Московский физико-технический институт

факультет инноваций и высоких технологий

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Лектор: М.Е. Жуковский

конспект лекций

автор: Александр Марков

30 января 2017 г.

Благодарности:

- М.Е. Жуковскому за проверку конспекта.
- Всем неравнодушным сокурсникам, сообщавшим об опечатках и ошибках, и в особенности Васильеву Александру, нашедшему наиболее неочевидные неточности.
- Марии Носаревой за предоставленные рукописные конспекты многих лекций.

Комментарии, предложения и найденные ошибки приветствуются по адресу `https://vk.com/id119012868`.

Содержание

1 Аксиоматика теории вероятностей

1.1 Аксиоматика Колмогорова

Определение 1.1. Произвольное множество Ω (не обязательно конечное) называется *пространством элементарных исходов.*

Определение 1.2. $\mathscr{F}$ — система подмножеств Ω — называется *алгеброй* на Ω, если выполнены следующие свойства:

1. $\Omega \in \mathscr{F}$
2. $A \in \mathscr{F} \Rightarrow \overline{A} \in \mathscr{F}$.
3. $A_1,\ A_2 \in \mathscr{F} \Rightarrow A_1 \cup A_2$ и $A_1 \cap A_2 \in \mathscr{F}$. (Причем достаточно лишь одного из этого).

Определение 1.3. Алгебра $\mathscr{F}$ называется *σ-алгеброй (сигма-алгеброй)*, если $A_1, A_2, \ldots \in \mathscr{F} \Rightarrow \bigcup\limits_{i=1}^{\infty} A_i \in \mathscr{F}$.

Замечание: Не важно, взять ли в определении сигма-алгебры объединение или пересечение множеств, т.к. они выражаются через друг друга и дополнение множеств согласно законам де Моргана.

Докажем некоторые свойства алгебр и σ-алгебр.

Утверждение 1.1.1. $\varnothing \in \mathscr{F}$

Доказательство. $\overline{\Omega} = \varnothing \in \mathscr{F}$ □

Утверждение 1.1.2. $\forall A,\ B \in \mathscr{F}:\ A \setminus B \in \mathscr{F}$

Доказательство. $A \setminus B = A \cap \overline{B}$ □

Определение 1.4. Пусть $(\Omega,\ \mathscr{F})$ — пространство элементарных событий и алгебра на нем. Тогда функция $\mu : \mathscr{F} \to \mathbb{R}_+ \cup \{+\infty\}$ называется *конечно-аддитивной мерой*, если $\forall A_1,\ A_2 \in \mathscr{F},\ A_1 \cap A_2 = \varnothing$ то $\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2)$.

Если $\mathscr{F}$ — алгебра и $\forall A_1,\ A_2, \ldots \in \mathscr{F}, \quad \forall i \neq j \quad A_i \cap A_j = \varnothing$, $\bigcup A_i \in \mathscr{F} \Rightarrow \mu(\bigcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} \mu(A_i)$, то μ называется *счетно-аддитивной мерой.*

Определение 1.5. Если $\mu(\Omega) = 1$, то μ называется конечно-аддитивной верятностью (в случае если μ конечно-аддитивная мера) и просто *вероятностью*, если μ — счетно-аддитивная. В таком случае, будем обозначать μ символом P.

Утверждение 1.1.3. $\mathsf{P}(\varnothing) = 0$

Доказательство. $\varnothing \cap \Omega = \varnothing \Rightarrow \mathsf{P}(\varnothing \cup \Omega) = \mathsf{P}(\varnothing) + \mathsf{P}(\Omega) = \mathsf{P}(\Omega) = 1 \Rightarrow \mathsf{P}(\varnothing) = 0$ □

Утверждение 1.1.4. Пусть $A,\ B \in \mathscr{F}$, такие, что: $A \subset B$. Тогда $\mathsf{P}(B) \geq \mathsf{P}(A)$

Доказательство. Следует из $\mathsf{P}(B) = \mathsf{P}(A) + \mathsf{P}(B \setminus A)$ и того, что вероятность — неотрицательная функция. □

Утверждение 1.1.5. $\mathsf{P}(\bigcup\limits_{i=1}^{\infty} A_i) \le \sum\limits_{i=1}^{\infty} \mathsf{P}(A_i)$

Доказательство. Рассмотрим события $B_i = A_i \setminus (A_1 \cup \ldots \cup A_{i-1}) \quad i \ge 2$; $B_1 = A_1$. Тогда $B_i \cap B_j = \varnothing$, $\bigcup B_i = \bigcup A_i$.

Получается, что $\mathsf{P}(\bigcup A_i) = \mathsf{P}(\bigcup B_i) = \sum \mathsf{P}(B_i) \le \sum P(A_i)$.

Заметим, что ряд слева сходится, т.к. его значение ограничено 1, а все его члены положительны. □

Теорема 1.1. *(О непрерывности вