi)^2} = E_{\xi^2 - 2\xi E_\xi + (E_\xi)^2} = E_{\xi^2} - 2E_\xi E_\xi + E_{(E_\xi)^2} =$
   $E_{(E_\xi)^2}$ – это матожидание от константы, оно равно этой самой константе
   $= E_{\xi^2} - 2(E_\xi)^2 + (E_\xi)^2 = E_{\xi^2} - (E_\xi)^2$ □

2. $D_\xi \geqslant 0$ как интеграл от неотрицательной функции

3. Если $D_\xi = 0$, то $\xi = const$ с вероятностью 1.

   Доказательство. $E_{(\xi - E_\xi)^2} = 0$. Это значит, что интеграл неотрицательной функции равен нулю. Это значит, что функция равна нулю почти везде. На языке вероятностей это значит, что функция равна нулю с нулевой вероятностью.

   Т. е. $\xi - E_\xi = 0$ с вероятностью 1. Вспомним, что E_ξ – это константа. Получили, что $\xi = const$ с вероятностью 1. □

4. $D_{\xi+c} = D_\xi$

   Доказательство. $E_{\xi+c} = c + E_\xi \Rightarrow \xi - E_\xi = (\xi + c) - E_{\xi+c} = \xi + c - E_\xi - c$ □

5. $D_{c\xi} = c^2 D_\xi$

   Доказательство. $D_{c\xi} = E_{(c\xi - E_{c\xi})^2} = E_{(c(\xi - E_\xi))^2} = c^2 E_{(\xi - E_\xi)^2}$ □

6. $D(-\xi) = D(\xi)$ (Следует из того, что $(-1)^2 = 1$)

7. Если ξ, η независимы, то $D_{\xi+\eta} = D_\xi + D_\eta$

   Доказательство. $E_{(\eta+\xi)^2} = E_{\xi^2 + 2\xi\eta + \eta^2} = E_{\xi^2} + 2E_{\xi\eta} + E_{\eta^2} \underset{\text{по незав.}}{=} E_{\xi^2} + 2E_\xi E_\eta + E_{\eta^2}$
   Посмотрим теперь на $(E_{\eta+\xi})^2 = (E_\xi + E_\eta)^2 = (E_\xi)^2 + 2E_\xi E_\eta + (E_\eta)^2$
   По первому свойству, $D_{\eta+\xi} = E_{(\eta+\xi)^2} - (E_{\eta+\xi})^2 =$
   $= E_{\xi^2} + 2E_\xi E_\eta + E_{\eta^2} - ((E_\xi)^2 + 2E_\xi E_\eta + (E_\eta)^2) = E_{\xi^2} - (E_\xi)^2 + E_{\xi^2} - (E_\eta)^2 = D_\xi + D_\eta$ □

8. $E_{|\xi - E_\xi|} \leqslant \sqrt{D_\xi}$

   Доказательство. По неравенству Ляпунова, $r = 1, s = 2, \overline{\xi} := |\xi - E_\xi|$ □

Теорема 3.1 (Неравенство Чебышёва (Маркова)).
$t, p > 0 \Rightarrow R(|\xi| \geqslant t) \leqslant \frac{E_{|\xi|^p}}{t^p}$

Доказательство. У нас оно уже было, только вместо матожидания был интеграл. □

Следствие. $P(|\xi - E_\xi| \geqslant t) \leqslant \frac{D_\xi}{t^2}$.

Доказательство. В неравенство Чебышёва напишем $\xi - E_\xi$ вместо ξ, а вместо p напишем 2. □

Пример.

1. $\xi \sim U[0,1]$
   $E_\xi = \int_{\mathbb{R}} x dP_\xi(x) =.$
   В нашем случае $P_\xi(x)$ – это мера Лебега на $[0,1]$. Тогда
   $= \int_0^1 x\,dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2}$
   $E_{\xi^2} = \int_{\mathbb{R}} x^2 dP_\xi(x) = \int_0^1 x^2 dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}$
   $D_\xi = E_{\xi^2} - (E_\xi)^2 = \frac{1}{3} - (\frac{1}{2})^2 = \frac{1}{12}$

2. $\xi \sim U[a,b]$.
   $\xi = (b-a)\eta + a$, где $\eta \sim U[0,1]$.
   $E_\xi = E_{(b-a)\eta+a} = (b-a)E_\eta + a = (b-a)\frac{1}{2} + 1 = \frac{a+b}{2}$
   $D_\xi = D_{(b-a)\eta} + a = D_{(b-a)\eta} = (b-a)^2 D_\eta = \frac{(b-a)^2}{12}$

3. $\xi \sim \mathcal{N}(0,1)$
   $E_\xi = \int_{\mathbb{R}} x\,P_\xi(x) = \int_{\mathbb{R}} x \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 0$ т.к. нечётная ф-я.
   $E_{\xi^2} = \int_{\mathbb{R}} x^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (-x) d(e^{-x^2/2}) = -\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-x^2/2} d(-x) = 1$

   $D_\xi = E_{\xi^2} - E_\xi^2 = 1$

4. $\xi \sim \mathcal{N}(a, \sigma^2)$
   $\xi = \sigma\eta + a$, где $\eta \sim \mathcal{N}(0,1)$
   $E_\xi = E_{\sigma\eta+a} = \sigma E_\eta + a = a$
   $D_\xi = D_{\sigma\eta+a} = D_{\sigma\eta} = \sigma^2 D_\eta = \sigma^2$

Пример Независимое множество в графе.

$G = (V, E)$ – граф, n вершин и $\frac{nd}{2}$ рёбер для $d \geqslant 1$

Тогда в графе можно выбрать антиклику размера $\frac{n}{2d}$.

Доказательство. Выберем некоторый параметр $p \in [0, 1]$

Будем брать случайное множество вершин S: $P(s \in S) = p$

Рассмотрим подграф на S.

Если $\{x, y\} \in E$, то

$$\xi_{xy} = \begin{cases} 1 & \text{если } x, y \in S \\ 0 & \text{иначе} \end{cases}$$

$\xi = \sum_{\{x,y\} \in E} \xi_{xy}$

$E_\xi = \sum_{\{x,y\} \in E} E_{\xi_{xy}} = \frac{nd}{2} p^2$

Если $v \in V$, то

$$\eta_v = \begin{cases} 1 & \text{если } v \in S \\ 0 & \text{иначе} \end{cases}$$

$\eta = \sum_{v \in V} \eta_v$

$E_\eta = \sum_{v \in V} E_{\eta_v} = pn$

$E_{\eta - \xi} = pn - \frac{ndp^2}{2}$

Проще говоря, мы только что ввели величины количеств вершин и рёбер в случайном множестве S.

Выберем такое p, чтобы выражение выше имело максимальное значение

$p \coloneqq \frac{1}{d}$

$E_{\eta - \xi} = \frac{n}{2d}$.

Если среднее значение равно $\frac{n}{2d}$, то существует мно-во S, для которого $\eta - \xi \geqslant \frac{n}{2d}$. Величина $\eta - \xi$ – это количество вершин минус количество рёбер (для S она $\geqslant \frac{n}{2d}$). Внутри множества S есть сколько рёбер и сколько-то вершин. Давайте для каждого ребра в S выкинем какую-то одну вершину. Тогда рёбер не останется, а вершин будет хотя бы $\frac{n}{2d}$. И оставшиеся вершины будут образовывать независимое множество. □

Замечание. Обычно с помощью данного подсчёта матожидания можно многое относительно точно понять о комбинаторных объектах с неизвестной структурой. Так, для примера выше, очевидно, что детерминированным алгоритмом можно выбрать независимое множество получше. Но порядок величины очень часто оказывается весьма точным.

28.03.2018

Замечание. Мы знаем, что $E_\xi = \int_{\mathbb{R}} x dP_\xi(x)$. Как будет выглядеть эта формула, если мы что-нибудь знаем о распределении?

1. P_ξ – дискретная мера.
   $\xi : \Omega \to \{y_1, y_2, \dots\}$
   Тогда $E_\xi = \sum_{k=1}^\infty y_k P(\xi = y_k)$
2. P_ξ – абсолютно непрерывна, т.е. есть плотность $p_\xi(x)$
   Тогда $E_\xi = \int_{\mathbb{R}} x p_\xi(x) dx$

3.1. Корреляция и ковариация

***Определение* 3.4** (Ковариация)**.**
η, ξ – с.в. такие, что $E_{\xi^2} < +\infty$, $E_{\eta^2} < +\infty$
$cov(\xi, \eta) \coloneqq E((\xi - E_\xi)(\eta - E_\eta))$

Свойства.

1. $cov(\xi, \xi) = D_\xi$
2. $cov(\xi, \eta) = E(\xi\eta) - E_\xi E_\eta$

   Доказательство. $cov(\xi, \eta) = E(\xi\eta - \xi E_\eta - \eta E_\xi + E_\xi E_\eta) = E(\xi\eta) - E_\xi E_\eta - E_\eta E_\xi + E_\xi E_\eta$ $\square$
3. Если ξ и η независимы, то $cov(\xi, \eta) = 0$

   Доказательство. Для независимых η и ξ верно, что $E_{\xi\eta} = E_\xi E_\eta$. И см. пункт выше $\square$
4. $cov(\xi_1 + \xi_2, \eta) = cov(\xi_1, \eta) + cov(\xi_2, \eta)$
   $cov(c\xi, \eta) = c \cdot cov(\xi, \eta)$
5. $cov(\xi, \eta) = cov(\eta, \xi)$
6. $D(\xi + \eta) = D\xi + D\eta + 2cov(\xi, \eta)$
   $D(\sum_{k=1}^n \xi_k) = \sum_{k=1}^n D\xi_k + 2\sum_{i<k} cov(\xi_i, \xi_j)$

   Доказательство. Индукция.
   База $n = 2$.
   $D(\xi + \eta) = E(\xi + \eta)^2 - (E(\xi + \eta))^2 = E(\xi^2 + 2\xi\eta + \eta^2) - (E\xi)^2 - 2E\xi E\eta - (E\eta)^2$
   Переход $n \to n + 1$

$$D(\sum_{k=1}^{n+1} \xi_k) = D(\sum_{k=1}^{n} \xi_k + \xi_{n+1}) = D(\sum^{n} \xi_k) + D\xi_{n+1} + 2cov(\sum_{k=1}^{n} \xi_k, \xi_{n+1}) =$$
$$\underset{\text{по инд. и линейн.}}{=} \sum_{k=1}^{n} D\xi_k + 2\sum_{i<k\leqslant n} cov(\xi_i, \xi_k) + D\xi_{n+1} + 2\sum_{k=1}^{n} cov(\xi_k, \xi_{n+1})$$ □

Замечание. $cov(\xi, \eta) = 0 \not\Rightarrow \xi$ и η независимы.

Доказательство.

$\xi(w) = \cos(w)$

$\eta(w) = \sin(w)$

$\Omega = \{0, \pi/2, \pi\}$

$E_\xi = \frac{\cos 0 + \cos \pi/2 + \cos \pi}{3} = 0$

$\xi\eta \equiv 0 \Rightarrow E_{\xi\eta} = 0$

$E_\xi E_\eta = 0 \Rightarrow cov(\xi, \eta) = 0$

Но нет независимости:

$P(\xi = 1)P(\eta = 1) = \frac{1}{3} \cdot \frac{1}{3} \neq P(\xi = 1, \eta = 1) = 0$ □

***Определение* 3.5** (Коэффициент корреляции)**.**

$\rho(\xi, \eta) \coloneqq \frac{cov(\xi,\eta)}{\sqrt{D_\xi}\sqrt{D_\eta}}$

Случайные величины не коррелируют, если $cov(\xi, \eta) = 0$

Замечание.

$\rho(\xi, \eta) \in [-1, 1]$ т.к. по неравенству Коши-Буняковского:

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

Замечание. Рассмотрим $E_{\xi^2}, E_{\eta^2} < +\infty$

Тогда на пространстве таких величин (с конечным матожиданием квадрата), пофакторизованных по константе, можно определить скалярное произведение:

$\langle \xi, \eta \rangle \coloneqq cov(\xi, \eta)$

$\sqrt{D_\xi}$ – норма в таком пространстве (называется стандартным отклонением $\sigma(\xi)$)

Фактор по константе нужен для того, чтобы соблюдалось свойство: $\langle \xi, \xi \rangle = 0 \iff \xi = 0$

Пример. $\Omega \coloneqq \{1, 2, 3, \dots, n\}$

$\nu(k) \coloneqq$ количество различных простых в разложении k.

А теперь смотрим на теорему ниже:

Теорема 3.2 (Харди-Рамануджана)**.**

$\omega(n) \to +\infty$

Тогда $P(|\nu(k) - \ln\ln n| > \omega(n)\sqrt{\ln\ln n}) \to 0$ при $n \to +\infty$

Доказательство. Туран, 1935

Здесь p – простое число, а не плотность. $\xi_p(k) = \begin{cases} 1, \text{ если } p|k \\ 0, \text{ если } p \nmid k \end{cases}$

$M := \sqrt[10]{n}$

$\xi := \sum_{p\in\mathbb{P},\ p\leqslant M} \xi_p$

Это почти все делители n. Не учтены только те, которые больше M. Сколько таких? Не больше 10 т.к. каждое из них не меньше $\sqrt[10]{n}$.

$0 \leqslant \nu(k) - \xi(k) \leqslant 10$

Поэтому доказывать теорему что для ν, что для ξ без разницы.

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

E_{ξ_p} – это количество чисел, которые делятся на p, каждое из них с вероятностью $\frac{1}{n}$.

$E_{\xi_p} = \frac{[n/p]}{n} = \frac{n/p+\mathcal{O}(1)}{n} = 1/p + \mathcal{O}(1/n)$

$E_{\xi} = \sum_{p\in\mathbb{P},\ p\leqslant M} E_{\xi_p} = \sum_{p\in\mathbb{P},\ p\leqslant M} 1/p + \mathcal{O}(M/n) \underset{\Sigma \text{ обратных к простым}}{=} \ln\ln M + \mathcal{O}(1) = \ln\ln n + \mathcal{O}(1)$

Получили, что в левой части неравенства написано практически $|\xi - E_\xi|$.

$D_\xi = \sum_{p\in\mathbb{P},\ p\leqslant M} D_{\xi_p} + 2\sum_{p,\,q\in\mathbb{P},\ p<q\leqslant M} cov(\xi_p, \xi_q)$

$D_{\xi_p} = E_{\xi_p^2} - (E_{\xi_p})^2 \underset{\text{т.к. } \xi_p=0|1}{=} E_{\xi_p} - (E_{\xi_p})^2 = 1/p + \mathcal{O}(1/n) - 1/p^2 + \mathcal{O}(1/n)$

$cov(\xi_p, \xi_q) = E(\xi_p\xi_q) - E_{\xi_p}E_{\xi_q} \underset{\xi_p\xi_q=1 \iff (pq)|k}{=} \frac{[\frac{n}{pq}]}{n} - \frac{[n/p]}{n}\frac{[n/q]}{n}$

$cov(\xi_p, \xi_q) \geqslant \frac{\frac{n}{pq}-1}{n} - \frac{n/p}{n}\frac{n/q}{n} = \frac{1}{pq} - \frac{1}{n} - \frac{1}{p}\frac{1}{q} = -\frac{1}{n}$

$2\sum cov(\xi_p, \xi_q) \geqslant -\frac{M^2}{n} = -\frac{\sqrt[5]{n}}{n} = \mathcal{O}(1)$

$cov(\xi_p, \xi_q) \leqslant \frac{\frac{n}{pq}}{n} - \frac{(\frac{n}{p}-1)(\frac{n}{q}-1)}{n^2} = \frac{1}{pq} - \frac{1}{p}\frac{1}{q} + \frac{1}{n}\frac{1}{p} + \frac{1}{n}\frac{1}{q} - \frac{1}{n^2} \leqslant \frac{1}{n}(\frac{1}{p} + \frac{1}{q})$

$2\sum cov(\xi_p, \xi_q) \leqslant \frac{1}{n}\sum_{q,p\in\mathbb{P},p\neq q}(\frac{1}{p} + \frac{1}{q})$

Сколько раз мы посчитаем $\frac{1}{p}$? С каждой $\frac{1}{q}$ по разу. То есть не больше, чем количество простых.

$2\sum cov(\xi_p, \xi_q) \leqslant \frac{1}{n}\sum_{q,p\in\mathbb{P},p\neq q}(\frac{1}{p} + \frac{1}{q}) \leqslant 2\frac{1}{n}M\sum_{p\leqslant M,p\in\mathbb{P}} \frac{1}{p} = 2\frac{\sqrt[10]{n}\ln\ln n}{n} = \mathcal{O}(1)$

$D_\xi = \sum_{p\in\mathbb{P},\ p\leqslant M} D_{\xi_p} + 2\sum_{p,\,q\in\mathbb{P},\ p<q\leqslant M} cov(\xi_p, \xi_q) = \sum_{p\leqslant M,p\in\mathbb{P}}(\frac{1}{p} - \frac{1}{p^2}) + \mathcal{O}(1)$

Ряд из обратных квадратов сходится, значит, они тоже наберут не больше, чем константу, положим $\frac{1}{p^2}$ в $\mathcal{O}(1)$

$D_\xi = \sum_{p\leqslant M,p\in\mathbb{P}} \frac{1}{p} + \mathcal{O}(1) = \ln\ln n + \mathcal{O}(1)$

Неравенство Чебышёва: $P(|\xi - E_\xi| \geqslant \lambda\sqrt{D_\xi}) \leqslant \frac{D_\xi}{(\lambda\sqrt{D_\xi})^2} = \frac{1}{\lambda^2}$, где

$\lambda = \omega(n)$

$E_\xi = \ln\ln n + \mathcal{O}(1)$

$\sqrt{D_\xi} = \sqrt{\ln\ln n} + \mathcal{O}(1)$

Почему $\mathcal{O}(1)$ можно стереть? Потому что λ – это штука стремящаяся к $+\infty$. Поэтому если стереть $\mathcal{O}(1)$ в левой части, то нужно будет прибавить в правой. Но оно попадёт туда само т.к. $\mathcal{O}(1)$ – это нечто ограниченное, а λ стремится к бесконечности. □

***Определение* 3.6.** k-й момент

E_{ξ^k} – k-й момент случайной величины ξ

$E_{\xi^k} = \int_{\mathbb{R}} x^k dP_\xi(x)$

Замечание. Польза их в том, что многие теоремы в формулировке требуют, чтобы какой-то момент был конечен. А ещё например, неравенство Ляпунова говорит о том, что если конечен момент какого-то порядка, то моменты всех порядков меньше тоже конечны.

***Определение* 3.7.**

$E|\xi - E_\xi|^k$ – k-й центральный момент

D_ξ – второй центральный момент

3.2. Сходимость случайных величин и закон больших чисел

***Определение* 3.8.** $\xi_n, \xi : \Omega \to \mathbb{R}$

1. ξ_n сходится к ξ почти наверное (= с вероятностью 1), если $P(\omega \in \Omega : \lim\limits_{n\to\infty} \xi_n(\omega) = \xi(\omega)) = 1$

2. ξ_n сходится к ξ в среднем порядка $r > 0$, если $E|\xi_n - \xi|^r \to 0$ при $n \to \infty$

3. ξ_n сходится к ξ по вероятности

   $\forall \varepsilon > 0\ P(|\xi_n - \xi| > \varepsilon) \to 0$ при $n \to \infty$

4. ξ_n сходится к ξ по распределению

   F_{ξ_n} сходится к F_ξ во всех точках непрерывности F_ξ. ($\sim$ теорема Муавра-Лапласа)

Замечание. Связь между сходимостями TODO: [Вычитать доказательства]

$1 \Rightarrow 3$ теорема Лебега

$3 \not\Rightarrow 1$ был пример

$2 \Rightarrow 3$ $P(|\xi_n - \xi| > \varepsilon) = P(|\xi_n - \xi|^r > \varepsilon^r) \leqslant \frac{E|\xi_n - \xi|^r}{\varepsilon^r} \to 0$ (Неравество Чебышёва)

$1 \not\Rightarrow 2$ (а значит, и $3 \not\Rightarrow 2$ т.к. из $1 \Rightarrow 3$ и мы получим пример, где 3, но не 2)

$\Omega = [0, 1]$

$\xi_n = n^{1/r} \mathbb{1}_{[0,1/n]} \to \xi \equiv 0$ сходится почти наверное

Но $E_{\xi_n^r} = E(n \mathbb{1}_{[0,1/n]}) = 1 \not\to 0$

$2 \not\Rightarrow 1$ (а значит $3 \not\Rightarrow 1$)

Контрпример: $\xi_{n,k} \coloneqq \mathbb{1}_{[k/n,(k+1)/n]}$

$E\xi_{n,k}^r = 1/n^r \to 0$

$3 \Rightarrow 4$ Пусть F_ξ непрерывна в точке x.

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

$\{\xi_n > x\} \supset \{\xi > x + \varepsilon\} \cap \{|\xi_n - \xi < \varepsilon|\}$

$P(\xi_n > x) \geqslant P(\dots)$

$\{\xi_n \leqslant x\} \subset \{\xi \leqslant x + \varepsilon\} \cup \{|\xi_n - \xi| \geqslant \varepsilon\}$

$F_{\xi_n}(x) \leqslant F_\xi(x + \varepsilon) + P(|\xi_n - \xi| \geqslant \varepsilon)$

$\overline{\lim\limits_{n\to\infty}} F_{\xi_n}(x) \leqslant F_\xi(x_\varepsilon) + \overline{\lim\limits_{n\to\infty}} P(|\xi_n - \xi| \geqslant \varepsilon) = F_\xi(x_\varepsilon)$

$\{\xi_n \leqslant x\} \supset \{\xi \leqslant x - \varepsilon\} \cap \{|\xi_n - \xi| < \varepsilon\}$

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

$1 - F_{\xi_n}(x) \leqslant 1 - F_\xi(x - \varepsilon) + P(|\xi_n - \xi| \geqslant \varepsilon)$

$F_{\xi_n}(x) \geqslant F_\xi(x - \varepsilon) - P(|\xi_n - \xi| \geqslant \varepsilon)$

$\underline{\lim}_{n\to\infty} F_{\xi_n}(x) \geqslant F_\xi(x - \varepsilon)$

$F_\xi(x + \varepsilon) \geqslant \overline{\lim}_{n\to\infty} F_{\xi_n}(x) \geqslant \underline{\lim}_{n\to\infty} F_{\xi_n}(x) \geqslant F_\xi(x - \varepsilon) \to F_\xi(x)$

$F_\xi(x + \varepsilon) \to F_\xi(x)$ при $\varepsilon \to 0$

$4 \not\Rightarrow 3$ Некорректна такая постановка вопроса

3.3. Закон больших чисел

Теорема 3.3. $\xi_1, \xi_2, \xi_3, \ldots$ попарно независимы

$\sup_n D_{\xi_n} < +\infty$, $S_n := \xi_1 + \xi_2 + \ldots + \xi_n$

Тогда $P(|\frac{S_n}{n} - E(\frac{S_n}{n})| > t) \underset{n\to\infty}{\to} 0$

Доказательство. Неравенство Чебышёва:

$P(|\frac{S_n}{n} - E\frac{S_n}{n}| > t) \leqslant \frac{D(\frac{s_n}{n})}{t^2} = \frac{D(S_n)}{n^2t^2} \underset{\text{т.к. } cov=0}{=} \frac{\sum_{k=1}^n D_{\xi_k}}{n^2t^2}$

Ну а т.к. дисперсии ограничены, то

$\frac{\sum_{k=1}^n D_{\xi_k}}{n^2t^2} \leqslant \frac{cn}{n^2t^2} = \frac{c}{nt^2} \to 0$ □

Замечание. Утверждения такого сорта называются законом больших чисел. А именно: на бесконечности вероятность среднего значения сильно отличаться от матожидания среднего значения стремится у нулю.

Следствие Закон больших чисел в форме Чебышёва.

$\xi_1, \xi_2, \ldots$ – независимые одинаково распредленные с.в.

$a := E_{\xi_1}, \sigma^2 := D_{\xi_1}, S_n := \xi_1 + \xi_2 + \ldots + \xi_n$

Тогда $P(|\frac{S_n}{n} - a| > t) \to 0$

Доказательство. Т.к. все величины одинаково распределены, то матожидание и дисперсия у них у всех одинаковые. Значит, $E(\frac{S_n}{n}) = \frac{nE_{\xi_1}}{n} = a$. □

Следствие ЗБЧ для схемы Бернулли.

p – вероятность успеха в схеме Бернулли.

Тогда $P(|\frac{S_n}{n} - p| > t) \to 0$

Доказательство. Вообще, это частный случай предыдущего следствия. $E_\xi = p$. □

Теорема 3.4 (Усиленный ЗБЧ)**.** $\xi_1, \xi_2, \ldots$ – независимые случайные величины.

$E|\xi_k - E_{\xi_k}|^4 \leqslant C, S_n := \xi_1 + \xi_2 + \ldots + \xi_n$

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

Доказательство. $\overline{\xi_k} := \xi_k - E_{\xi_k}$ – всё ещё независимы т.к. мы из каждой просто вычли константу.

$E_{\overline{\xi_k}} = 0$, надо доказать, что $\frac{\bar{\xi}_1 + \bar{\xi}_2 + \ldots + \bar{\xi}_n}{n} \to 0$ почти наверное.

Будем считать, что $E_{\xi_k} = 0 \,\forall k$.

$A_n := \{|\frac{S_n}{n}| > \varepsilon\}$ – вот такое множество событий

Как устроено множество $\{\omega\}$, на которых нет стремления к нулю? Это ровно те ω, которые попадут в бесконечное количество A_n

$A := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$ – в точности то множество, на котором нет стремления к нулю.

$A = \{w \in \Omega : |\frac{S_n}{n}| \nrightarrow 0\}$

по лемме Бореля-Кантелли, из $\sum_{n=1}^{\infty} P(A_n) < +\infty$ следует, что $P(A) = 0$

Проверим, что $\sum_{n=1}^{\infty} P(A_n) < +\infty$.

$P(A_n) = P(|\frac{S_n}{n}| > \varepsilon^4) \underset{\text{по нер-ву Чернышёва}}{\leqslant} \frac{E(\frac{S_n^4}{n^4})}{\varepsilon^4} = \frac{E_{S_n^4}}{n^4\varepsilon^4}$

Достаточно доказать, что $E_{S_n^4} \leqslant cn^2$ для некоторой константы c, т.к. тогда мы получим, что $\frac{E_{S_n^4}}{n^4\varepsilon^4} \leqslant \frac{c}{\varepsilon^4}\frac{1}{n^2}$, ну а это сходится.

$S_n^4 = (\xi_1+\xi_2+\ldots+\xi_n)^4 \approx \sum_{k=1}^{n} \xi_k^4 + 6\sum_{k<j} \xi_k^2\xi_j^2 + 12\sum \xi_i\xi_j\xi_k + 24\sum \xi_i\xi_k\xi_k\xi_l + 4\sum \xi_j^3\xi_k$

$E_{S_n^4} = \sum E_{\xi_k^4} + 6\sum E(\xi_j^2\xi_k^2)$ – все остальные слагаемые самоуничтожились за счёт того, что матожидание произведения независимых с.в. превратится в произведение матожиданий, а матожидания величин в первой степени у нас нули.

$E_{S_n^4} = \sum E_{\xi_k^4} + 6\sum E(\xi_j^2\xi_k^2) \leqslant Cn + 6\sum \sqrt{E_{\xi_k^4}}\sqrt{E_{\xi_j^4}}$, где оценка на $E_{\xi_k^4}$ у нас есть из условия, а на сумму произведений – из неравенства Коши-Буняковского. Ну а $\sqrt{E_{\xi_j^4}} \leqslant \sqrt{C}$. Получим:

$E_{S_n^4} \leqslant Cn + 6\frac{n(n-1)}{2}C \leqslant 3Cn$ □

Следствие УЗБЧ для схемы Бернулли.

p –вероятность успеха.

$\frac{S_n}{n} \to p$ почти наверное.

Доказательство. Нужно понять, что 4-й момент будет конечен.

$E(\xi_{-}p)^4 = E\xi_k^4 - \ldots = E\xi_k - \ldots = p - \ldots$ т.к. ξ_k принимает значения 0 или 1. Поэтому у подбрасывания монетки все нецентральные моменты одинаковы, а центральные моменты конечны. □

Замечание. Для любой случайной величины, которая принимает конечное количество значений можно понять, что все моменты существуют и конечны.

Теорема 3.5 (УЗБЧ в форме Колмогорова).

$\xi_1, \xi_2, \ldots$ – независимые одинаково распределённые с.в.

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

Тогда $\frac{S_n}{n} \to a$ почти наверное $\iff a = E_{\xi_1}$

Доказательство. Его нет и не будет. □

Следствие Метод Монте-Карло. Φ – фигура на плоскости.

Пусть мы легко умеем отвечать на запрос: по заданной точке сказать, лежит она в фигуре или нет.

Например, если фигура задаётся системой неравенств.

Мы хотим найти площадь фигуры.

Возьмём прямоугольник, который ограничивает фигуру и начинаем кидать в него случайные точки (желательно как можно более плотно описывающий фигуру).

$\xi_k = 1$, если мы попали в Ф и $\xi_k = 0$, если нет.

Закон больших чисел говорит, что $\frac{S_n}{n} \to p$ почти наверное.

$p = \frac{\text{площадь } \Phi}{\text{площадь прямоугольника}}$

Для крупных прикидок метод работает хорошо.

4. Характеристическая функция

Теорема 4.1. $\xi_1, \xi_2, \dots$ –независимые случайные величины.

$f_i : \mathbb{R}^{n_i} \to \mathbb{R}$ – измеримые относительно Борелевской σ-алгебры.

Тогда $f_1(\xi_1, \dots, \xi_{n_1}), f_2(\xi_{n_1+1}, \dots, \xi_{n_2+n_2})$ – независимые случайные величины.

Доказательство. Докажем для двух функции, остальное аналогично.

$f(\xi_1, \dots, \xi_n)$ и $g(\eta_1, \dots, \eta_m)$ независимы.

$\{f(\xi_1, \dots, x_n) \in A\}$ и $\{g(\eta_1, \dots, \eta_m) \in B\}$ независимые события.

A и B – борелевские множества на прямой.

Т.к. такие множества порождаются интервалами, то можно считать, что A и B – интервалы.

Нас интересуют прообразы этих штук. Т.е. $(\xi_1, \dots, x_n) \in f^{-1})(A)$ и $(\eta_1, \dots, \eta_m) \in g^{-1}(B)$.

Эти прообразы – Борелевские множества (TODO: почему?). Значит, нужно проверить, что независимые события: один вектор попадает в некоторое борелевское подмножество в $\mathbb{R}^n$ и второй в другое борелевское подмножество в $\mathbb{R}^m$.

Снова, т.к. множества Борелевские, то достаточно проверить, что всё выполняется на ячейках.

Т.е.:

$(\xi_1, \dots, \xi_n) \in [a, b)^n$ и $(\eta_1, \dots, \eta_m) \in [c, d)^m$.

Т.к. события независимы, то

$P(X) := P((\xi_1, \dots, \xi_n) \in [a, b)^n) = P(\xi_1 \in [a, b))P(\xi_2 \in [a, b)) \dots P(\xi_n \in [a, b))$ и

$P(Y) := P((\eta_1, \dots, \eta_n) \in [c, d)^m) = P(\eta_1 \in [c, d))P(\eta_2 \in [c, d)) \dots P(\eta_m \in [c, d))$

$P(X \cap Y) = P((\eta_1, \dots, \eta_n) \in [c, d)^m \wedge (\xi_1, \dots, \xi_n) \in [a, b)^n) =$

$= P(\xi_1 \in [a, b)) \dots P(\xi_n \in [a, b))P(\eta_1 \in [c, d)) \dots P(\eta_m \in [c, d)) = P(X)P(Y)$ □

***Определение* 4.1.** ξ – случайная величина, принимающая значения из $\mathbb{N}_0$.

$G_\xi(t) := \sum_{n=0}^{\infty} P(\xi = n)t^n$ – произв. ф-я для ξ

Свойства.

1. $G_\xi = Et^\xi$

   Доказательство.

   $G_\xi(t) = \sum_{n=0}^{\infty} P(\xi = n)t^n$

   $E_{t^\xi} = \sum P(t^\xi = t^n)t^n$

   $E_{t^\xi} = G_\xi(t)$ при $t \in (0, 1)$.

   Этого достаточно, чтобы сказать, что эти два ряда равны везде (см. ТФКП). □

2. $G_\xi(1) = 1$ и ряд сходится в единичном круге.

   Доказательство.

   $G_\xi(1) = \sum_{n=0}^{\infty} P(\xi = n) = 1$

   (вообще, это вероятность того, что ξ принимает значение из $\mathbb{N}_0$) $\square$

3. $G'_\xi(1) = E_\xi$

   Доказательство.

   $E_\xi = \sum_{n=0}^{\infty} nP(\xi = n)$

   $G'_\xi(t) = \sum_{n=0}^{\infty} nP(\xi = n)t^{n-1}$ $\square$

4. $E_{\xi^2} = G''_\xi(1) + G'_\xi(1)$

   Доказательство.

   $E_{\xi^2} = \sum n^2 P(\xi = n) = \sum n(n-1)P(\xi = n) + \sum nP(\xi = n) = G''_\xi(1) + G'_\xi(1)$ $\square$

5. $D_\xi = G''_\xi(1) + G'_\xi(1) - G'_\xi(1)^2$

   Доказательство.

   $D_\xi = E_{\xi^2} - (E_\xi)^2$. $\square$

6. $G_{\xi+\eta}(t) = G_\xi(t)G_\eta(t)$, если ξ и η независимые.

   Доказательство.

   t^ξ и t^η – независимы, тогда $G_{\xi+\eta}(t) = E(t^\xi t^\eta) = (Et^\xi)(Et^\eta) = G_\xi(t)G\eta(t)$ $\square$

***Определение* 4.2** (Комплекснозначная с.в.)**.**

$\xi : \Omega \to \mathbb{C}$ измеримая.

$\xi = \operatorname{Re}\xi + i\operatorname{Im}\xi$, где $\operatorname{Re}\xi$ и $\operatorname{Im}\xi$ – это тоже какие-то случайные величины.

***Определение* 4.3.** $E_\xi := E(\operatorname{Re}\xi) + iE(\operatorname{Im}\xi)$ Все свойства с равенствами для матожидания сохраняются.

Упражнение. Доказать комплексную линейность E: $E(i\xi) = iE(\xi)$.

Свойство. $|E_\xi| \leqslant E_{|\xi|}$

Доказательство. Подберём $c \in \mathbb{C}$, $|c| = 1$ такое, что $E(c\xi) = |E_\xi|$.

Например, возьмём $c := \frac{\overline{E_\xi}}{|E_\xi|}$ Тогда $E(c\xi) = cE(\xi) = \frac{E_\xi \overline{E_\xi}}{|E_\xi|} = |E_\xi|$.

При этом т.к. $|c| = 1$, то $|c\xi| = |\xi|$.

$E(\operatorname{Re}(c\xi)) \underset{\text{нер-во для вещ. матож.}}{\leqslant} E|\operatorname{Re}(c\xi)| \leqslant E|c\xi| = E|\xi|$.

С другой стороны, $E(\operatorname{Re}(c\xi)) = \operatorname{Re} E(c\xi) = |E_\xi|$ т.к. мы правильно подобрали c. □

***Определение* 4.4** (Комплекснозначная ковариация)**.**

$cov(\xi, \eta) := E(\xi - E_\xi)\overline{(\eta - E_\eta)}$

$cov(\xi, \xi) = D_\xi$

***Определение* 4.5.** $\varphi_\xi(t) := Ee^{it\xi}$ – характеристическая функция ξ.

Свойства.

1. $\varphi_\xi(0) = 1$ и $|\varphi_\xi(t)| \leqslant 1\, \forall t \in \mathbb{R}$

   Доказательство. $|\varphi_\xi(t)| = |Ee^{it\xi}| \leqslant E|e^{it\xi}| = E1 = 1$ □

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

   Доказательство. $\varphi_{a\xi+b}(t) = E(e^{i(e\xi+b)t}) = e^{ibt}Ee^{ia\xi t} = e^{ibt}\varphi_\xi(at)$ □

3. Если ξ_1,$\xi_2, \ldots$,ξ_n – независимые, то $\varphi_{\xi_1+\xi_2+\ldots+\xi_n}(t) = \varphi_{\xi_1}(t)\varphi_{\xi_2}(t)\ldots\varphi_{\xi_n}(t)$

   Доказательство. $\varphi_{\xi+\eta}(t) = Ee^{it(\xi+\eta)} = E(e^{it\xi}e^{it\eta}) \underset{\text{т.к. } f_1(\xi) \text{ и } f_2(\eta) \text{ незав.}}{=} = (Ee^{et\xi})(Ee^{it\eta}) = \varphi_\xi(t)\varphi_\eta(t)$ □

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

   Доказательство. $\varphi_\xi(-t) = Ee^{-it\xi} = E\overline{e^{it\xi}} \underset{\text{т.к. } E = \operatorname{Re} E + i \operatorname{Im} E}{=} \overline{Ee^{it\xi}} = \overline{\varphi_\xi(t)}$ □

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

   Доказательство. $\varphi_\xi(t+h) - \varphi_\xi(t) = Ee^{i\xi(t+h)} - Ee^{it\xi} = E(e^{it\xi}e^{ih\xi} - e^{it\xi}) = E(e^{it\xi}(e^{ith} - 1))$

   Мы хотим понять, что это равномерно стремится к нулю.

   $|\varphi_\xi(t+h) - \varphi_\xi(t)| \leqslant E|e^{it\xi}(e^{ih\xi-1})| \underset{\text{т.к. } |e^{it\xi}|=1}{=} E|e^{ih\xi} - 1| \underset{h \to 0}{\to} 0$, где последний переход сделан по теореме Лебега о мажорируемой сходимости: по каждому фиксированному ξ сходимость к нулю есть, предъявим мажоранту: $|e^{ih\xi} - 1| \leqslant 2$.

   От t при этом наша функция уже не зависит. Значит, из того, что это матожидание просто стремится к нулю, следует равномерная сходимость по t к нулю, а значит и равномерная непрерывность.

$\square$

11.04.2018

Пример.

Есть $\xi \sim \mathcal{N}(a, \sigma^2)$, хочется посчитать характеристическую функцию для такого распределения.

Представим, что мы уже знаем, как посчитать хар. ф-ю от $\eta \sim \mathcal{N}(0,1)$

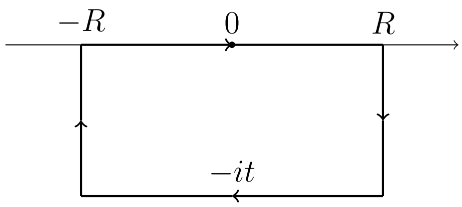
$\xi = \sigma\eta + a$

Тогда можно выразить новую хар. ф-ю через известную: $\varphi_\xi(t) = e^{iat}\varphi_\eta(\sigma t)$

Осталось посчитать хар. ф-ю для η:

$\varphi_\eta(t) = Ee^{it\eta} = \int\limits_{-\infty}^{\infty} e^{itx}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx = \frac{e^{-t^2/2}}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{-(x-it)^2/2}dx$

Давайте считать интеграл $\int\limits_{\Gamma_R} e^{-z^2/2}dz$ по следующему контуру:



Т.к. особых точек нету, то $\int = 0$.

$0 = \int\limits_{\Gamma_R} e^{-z^2/2}dz = \int\limits_{-R}^{R}$ верхняя часть $- \int\limits_{-R}^{R}$ нижняя часть $+ \int\limits_{0}^{-t}$ правая часть $- \int\limits_{0}^{-t}$ левая часть $=$
$= \int\limits_{-R}^{R} e^{-x^2/2}dx - \int\limits_{-R}^{R} e^{-(x-it)^2/2}dx + \int\limits_{0}^{-t} e^{-(R+iy)^2/2}dy + \int\limits_{0}^{-t} e^{-(-R+iy)^2/2}dy \underset{\text{см. ниже}}{\rightarrow} \sqrt{2\pi} - \int\limits_{-R}^{R} e^{-(x-it)^2/2}dx + 0$

$|e^{-(\pm R+iy)^2/2}| = e^{-R^2/2}e^{y^2/2} \to 0$ т.к. $e^{y^2/2}$ ограничено, а $e^{-R^2/2} \underset{R\to\infty}{\rightarrow} 0 \Rightarrow$ интегралы от левой и правой частей уходят в 0.

$\sqrt{2\pi} = \int\limits_{-\infty}^{\infty} e^{-x^2/2}dx$ – почти Гауссов интеграл, несложно проверить.

Итого: $0 = \sqrt{2\pi} - \int\limits_{-R}^{R} e^{-(x-it)^2/2}dx + 0 \Rightarrow \sqrt{2\pi} = \int\limits_{-\infty}^{\infty} e^{-(x-it)^2/2}dx$

$\varphi_\eta(t) = \frac{e^{-t^2/2}}{\sqrt{2\pi}}\sqrt{2\pi} = e^{-t^2/2}$

$\varphi_\xi(t) = e^{iat}\varphi_\eta(\sigma t) = e^{iat-\sigma^2t^2/2}$

Теорема 4.2. Пусть $E|\xi|^n < +\infty$. Тогда $\forall k = 1, 2, \dots, n$

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

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

Доказательство. Индукция по k. База $k = 0$ – определение φ_ξ.

Переход $k \to k+1$:

$(\varphi_\xi(t))^{(k+1)} = \lim\limits_{h\to 0} \frac{(\varphi_\xi(t+h))^{(k)} - (\varphi_\xi(t))^{(k)}}{h} = \lim\limits_{h\to 0} \frac{E((i\xi)^k e^{i(t+h)\xi}) - E((i\xi)^k e^{it\xi})}{h} \underset{\text{по линейности}}{=}$

$= \lim\limits_{h\to 0} E((i\xi)^k e^{it\xi} \frac{e^{ih\xi}-1}{h}) = \lim\limits_{h\to 0} \int_{\mathbb{R}} (ix)^k e^{itx} \frac{e^{ihx}-1}{h} dP_\xi(x)$

Хотим показать, что интеграл можно переставлять местами с пределом. Если мы это покажем, то $\lim\limits_{h\to 0} \frac{e^{ih\xi}-1}{h} = i\xi$ (см. второй замечательный предел) даст нам нужную формулу. Чтобы это показать, предъявим суммируемую мажоранту.

$|(ix)^k e^{itx} \frac{e^{ihx}-1}{h}| = x^k |\frac{e^{ihx}-1}{h}| \underset{\text{см. ниже}}{=} x^k \mathcal{O}(x) = \mathcal{O}(x^{k+1})$ – суммируемая мажоранта?

Есть два принципиально разных случая.

Если $|xh| \geqslant 1 \Rightarrow \frac{1}{|h|} \geqslant |x|$ и т.к. $|e^{ihx} - 1| \leqslant 2$, то $|\frac{e^{ihx}-1}{h}| \leqslant \frac{2}{|h|} = \mathcal{O}(x)$.

Если же $|xh| < 1$, то e^{ihx} раскладывается в ряд и мы получим $\frac{e^{ihx}-1}{h} = \frac{1+ihx+\mathcal{O}(ihx)-1}{h} = ix+\mathcal{O}(x)$. Ну и тогда $|ix + \mathcal{O}(ix)| = \mathcal{O}(x)$

Итого, мы поняли, что $|(ix)^k e^{itx} \frac{e^{ihx}-1}{h}| = \mathcal{O}(x^{k+1})$. Из того, что все моменты вплоть до n конечны и $k+1 \leqslant n$ следует, что $\int\limits_{\mathbb{R}} x^{k+1} dP_\xi(x) < +\infty$, а значит, x^{k+1} действительно мажоранта и можно переставить предел с интегралом.

$\square$

Следствие. $E\xi = -i\varphi'_\xi(0)$

$D\xi = E\xi^2 - (E\xi)^2 = -\varphi''_\xi(0) + (\varphi'_\xi(0))^2$

Теорема 4.3. Если существует $\varphi''_\xi(0)$, то $E\xi^2 < +\infty$.

Доказательство. $E\xi^2 = \int\limits_{-\infty}^{+\infty} x^2 dP_\xi(x) = \int\limits_{-\infty}^{+\infty} \lim\limits_{t\to 0} (\frac{\sin(tx)}{t})^2 dP_\xi(x) \leqslant$

$(\sin(tx))^2 = (\frac{e^{itx} - e^{-itx}}{2i})^2 = -\frac{1}{4}(e^{2itx} + e^{-2itx} - 2)$

По лемме Фату, можно оценить интеграл от предела сверху, если переставить их местами:

$\leqslant \lim\limits_{t\to 0} \int\limits_{-\infty}^{+\infty} (\frac{\sin tx}{t})^2 dP_\xi(x) = \lim\limits_{t\to 0} -\frac{1}{4t^2} \int\limits_{-\infty}^{+\infty} (e^{2itx} + e^{-2itx} - 2) dP_\xi(x) =$

$= -\frac{1}{4} \lim\limits_{t\to 0} \frac{1}{t^2}(\varphi_\xi(2t) + \varphi_\xi(-2t) - 2\varphi_\xi(0)) =$

$\varphi_\xi(s) = \varphi_\xi(0) + s\varphi'_\xi(0) + \frac{s^2}{2}\varphi''_\xi(0) + o(s^2)$ – это нужно подставить в формулу для $\varphi_\xi(2t)$ и $\varphi_\xi - (2t)$, поприводить подобные слагаемые и получим:

$$= -\frac{1}{4}\lim_{t\to 0}\frac{(2t)^2\varphi_\xi''(0)+o(t^2)}{t^2} = -\varphi_\xi''(0)$$

□

Замечание. Если существует $\varphi^{(2n)}(0)$, то $E\xi^{2n} < +\infty$ (Без док-ва.)

Теорема 4.4. Формула Обращения

$a < b$, $P_\xi(\{a\}) = P_\xi(\{b\}) = 0$

(иначе говоря $P(\xi = a) = P(\xi = b) = 0$)

$$P(a \leqslant \xi \leqslant b) = \lim_{T\to+\infty}\frac{1}{2\pi}\int\limits_{-T}^{+T}\frac{e^{-iat}-e^{-ibt}}{it}\varphi_\xi(t)dt$$

(Можно воспринимать это как интеграл по главному значению от $-\infty$ до $+\infty$)

Доказательство. Хотим линейное преобразование, после которого получится, что: $\{\xi \in [a,b]\} \Leftrightarrow \{\eta \in [-1,1]\}$

$$\xi = \frac{a+b}{2} + \frac{b-a}{2}\eta$$

$$\varphi_\xi(t) = e^{i\frac{a+b}{2}t}\varphi_\eta(\tfrac{b-a}{2}t)$$

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

$$= \int\limits_{-T}^{+T}\frac{e^{-i\frac{a-b}{2}t}-e^{-i\frac{b-a}{2}t}}{it}\varphi_\eta(\tfrac{b-a}{2}t)dt \underset{s:=\frac{b-a}{2}t}{=} \int\limits_{-\frac{b-a}{2}T}^{\frac{b-a}{2}T}\frac{e^{is}-e^{is}}{is}\varphi_\eta(s)ds$$

Пределы интегрирования по-прежнему симметричны и стремятся к бесконечности, $\frac{b-a}{2} > 0$. Тогда достаточно проверить для $[a,b] = [-1,1]$

$$\int\limits_{-T}^{T}\frac{e^{it}-e^{-it}}{it}\varphi_\xi(t)dt = \int\limits_{-T}^{T}\frac{e^{it}-e^{it}}{-it}\int\limits_{-\infty}^{+\infty}e^{itx}dP_\xi(x)dt \underset{\text{т. Фубини}}{=} \int\limits_{-\infty}^{+\infty}\int\limits_{-T}^{T}\frac{e^{it}-e^{-it}}{it}e^{itx}dtdP_\xi(x)$$

$$\Phi_T(x) := \int\limits_{-T}^{T}\frac{e^{it}-e^{-it}}{it}e^{itx}dt$$

$$\Phi_T(x) = \int\limits_{-T}^{T}\int\limits_{-1}^{1}e^{itu}du e^{itx}dt = \int\limits_{-1}^{1}\int\limits_{-T}^{T}e^{it(u+x)}dtdu = \int\limits_{-1}^{1}\frac{e^{iT(u+x)}-e^{-iT(u+x)}}{i(u+x)}du = \int\limits_{-1}^{1}\frac{2\sin(T(u+x))}{u+x}du \underset{v:=T(u+x)}{=} 2\int\limits_{T(x-1)}^{T(x+1)}\frac{\sin v}{v}dv$$

$$\int\limits_{T(x-1)}^{T(x+1)}\frac{\sin v}{v}dv \underset{T\to\infty}{\to} = \begin{cases} 0, \text{ если} & x > 1 (\text{оба конца интегрирования уходят на} +\infty \text{ сходящегося интегра} \\ \pi, \text{ если} & -1 < x < 1 (\text{концы разных знаков, а } \int_{-\infty}^{+\infty} \text{ считали где-то раньше}) \\ 0, \text{ если} & x < -1 (\text{оба конца интегрирования уходят на} -\infty \text{ сходящегося интегр} \\ \frac{\pi}{2} & x = \pm 1 \end{cases}$$

$$\lim_{T\to+\infty}\int\limits_{-T}^{T}\frac{e^{it}-e^{-it}}{it}\varphi_\xi(t)=\lim_{T\to\infty}\int\limits_{-\infty}^{+\infty}\Phi_T(x)dP_\xi(x)\underset{\text{см. ниже}}{=}\int\limits_{-\infty}^{+\infty}\lim_{T\to+\infty}\Phi_T(x)dP_\xi(x)=$$

$$=2\pi\int\limits_{-\infty}^{+\infty}\mathbb{1}_{[-1,1]}(x)dP_\xi(x)=2\pi P_\xi([-1,1])$$

Переставить интеграл с пределом можно по теореме Лебега. Суммируемая мажоранта $=$ 2π □

Следствие. Если $\varphi_\xi=\varphi_\eta$, то ξ и η одинаково распределены.

Доказательство. Надо проверить, что совпали функции распределения $F_\xi(x)$ и $F_\eta(x)$.

1) b такое, что $P(\xi=b)=0$

Проблемных точек (в которых $P(\xi=a)>0$) не более, чем счётное число.

Почему? Для любого n количество проблемных точек, значение в которых не меньше $\frac{1}{n}$, не превышает n. Получили счётное число конечных множеств, в которых лежат все такие точки.

Выберем последовательность $a_n<b$, $a_n\searrow-\infty$ и $P(\xi=a_n)=P(\eta=a_n)=0$

$P(a_n\leqslant\xi\leqslant b)=\lim\dots=P(a_n\leqslant\eta\leqslant b)\to P(\eta\leqslant b)=F_\eta(b)$

$P(a_n\leqslant\xi\leqslant b)\to P(\xi\leqslant b)=F_\xi(b)$

$\Rightarrow F_\xi=F_\eta$ во всех точках, где $P(\xi=b)=P(\eta=b)=0$

Но они непрерывны справа. Введём $b_n\searrow x$, $P(\xi=b_n)=P(\eta=b_n)=0$

$F_\xi(x)\leftarrow F_\xi(b_n)=F_\eta(b_n)\to F_\eta(x)$

Получили равенство во всех точках.

□

Следствие. Пусть $\int\limits_{-\infty}^{+\infty}|\varphi_\xi(t)|dt<+\infty$.

Тогда ξ имеет плотность $p_\xi(x)=\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}e^{-itx}\varphi_\xi(t)dt$.

Доказательство. $P(a\leqslant\xi\leqslant b)=\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}\frac{e^{-iat}-e^{-iat}}{it}\varphi_\xi(t)dt$ т.к. φ суммируемая и интеграл из теоремы выше сходится от $-\infty$ до $+\infty$ как произведение суммируемой функции на ограниченную.

Проверим, что формула плотности из условия теоремы подходит:

$$\int\limits_a^b p_\xi(x)dx=\frac{1}{2\pi}\int\limits_a^b\int\limits_{-\infty}^{+\infty}e^{-itx}\varphi_\xi(t)dtdx=\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}\int\limits_a^b e^{-itx}dx\varphi_\xi(t)dt\underset{\text{т.Фубини}}{=}\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}\frac{e^{-iat}-e^{-iat}}{it}\varphi_\xi(t)dt$$

□

Теорема 4.5. $\xi_k\sim\mathcal{N}(a_k,\sigma_k^2)$ – независимые с.в.

$\eta = a_0 + \sum\limits_{k=1}^{n} c_k\xi_k$, причём не все $c_k = 0$

$A \coloneqq a_0 + \sum\limits_{k=1}^{n} c_k a_k$

$\sigma \coloneqq \sum\limits_{k=1}^{n} c_k^2\sigma_k^2$

Тогда $\eta \sim \mathcal{N}(A, \sigma)$

Доказательство. $\varphi_{\xi_k}(t) = e^{ia_kt}e^{-\sigma_k^2t^2/2}$

$$\varphi_\eta(t) = \varphi_{a_0}(t)\prod_{k=1}^{n}\varphi_{c_k\xi_k}(t) = e^{ia_0t}\prod_{k=1}^{n}\varphi(c_kt) = e^{ia_0t}\prod_{k=1}^{n}e^{ia_kc_kt}e^{-\sigma_k^2c_k^2t^2/2} = e^{iAt-\sigma^2t^2/2}$$

□

Замечание. $\varphi_\xi(t) = \int_{-\infty}^{+\infty} e^{itx}p_\xi(x)\,dx$

$p_\xi(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx}\varphi_\xi(t)\,dt$, если φ_ξ суммируема.

Эти действия называются преобразование Фурье и обратное преобразование Фурье.

Теорема 4.6. $\xi_1, \xi_2, \ldots$ – с.в.

Следующие условия равносильны:

1. $\xi_n \to \xi$ по распредлению (т.е. $F_{\xi_n}(x) \to F_\xi(x)$ во всех точках непрерывности F_ξ)
2. $\forall U \subset \mathbb{R}$ – открытое множества.

   $\varliminf\limits_{n\to\infty} P(\xi_n \in U) \geqslant P(\xi \in U)$
3. $\forall B \subset \mathbb{R}$ – замкнутого множества

   $\varlimsup\limits_{n\to\infty} P(\xi_n \in B) \leqslant P(\xi \in B)$
4. $\forall A \subset \mathbb{R}$ борелевское : $P(\xi \in \operatorname{Fr} A) = 0$ $(\operatorname{Fr} \coloneqq \operatorname{Cl} A \setminus \operatorname{Int} A)$

   $\lim\limits_{n\to\infty} P(\xi_n \in A) = P(\xi \in A)$
5. $\forall A \subset \mathbb{R}$ борелевское : $P(\xi \in \operatorname{Fr} A) = 0$

   $E(\mathbb{1}_A \circ \xi_n) \to E(\mathbb{1}_A \circ \xi)$
6. $\forall f : \mathbb{R} \to \mathbb{R}$ непрерывная и ограниченная

   $E(f(\xi_n)) \to E(f(\xi))$
7. $\varphi_{\xi_n} \to \varphi_\xi$ поточечно

Доказательство.

2) $\iff$ 3) $B = \mathbb{R} \backslash U$

$P(\xi_n \in B) = 1 - P(\xi_n \in U)$

$\overline{\lim}\, P(\xi_n \in B) = 1 - \underline{\lim}\, P(\xi_n \in U) \underset{2 \Rightarrow 3}{\leqslant} 1 - P(\xi \in U) = P(\xi \in B)$

$\underline{\lim}\, P(\xi_n \in U) = 1 - \overline{\lim}\, P(\xi_n \in B) \underset{3 \Rightarrow 2}{\geqslant} 1 - P(\xi \in B) = P(\xi \in U)$

$2 \cup 3 \Rightarrow 4$ $B := \operatorname{Cl} A$, $U := \operatorname{Int} A$

$\underline{\lim}\, P(\xi_n \in A) \underset{\text{т.к. } U \subset A}{\geqslant} \underline{\lim}\, P(\xi_n \in U) \underset{\text{по п. 3}}{\geqslant} P(\xi \in U)$

$\underline{\lim}\, P(\xi_n \in A) \leqslant \overline{\lim}\, P(\xi_n \in A) \underset{\text{т.к. } A \subset B}{\leqslant} \overline{\lim}\, P(\xi_n \in B) \underset{\text{по п. 2}}{\leqslant} P(\xi \in B)$

$P(\xi \in B) = P(\xi \in U) + P(\xi \in \operatorname{Fr} A) \underset{P(\xi \in \operatorname{Fr} A) = 0}{=} P(\xi \in U)$

Показали, что $\underline{\lim}\, P(\xi_n \in A)$ с двух сторон зажимается $P(\xi \in U)$. Значит, $\underline{\lim}\, P(\xi_n \in A) = P(\xi \in U) = P(\xi \in A)$.

$4 \Rightarrow 5$

$$\mathbb{1}_A \circ \xi_n = \begin{cases} 1, & \xi_n \in A \\ 0, & \xi_n \notin A \end{cases}$$

Тогда $E(\mathbb{1}_A \circ \xi_n) = P(\xi_n \in A)$ (складывать вероятности с множителем 1 не очень интересно), а по пунтку 4 получаем $P(\xi_n \in A) \to P(\xi \in A)$

$6 \Rightarrow 7$ $\varphi_{\xi_n}(t) = Ee^{it\xi} = E\cos(t\xi_n) + iE\xi_n(t\xi_n)$

Ну и по пункту 7 рассмотрим матожидания по отдельности т.к. cos и sin вполне непрерывны и ограничены...

$1 \Rightarrow 2$ U – открытое множество, D –точки, где F_ξ не является непрерывной, D – нбсч

Представим U следующим образом: $U = \bigsqcup_{k=1}^{\infty} (a_k, b_k]$, $a_k, b_k \notin D$.

Покажем, что добиться того, чтобы концы не входили в D несложно. Вспомним, как мы показывали, что есть замощение каким-то счётным числом ячеек. Мы замощали обрезками целой длины, потом половинной, пока помещается, потом четвертью.. Сделаем процесс немного аккуратнее. Если мы в какой-то момент концом отрезка попали в D, сдвинемся этим концом на $< 10\%$ от длины отрезка так, чтобы не попасть (т.к. D нбчс, то найдём). Поскольку от этого процесса нам требовалось только монотонное уменьшение длины отрезков, которыми мы всё замощаем, то всё по-прежнему будет хорошо.

$P(\xi_n \in U) = \sum_{k=1}^{\infty} P(\xi_n \in (a_k, b_k]) = \sum_{k=1}^{\infty} (F_{\xi_n}(b_k) - F_{\xi_n}(a_k)) \geqslant$

$\geqslant \sum_{k=1}^{m} (F_{\xi_n}(b_k) - F_{\xi_n}(a_k)) \underset{n \to \infty, \text{по п. 1}}{\to} \sum_{k=1}^{m} (F_\xi(b_k) - F_\xi(a_k)) = P(\xi \in \bigcup_{k=1}^{m} (a_k, b_k)$

Когда мы делали переход по пункту 1, мы как раз воспользовались тем, что точки a_k, b_k хорошие.

$\underline{\lim} P(\xi_n \in U) \geqslant \underline{\lim} \sum_{k=1}^m (F_\xi(b_k) - F_\xi(a_k)) \underset{\text{т.к. эта штука имеет предел}}{=} P(\xi \in \bigcup_{k=1}^m (a_k, b_k]) =$

$= P_\xi(\bigcup_{k=1}^m (a_k, b_k]) \underset{m\to\infty}{\to} P_\xi(\bigcup_{k=1}^\infty (a_k, b_k]) = P_\xi(U) = P(\xi \in U)$

$5 \Rightarrow 6$ 6 верна для $f = \mathbb{1}_A$. Тогда верно и для ступенчатых функций.

$f \in C(\mathbb{R}), |f| \leqslant M$, $D' := \{x : P(f(\xi) = x) > 0\}$, т.е. $D' = f^{-1}(D)$ – нбчс.

Порежем отрезок $[-M, M]$ точками не из D': $-M = t_0 < t_1 < \ldots < t_n = M, t_i \notin D'$

Пусть $A_j := \{x : t_{j-1} \leqslant f(x) < t_j\} \supset \{x : t_{j-1} < x < t_j\} =: U_j$ – открытое

$A_j := \{x : t_{j-1} \leqslant f(x) < t_j\} \supset \{x : t_{j-1} \leqslant x \leqslant t_j\} =: B_j$ – дополнение открытого (которое из точек либо левее, либо правее), а значит замкнутое.

$B_j \backslash U_j \subset \{x : f(x) = t_{j-1}\} \sup\{x : f(x) = t_j\}$.

$P(\xi \in (B_j \backslash U_j)) = 0$ по выбору точек t_i (они не из D')

Тогда $P(\xi \in \operatorname{Fr} A_j) = 0$ и A_j можно подставлять в 5 пункт.

Возьмём ступенчатую функцию, похожую на f:

$g(x) := \sum_{j=1}^n t_j \mathbb{1}_{A_j}(x)$, $Eg(\xi_n) \to E(g(\xi))$

$Eg(\xi_n) \to Eg(\xi)$ по линейности и пункту 5.

$|f(x) - g(x)| \leqslant \max(t_j - t_{j-1})$ по разбиению на отрезочки.

$E|f(x) - f(g)| \leqslant \max\limits_{j=0,\ldots,n} (t_j - t_{j-1})$

Чего теперь хочется? Хочется для $\forall \varepsilon$ уметь предъявлять N, такое, что для любого $n > N$ существует подходящее разбиение разбиение отрезка $-M = t_0 < t_1 < \ldots < t_n = M, t_i \notin D'$ такое, что $\max\limits_{j=0,\ldots,n} (t_j - t_{j-1}) < \varepsilon$. На самом деле это можно делать также, как и в предыдущем доказательстве: сначала пытаемся поставить отметку на расстоянии ε от предыдущей, если не получилось, поставим где-то на отрезке $[t_{j-1} + 0.9\varepsilon, t_{j-1} + \varepsilon]$. Здесь важно, что D' нбчс. $\frac{2M}{0.9\varepsilon} \geqslant N \geqslant \frac{2M}{\varepsilon}$.

$|Ef(\xi_n) - Ef(\xi)| \leqslant |Ef(\xi_n) - Eg(\xi_n)| + |Eg(\xi_n) - Eg(\xi)| + |Eg(\xi) - Ef(\xi)| < 2\varepsilon + |Eg(\xi_n) - Eg(\xi)|$

Это мы подобрали такое g (точнее, n), что $|g(x) - f(x)| \leqslant \varepsilon$. А ещё мы знаем, что $Eg(\xi_n) \to Eg(\xi)$, значит с какого-то n это тоже будет $< \varepsilon$. Получили, что $Ef(\xi_n) \to Ef(\xi)$.

$7 \Rightarrow 1$ $P(\xi_n \in (a, b]) = \lim\limits_{T\to\infty} \frac{1}{2\pi} \frac{e^{-ait} - e^{-ibt}}{it} \varphi_{\xi_n}(t) dt$, если $P(\xi_n = a) = P(\xi_n = b) = 0$

Возьмём с.в. $\eta_\sigma \sim \mathcal{N}(0, \sigma^2)$ не зависит от $\xi_1, \xi_2, \ldots$.

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

$\varphi_{\xi+\eta_\sigma}(t) = \varphi_{\xi}(t)\varphi_{\eta_\sigma}(t) = \varphi_{\xi}(t)e^{-\sigma^2t^2/2}$

$\Rightarrow \varphi_{\xi_n+\eta_\sigma} \to \varphi_{\xi+\eta_\sigma}$ поточечно.

$P(\xi_n+\eta_\sigma \in (a,b]) = \lim\limits_{T\to\infty} \frac{1}{2\pi}\int_{-T}^{T} \frac{e^{-iat}-e^{-iab}}{it}\varphi_{\xi_n}(t)e^{-\sigma^2t^2/2}dt$

Заметим, что получившийся интеграл сходится и даже абсолютно т.к. $e^{-\sigma^2t^2/2}$ сходится, а всё остальное ограничено некоторой универсальной константой, которая даже от n не зависит: $|\varphi| \leqslant 1$, а $\frac{e^{-iat}-e^{-ibt}}{it}$ около нуля как-то ограничим, а на бесконечностях всё совсем хорошо.

$P(\xi_n+\eta_\sigma \in [a,b)) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{e^{-iat}-e^{-iab}}{it}\varphi_{\xi_n}(t)e^{-\sigma^2t^2/2}dt$

Воспользуемся тогда теоремой Лебега: пусть c – константа, ограничивающая$\frac{e^{-iat}-e^{-ibt}}{it}$, тогда мажорантой будет $ce^{-\sigma^2t^2/2}$. Тогда

$\frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{e^{-iat}-e^{-iab}}{it}\varphi_{\xi_n}(t)e^{-\sigma^2t^2/2}dt \underset{\text{т. Лебега}}{\to} \frac{1}{2\pi}\frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{e^{-iat}-e^{-iab}}{it}\varphi_{\xi}(t)e^{-\sigma^2t^2/2}dt = P(\xi+\eta_\sigma \in [a,b))$

Таким образом мы всё знаем про $\xi+\eta_\sigma$ в хороших отрезках (концы которых не лежат в точках разрыва функции распределения).

Осталось две вещи: понять, почему можно добавить η_σ и почему плохие точки нам не мешают.

$G_n := F_{\xi+\eta_\sigma}$, $G := F_\xi + \eta_\sigma$

Мы поняли, что $G_n(x) \to G(x)$ за исключением нбчс множества.

Пусть x – точка непрерывности $F_\xi \Rightarrow \exists\delta > 0 : x+\delta$ и $x-\delta$ хорошие точки и $F_\xi(x)-F_\xi(x\pm\delta) < \varepsilon$

$P(|\eta_0| \geqslant \varepsilon) \underset{\text{нер-во Чебышёва} \frac{D\eta_0}{\varepsilon^2}}{\leqslant} = \frac{\sigma^2}{\delta^2}$. Тогда мы можем подобрать такую σ, для которой $\frac{\sigma^2}{\delta^2} < \varepsilon$.

TODO: Мне надоело писать это нудное доказательство, потом допишу. Всё равно оно пока никому не нужно. □

5. Центральная предельная теорема

Теорема 5.1 (ЦПТ в форме Леви)**.** $\xi_1, \xi_2, \ldots$ – независимые одинаково распределенные величины.

$a = E_{\xi_1}$, $\sigma^2 = D_{\xi_1}$, $S_n := \xi_1 + \xi_2 + \ldots + \xi_n$

Тогда $F_{\frac{S_n - na}{\sqrt{n}\sigma}} \to \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$ во всех точках.

Доказательство. Мы хотим проверить, что последовательность функций распределения сходится во всех точках к некоторой непрерывной функции распределения. Значит, это просто сходимость по распределению. Значит, нам надо проверить, что $\frac{S_n - an}{\sqrt{a}\sigma}$ сходится к $\mathcal{N}(0,1)$ по распределению.

$\varphi_{\frac{S_n - an}{\sqrt{a}\sigma}} = \Pi \varphi_{\frac{\xi_k - a}{\sqrt{n}\sigma}} = (\varphi_{\frac{\xi_1 - a}{\sqrt{n}\sigma}})^n$ т.к. с.в. одинаково распределены.

Мы знаем, что у величины $\frac{\xi_k - a}{\sqrt{n}\sigma}$ нулевое матожидание т.к. $E(\xi_1 - a) = 0$

$D(\xi_1 - a) = D_{\xi_1} = \sigma^2$

Где-то выше мы узнали, что $E\xi = -i\varphi'_\xi(0)$ и $D\xi = -\varphi''_\xi(0) + (\varphi'_\xi(0))^2$. Воспользуемся этими знаниями, чтобы разложить φ в ряд в нуле до второго члена:

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

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

Нам в результате хочется показать, что $\varphi_{\frac{S_n - an}{\sqrt{a}\sigma}} \to e^{-t^2/2}$

То есть $(1 - \frac{t^2}{2n} + o(\frac{t^2}{n\sigma^2}))^n \to e^{-\frac{t^2}{2}}$

Прологарифмируем

$n \ln(1 - \frac{t^2}{2n} + o(\frac{t^2}{n})) \sim n(-\frac{t^2}{2n} + o(\frac{t^2}{n})) \to -\frac{t^2}{2}$ □

18.04.2018

Теорема 5.2. $F_n, F : \mathbb{R} \to [0,1]$ монотонны и $F \in C(\mathbb{R}) \Rightarrow F_n \rightrightarrows F$

Доказательство. Возьмём $\varepsilon = \frac{1}{m}$

Пусть $t_k : F(t_k) = \frac{k}{m} \quad k = 1, 2, \ldots, m-1$ т.е. просто порезали на кусочки с одинаковым приростом функции. Заметим, что так можно ровно потому, что $F \in C(\mathbb{R})$. Тогда т.к. F ещё и монотонна, у каждой точки на отрезке $[0,1]$ существует прообраз в виде точки или непрерывного отрезка. (TODO: на лекции уже после доказательства мы почему-то решили, что непрерывность F не нужна. Тем не менее, я без неё действие выше делать не умею)

TODO: картинка монотонной функции

Мы знаем про то, что в этих точках есть сходимость: $F_n(t_k) \to F(t_k)$

Выберем N, начиная с которого $|F_n(t_k) - F(t_k)| < \varepsilon \quad \forall k = 1, 2, \dots, m-1$. Т.е. функции в этих точках не сильно отличаются от F.

Покажем, что и в остальных точках разница небольшая. Рассмотрим t, $t_k \leqslant t < t_{k+1}$

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

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

$\Rightarrow |F_n(t) - F(t)| < 2\varepsilon$

□

Замечание. Значит, в центральной предельной т. равномерная сходимость.

Теорема 5.3 (Муавра-Лапласа). $p \in (0,1)$ схема Бернулли с вероятностью успеха p

$p(\frac{S_n - np}{\sqrt{npq}} \leqslant x) \to \Phi(x)$

$p(a < \frac{S_n - np}{\sqrt{npq}} \leqslant b) \to \Phi(b) - \Phi(a)$ (просто разность предыдущих выражений)

Доказательство. $S_n = \xi_1 + \xi_2 + \dots + \xi_n$

$E\xi_1 = p$, $D\xi_1 = pq$

По ЦПТ, $P\left(\frac{S_n - nE\xi_1}{\sqrt{nD\xi_1}} \leqslant x\right) \to \Phi(x)$ □

Теорема 5.4 (Пуассона). $\xi_{11}, \xi_{21}, \xi_{22}, \xi_{31}, \xi_{32}, \xi_{33}, \dots$

$P(\xi_{nk} = 1) = p_{nk}$, $P(\xi_{nk} = 0) = q_{nk} = 1 - p_{nk}$

Испытания в каждой серии независимы: $\xi_{n1}, \xi_{n2}, \dots, \xi_{nn}$ – независимы

$\max\{p_{n1}, p_{n2}, p_{nn}\} \to 0$, $p_{n1} + p_{n2} + \dots + p_{nn} \to \lambda > 0$

$S_n \coloneqq \xi_{n1} + \xi_{n2} + \dots + \xi_{nn}$

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

Доказательство. Заметим, что при распределении Бернулли считать характеристическую функцию тоже очень просто:

$\varphi_{\xi_{nk}}(t) = Ee^{it\xi_{nk}} = e^{it\cdot 1} p_{nk} + e^{it\cdot 0} q_{nk} = e^{it} p_{nk} + q_{nk} = 1 + p_{nk}(e^{it} - 1)$

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

$\ln(\prod\limits_{k=1}^{n} (1 + p_{nk}(e^{it} - 1))) = \sum\limits_{k=1}^{n} \ln(1 + p_{nk}(e^{it} - 1)) \underset{\text{р. Тейлора}}{=} \sum\limits_{k=1}^{n} p_{nk}(e^{it} - 1) + \sum\limits_{k=1}^{n} \mathcal{O}(p_{nk}^2)$

$\sum\limits_{k=1}^{n} p_{nk}(e^{it}-1) \to \lambda(e^{it}-1)$ т.к. $(e^{it}-1)$ – фиксировано, а $\sum p \to \lambda$

$\sum\limits_{k=1}^{n} \mathcal{O}(p_{nk}^2) \leqslant C \sum p_{nk} \max p_{nk}$, где C – константа из $\mathcal{O}$, а всё остальное, вместо того, чтобы считать сумму квадратов, мы умножили на максимум из вероятностей для n-й серии экспериментов. $\max p_{nk} \to 0$ по условию.

Итого, получили, что $\phi_{S_n} \to e^{\lambda(e^{it}-1)}$, поймём, что это х.ф. распределения Пуассона

Заведём $\eta \sim Poisson(\lambda)$

$\phi_\eta(t) = Ee^{it\eta} = \sum\limits_{m=0}^{\infty} e^{itm}P(\eta=m) = \sum\limits_{m=0}^{\infty} e^{itm}e^{-\lambda}\frac{\lambda^m}{m!} = e^{-\lambda}\sum\limits_{m=0}^{\infty}\frac{(e^{it}\lambda)^m}{m!} \underset{\text{р. Тейлора для } e^x}{=} e^{\lambda e^{it}-\lambda} \Rightarrow$ есть сходимость по распределению

Осталось понять, что есть сходимость во всех точках. Из сходимости по распределению мы получили сходимость во всех точках, где есть непрерывность. Но наше распределение есть по сути ступеньки, поэтому непрерывности на разрыве ступенек нет. Возьмём такую точку разрыва $P(S_n = m)$. Что есть это значение? По сути, это $F(m+\varepsilon) - F(m-\varepsilon)$ при $\varepsilon \to 0$. А в точках $m+\varepsilon$ и $m-\varepsilon$ есть сходимость к тому, что нужно (т.к. в них непрерывно). В пределе получится то, что хотели. (TODO: формально осознать это рукомахательство). $\square$

Теорема 5.5 (ЦПТ в форме Линденберга)**.** ξ_1, ξ_2, ...– независимы с.в.

$a_n := E\xi_n$, $\sigma_n^2 := D\xi_n$, $D_n^2 := \sum\limits_{k=1}^{n}\sigma_k^2$

Если $Lind(\varepsilon, n) := \frac{1}{D_n^2}\sum\limits_{k=1}^{n} Ef(\xi_k - a_k) \underset{n\to\infty}{\to} 0$ (условие Линденберга)

и $f(x) := x^2\mathbb{1}_{\{|x|\geqslant\varepsilon D_n\}}$, $\varepsilon > 0$

Тогда $P(\frac{S_n - ES_n}{\sqrt{DS_n}} \leqslant x) \to \Phi(x)$

Доказательство. Не будем доказывать, это делается аналогично ЦПТ. $\square$

Теорема 5.6 (ЦПТ в форме Ляпунова)**.** ξ_1, ξ_2, ...– независимые с.в.

$a := E\xi_n$, $\sigma_n^2 := D\xi_n$, $D_n := \sqrt{\sum\limits_{k=1}^{n}\sigma_k^2}$

Если $\delta > 0$, $L(\delta, n) := \frac{1}{D_n^{2+\delta}}\sum\limits_{k=1}^{n} E|\xi_k - a_k|^{2+\delta} \underset{n\to\infty}{\to} 0$

Тогда $P(\frac{S_n - ES_n}{\sqrt{DS_n}} \leqslant x) \to \Phi(x)$

Доказательство. Линденберг $\Rightarrow$ Ляпунов

$Lind(\varepsilon, n) = \frac{1}{D_n^2}\sum\limits_{k=1}^{n} E((\xi_k - a_k)^2\mathbb{1}_{\{|\xi_k - a_k|\geqslant\varepsilon D_n\}}) \leqslant$

Т.к. из-за характеристической функции на имеющей значение части верно, что $|\xi_k - a_k| \geqslant \varepsilon D_n$, то можно домножить функцию под матожиданием на $\left(\frac{|\xi_k - a_k|}{\varepsilon D_n}\right)^\delta$, $\delta > 0$, станет только больше.

$$\leqslant \frac{1}{D_n^2}\sum_{k=1}^{n} E\left((\xi_k-a_k)^2\frac{|\xi_k-a_k|^\delta}{\varepsilon^\delta D_n^\delta}\mathbb{1}_{\{}|\xi_k-a_k|\geqslant\varepsilon D_n\}\right)\leqslant$$

Если убрать характеристическую функцию, то т.к. под матожиданием нечто неотрицательное, то множество увеличится, значит и значение не уменьшится.

$$\leqslant \frac{1}{\varepsilon^\delta D^{2+\delta}n}\sum_{k=1}^{n} E|\xi_k-a_k|^{2+\delta}=\frac{L(\delta,n)}{\varepsilon^\delta}\underset{\text{по условию}}{\to} 0 \qquad \square$$

Упражнение. Проверить, что для независимых одинаково распределённых с.в. выполняется условие Линденберга

Теорема 5.7. ξ_1, ξ_2, ...– независимые с.в. , $\delta\in(0,1]$

Тогда $\sup_{x\in\mathbb{R}}|P(\frac{S_n-ES_n}{\sqrt{DS_n}}\leqslant x)-\Phi(x)|\leqslant C_\delta L(\delta,n)$

Следствие Теорема Берри-Эссеена. ξ_1, ξ_2, ...– независимые одинаково распределённые с.в.

$a=E\xi_1$, $\sigma^2=D\xi_1$

$\sup_{x\in\mathbb{R}}|P\left(\frac{S_n-na}{\sqrt{n}\sigma}\leqslant x\right)-\Phi(x)|\leqslant C\frac{E|\xi_1-a|^3}{\sqrt{n}\sigma^3}$

Доказательство. Т.к. распределенности одинаковые, то

$\sum_{k=1}^n E(|\xi_k-a_k|^{2+\delta})=nE(\xi-a)^{2+\delta}$, а $D_n^2=n\sigma^2$

$$\sup_{x\in\mathbb{R}}|P\left(\frac{S_n-na}{\sqrt{n}\sigma}\leqslant x\right)-\Phi(x)|\leqslant C_\delta L(\delta,n)\underset{\text{см. стрчку выше}}{=}C_\delta\frac{nE|\xi_1-a|^{2+\delta}}{\sqrt{n}\sigma^2 D_n^\delta}\underset{\delta:=1}{=}C\frac{nE|\xi_1-a|^3}{n\sqrt{n}\sigma^3}=C\frac{E|\xi_1-a|^3}{\sqrt{n}\sigma^3}$$

$\square$

Замечание. Эссеен (1956) $C>0.40973$, Швецова (2011) $C<0.4748$

Пример. $\xi_k=\begin{cases}k^\alpha \text{ с вероятностью } 0.5\\ -k^\alpha \text{ с вероятностью } 0.5\end{cases}$ – независимые случайные величины

$E\xi_k=0$ (т.к. симмитричная с.в.), $D\xi_k=E_{\xi_k^2}-(E_{\xi_k})^2=E_{\xi_k^2}=\frac{1}{2}(k^\alpha)^2+\frac{1}{2}(-k^\alpha)^2=k^{2\alpha}$

$D_n^2=\sum_{k=1}^{n}k^{2\alpha}\sim\frac{n^{2\alpha+1}}{2\alpha+1}$ при $\alpha>-0.5$ (по Эйлеру-Маклорену)

$E|\xi_k|^3=k^{3\alpha}$

$$\sum_{k=1}^{n}E|\xi_k|^3=\sum_{k=1}^{n}k^{3\alpha}\sim\begin{cases}\frac{n^{3\alpha+1}}{3\alpha+1} \text{ при } \alpha>-\frac{1}{3} \text{ (тот же Эйлер-Маклорен)}\\ \ln n \text{ при } \alpha=-\frac{1}{3} \text{ (гармонический ряд)}\\ \mathcal{O}(1) \text{ при } \alpha<-\frac{1}{3} \text{ (сходящаяся геом. последовательность)}\end{cases}$$

$$L(\delta:=1,n)=\frac{1}{D_n^3}\sum E|\xi_k|^3\sim\frac{(2\alpha+1)^{3/2}}{n^{3\alpha+3/2}}\cdot\begin{cases}\frac{n^{3\alpha+1}}{3\alpha+1}\\ \ln n\\ \mathcal{O}(1)\end{cases}\to 0 \text{ при } \alpha>-0.5$$

Мы проверили условие Ляпунова и из неё делаем вывод:

$\frac{S_n - ES_n}{\sqrt{DS_n}} \sim S_n \frac{(2\alpha+1)^{1/2}}{n^{\alpha+1/2}} \to$ сходится к нормальному распределению

Теорема 5.8 (Хартмана-Винтера или Закон повторного логарифма)**.** ξ_1, ξ_2, ...– независимые одинаково распределённые с.в.

$E\xi_1 = 0$, $\sigma^2 := D\xi_1$

Тогда $\overline{\lim\limits_{n\to\infty}} \frac{S_n}{\sqrt{2n\ln\ln n}} = \sigma$

Замечание. В чём смысл этой теоремы?

При $E\xi_1 = 0$ ЦПТ говорит нам следующее: $\frac{S_n}{\sqrt{n}\sigma} \to \mathcal{N}(0,1)$. То есть что порядок роста отклонения – это примерно $\sqrt{n}\sigma$.

Закон же повторного логарифма говорит нам, что на самом деле порядок отклонения – это $\sqrt{2n\ln\ln n}\sigma$ и точек, отклонение для которых имеет порядок повторного логарифма достаточно много (бесконечно много, раз уж верхний предел вышел). А вот отклонение $\sqrt{n}$ на самом деле в жизни достаточно редкое. Бонусная картинка для медитации

Теорема 5.9 (Штрассена)**.** При тех же условиях множество предельных точек последовательности $\frac{S_n}{\sqrt{2n\ln\ln n}}$ есть $[-\sigma, \sigma]$

6. Случайные процессы

6.1. Определения

***Определение* 6.1** (Случайный процесс)**.** $(\Omega, \mathcal{F}, P)$ – вероятностное пространство, T – некоторое множество. Случайны процесс – набор случайных величин $\{\xi_t\}_{t\in T}$ из $(\Omega, \mathcal{F}, P)$.

***Определение* 6.2** (Случайный процесс с непрерывным временем)**.** Случайный процесс, для которого ξ_t измерима в произвольные моменты t

***Определение* 6.3** (Случайный процесс с дискретным временем)**.** Есть моменты t_1, t_2, ..., когда изменяется ξ_t.

Замечание. Имеется ввиду изменение именно с.в., а не её значения. Например – несколько секунд распределение было нормальным, а потом стало равномерным.

***Определение* 6.4** (Случайный процесс с дискретным набором состояний)**.** Случайный процесс, у которого множество значений ξ_t дискретно.

Пример. T – некоторый промежуток времени.

Пример с непрерывным временем: ξ_t – температура в момент времени t.

Пример с дискретным временем и дискретным набором состояний: пользователь прыгает по сайтам. ξ_t – страница, которая открыта у него в момент времени t.

***Определение* 6.5** (Траектория)**.** $f : \Omega \times T \to \mathbb{R}$

$\xi_t(w) \coloneqq f(w, t)$ – случайные величины

Фиксир. $w = w_0$ рассм. $f(w_0, t)$ – траектория (реализация)

Пример.

1. g – непрерывная функция : $\mathbb{R} \to \mathbb{R}$

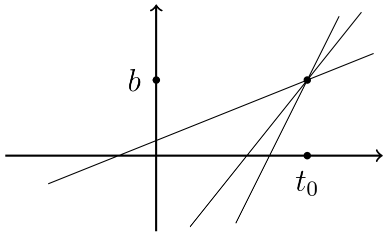
   ξ – случайная величина

   $\xi_t(w) \coloneqq g(t)\xi(w)$

2. Веерный процесс

   $\xi_t(w) \coloneqq (t - t_0)\xi(w) + b.$

Картинка траектории для такого процесса будет выглядеть как прямая, вращающаяся вокруг своей оси:



3. Последовательность случайных величин – случайный процесс с дискретным временем

6.2. Условные мат. ожидания

***Определение* 6.6.** $\xi : \Omega \to \mathbb{R}$ случайная величина на $(\Omega, \mathcal{F}, P)$

$\mathcal{A}$ – σ-алгебра – подмножество $\mathcal{F}$

Тогда $\eta := E(\xi|\mathcal{A})$ – с.в. на $(\Omega, \mathcal{A}, P)$ такая, что $\forall A \in \mathcal{A}$ $E(\eta \mathbb{1}_A) = E(\xi \mathbb{1}_A)$

Замечание Единственность. Пусть η_1, η_2 – условные матожидания

$E(\eta_1 \mathbb{1}_A) = E(\xi \mathbb{1}_A) = E(\eta_2 \mathbb{1}_A) \Rightarrow E((\eta_1 - \eta_2)\mathbb{1}_A) = 0$

$$\left.\begin{aligned} A := \{\eta_1 > \eta_2\} \Rightarrow P(A) = 0 \\ \tilde{A} := \{\eta_1 < \eta_2\} \Rightarrow P(\tilde{A}) = 0 \end{aligned}\right\} \Rightarrow \eta_1 = \eta_2 \text{ почти наверное}$$

Пример. $\mathcal{A} = \{\varnothing, \Omega\}$. Посчитаем на таком множестве $E(\xi|\mathcal{A})$.

Измеримые относительно этой $\mathcal{A}$ функции – константы и только они т.к. любое лебегово множество должно быть либо пустым, либо всем сразу.

Значит, $\eta := E(\xi|\mathcal{A}) = const =: c$. Что это за константа? Распишем $\forall A \in \mathcal{A}$:

$E(\eta \mathbb{1}_\varnothing) = E(c \mathbb{1}_\varnothing) = E(\xi \mathbb{1}_\varnothing)$ – всегда верно.

$c = E(c \mathbb{1}_\Omega) = E(\xi \mathbb{1}_\Omega) = E\xi$

Получили $\eta = E\xi$

18.04.2018

Пример. $\Omega := \bigsqcup\limits_{k=1}^{\infty} A_k$ натянем на $\{A_n\}$ σ-алгебру
η постоянна на A_k $\forall k$
$E(\xi \mathbb{1}_{A_k}) = E(\eta \mathbb{1}_{A_k}) = P(A_k) \cdot \eta|_{A_k}$

$\eta(w) = \frac{E(\xi \mathbb{1}_{A_c})}{P(A_k)}$ при $w \in A_k$

***Определение* 6.7.** μ и η – меры на σ-алгебре $\mathcal{B}$

μ абсолютно непрерывна относительно ν

$\mu << \nu$ если $\forall e \in \mathcal{B}$ $\nu e = 0 \Rightarrow \mu e = 0$

Теорема 6.1 (Радона-Никодима)**.** μ и ν конечные меры на $\mathcal{B}$ и $\mu << \nu$. Тогда $\exists w \geqslant 0$, суммируемая относительно ν, т.ч. $\mu A = \int_A w d\nu$

Теорема 6.2. Если $E|\xi| < +\infty$, то существует $E(\xi|A)$ и оно $< +\infty$ п.н.

Доказательство. $\xi_\pm$

$\mu_\pm A \coloneqq E(\xi_\pm \mathbb{1}_A)$ при $A \in \mathcal{A}$ – конечные меры на $\mathcal{A}$ и $\mu_\pm << P$

Тогда $\exists \eta_\pm \geqslant 0$ – измерима относительно $\mathcal{A}$, суммируема и т.ч. $E(\xi_\pm \mathbb{1}_A) = \mu_\pm A = \int_A \eta_\pm dP$

$\eta \coloneqq \eta_+ - \eta_-$ – измерима относительно $\mathcal{A}$

$E(\xi \mathbb{1}_A) = E(\xi_+ \mathbb{1}_A) - E(\xi_- \mathbb{1}_A) = \int_A (\eta_+ - \eta_-) dP = \int_A \eta dP = E(\eta \mathbb{1}_A)$

□

Свойства условного мат. ожидания.

1. $E(c|A) = c$
2. $E(\xi|A)$ линейно по ξ
3. $\mathcal{A}_1 \subset \mathcal{A}_2 \Rightarrow E(\xi|\mathcal{A}_1) = E(E(\xi|\mathcal{A}_2)|\mathcal{A}_1)$

   Доказательство. Надо доказать, что п.ч. есть $E(\xi|\mathcal{A}_1)$

   оно измеримо относительно $\mathcal{A}_1$

   $\nu_1 \coloneqq E(E(\xi|\mathcal{A}_2)|\mathcal{A}_1)$

   $\nu_2 \coloneqq E(\xi|\mathcal{A}_2$

   $E(\nu \mathbb{1}_A) = E(\tilde{\nu} \mathbb{1}_A) = E(\xi \mathbb{1}_A) \forall A \in \mathcal{A}_1$

   □

4. $E(E(\xi|\mathcal{A})) = E\xi$

   Доказательство. $E\xi = E(\xi|\{\varnothing, \Omega\})$ и далее по св-ву 3 для $\mathcal{A}_1 \coloneqq \{\varnothing, \Omega\}$

   □

5. Если ξ – измеримая относительно $\mathcal{A}$, то $E(\xi|\mathcal{A}) = \xi$

6. $\xi \leqslant \eta \Rightarrow E(\xi|\mathcal{A}) \leqslant E(\eta|\mathcal{A})$ п.н.

   Доказательство. $\xi \geqslant 0 \Rightarrow E(\xi|\mathcal{A}) \geqslant 0$ из доказательства существования □

***Определение* 6.8.** η – с.в.

$\sigma(\eta)$ – наименьшая σ-алгебра, относительно которой η измерима. Это σ-алгебра, натянутая на $\{\eta \leqslant t\}_{t\in\mathbb{R}}$

***Определение* 6.9.** $E(\xi|\eta) \coloneqq E(\xi|\sigma(\eta))$

Если η дискретная, то $\eta = \sum\limits_{k=1}^{\infty} c_k \mathbb{1}_{A_k}$, A_k дизъюнктны.

Пример. $\mathcal{A}$ натянута на $A_k : \Omega = \bigcup\limits_{k=1}^{\infty} A_k$

$E(\xi|\mathcal{A}) = \sum\limits_{k=1}^{\infty} a_k \mathbb{1}_{A_k}$

$a_k = \frac{E(\xi \mathbb{1}_{A_k})}{P(A_k)}$

$E(\sum\limits_{k=1}^{\infty} a_k \mathbb{1}_{A_k}) = \sum\limits_{k=1}^{\infty} a_k P(A_k)$

$E\xi = \sum\limits_{k=1}^{\infty} P(\eta = c_k) \cdot E(\xi|\eta = c_k)$

Свойства.

7. ξ и η независимы $\Rightarrow E(\xi|\eta) = E\xi$

Доказательство. Надо проверить равенства $E(\xi \mathbb{1}_A) = E(E\xi \mathbb{1}_A) = E\xi P(\eta \leqslant t)\ \forall A = \{\eta \leqslant t\}$

$E(\xi \mathbb{1}_{\{\eta \leqslant t\}}) = P(\eta \leqslant t) E\xi$

$P(\xi \geqslant s, \eta \leqslant t) = P(\xi \geqslant s) P(\eta \leqslant t)$

$E_{\xi_+} = \int_0^{\infty} P(\xi \geqslant s) ds$ □

Пример. $N, \xi_1, \xi_2, \ldots$ – независимые с.в.

$\xi_1, \xi_2, \ldots$ – одинаково распределены $E\xi_1 = a$

$E(\sum\limits_{k=1}^{N} \xi_k) = \sum\limits_{n=1}^{\infty} P(N = n) E(\sum\limits_{k=1}^{N} \xi_k | N = n) =$

$E(\sum\limits_{k=1}^{n} \xi_k) = \sum\limits_{k=1}^{n} E\xi_k = na = \sum\limits_{n=1}^{\infty} P(N = n) na = ENa$

Упражнение. $D(\sum\limits_{k=1}^{N} \xi_k) = ?$

Через $E\xi_1, D\xi_1, EN, DN$.

***Определение* 6.10** (Мартингал). $(\Omega, \mathcal{F}, P)$ – вероятностное пространство

$\mathcal{A}_0 \subset \mathcal{A}_1 \subset \ldots$ – σ-алгебры $\subset \mathcal{F}$

$\xi_0, \xi_1, \xi_2, \ldots$ – мартингал, если $E|\xi_n| < +\infty$

$\xi_n = E(\xi_{n+1}|\mathcal{A}_n)$

Пример.

1. Конечный мартингал

   $\xi_n \coloneqq E(\eta|\mathcal{A}_n)$

2. $\eta_1, \eta_2, \eta_3, \ldots$ – независимые с.в.

   $P(\eta_k = 1) = P(\eta_k = -1) = \frac{1}{2}$

   $S_n \coloneqq \eta_1 + \eta_2 + \eta_3 + \cdots + \eta_n$

   $\mathcal{A}_n \coloneqq \sigma(\eta_1, \eta_2, \ldots, \eta_n)$

   $E(S_{n+1}|\mathcal{A}_n) = E(S_n + \eta_{n+1}|\mathcal{A}_n) = E(S_n|\mathcal{A}_n) + E(\eta_{n+1}|\mathcal{A}_n)$

   $E(S_n|\mathcal{A}_n) = S_n$

   $E(\eta_{n+1}|\mathcal{A}_n) = E\eta_{n+1} = 0$

6.3. Марковские цепи

***Определение* 6.11.** Y – нбчс множество (множество состояний/фазовое пространство)

ξ_1, ξ_2, ...– последовательность с.в. : $\Omega \to Y$ – цепь Маркова, если $P(\xi_n = a_n|\xi_{n-1} = a_{n-1}, \ldots, \xi_0 = a_0) = P(\xi_n = a_n|\xi_{n-1} = a_{n-1})$

Замечание. "Будущее зависит от настоящего, но не зависит от прошлого"

Замечание. Распределения вероятности ξ_n зависят от распределений вероятности π_0, т.е. вероятности $P\xi_0$ функций перехода $P(\xi_n = b|\xi_{n-1} = a) = p_{ab}^{(n)}$

Будем рассматривать однородные цепи Маркова, где $p_{ab}^{(n)} = p_{ab}$ (не зависит от n)

Пример.

1. Случайные величины на $\mathbb{Z}$

   $Y = \mathbb{Z}$

   Прыгаем влево или вправо по прямой:

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

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

   $\xi_n \coloneqq \eta_1 + \eta_2 + \cdots + \eta_n$, где $P(\eta_k = 1) = p$, $P(\eta_k = -1) = 1-p$

2. Случайное блуждание с поглощением

   Прыгаем вправо или влево по отрезку, пока не дойдём до одного из концов. В конце стоим.

   $Y = \mathbb{Z} \cap [a, b]$

3. Случайное блуждание с отражением

   Прыгаем вправо или влево по отрезку. Дойдя до конца, с вероятностью 1 отскакиваем назад.

Какова вероятность фиксированной траектории?

Теорема 6.3. $P(\xi_0 = a_0, \xi_1 = a_1, \dots, \xi_n = a_n) = \pi_0(a_0)p_{a_0a_1}p_{a_1a_2}\dots p_{a_{n-1}a_n}$

Доказательство. Индукция.

Переход:

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

$P(\xi_n = a_n|\xi_{n-1} = a_{n-1}) = p_{a_{n-1}a_n}$ □

Почему существуют Марковские цепи?

Теорема 6.4. $\pi_0 : Y \to [0, 1]$

$p : Y \times Y \to [0, 1]$

$\sum\limits_{a \in Y} \pi_0(a) = 1$

$\sum\limits_{b \in Y} p_{ab} = 1 \ \forall a \in Y$

Тогда существует цепь Маркова с начальным распределением π_0 и переходной вероятностью p_{ab}.

Теперь в терминах бесконечных матриц.

Теорема 6.5. ρ – строка переходных вероятностей

$\pi_n := P_{\xi_n} = \pi\rho^n$

Доказательство. Индукция.

Переход $n - 1 \to n$:

$\pi_n = \pi_{n-1}\rho$

$\pi_n(b) = \sum\limits_{a \in Y} \pi_{n-1}(a)p_{ab} = \sum\limits_{a \in Y} P(\xi_n = a)P(\xi_n = b|\xi_{n-1} = a) \underset{\text{ф-ла полной в-ти}}{=} P(\xi_n = b)$ TODO: картинка перемножения строки из $\pi_{n-1}(y_0), \pi_{n-1}(y_1), \dots$ на матрицу из $p_{y_iy_j}$ □

Замечание Обозначения. $p_{ab}(b) := P(\xi_n = b|\xi_0 = a) = P(\xi_{n+k} = b|\xi_k = a)$

***Определение* 6.12.** π – стационарное распределение, если $\pi = \pi\rho$

Пример. Стационарное распределение есть не всегда.

Случайное блуждание на прямой с равным шансом пойти вправо и влево.

Пусть существует стационарное распределение π

$\pi(y) = \pi(y-1)\frac{1}{2} + \pi(y+1)\frac{1}{2}$

$2\pi(y) = \pi(y-1) + \pi(y+1)$

$\pi(y+1) - \pi(y) = \pi(y) - \pi(y-1) \Rightarrow$ первая разность $= const$

Так не бывает.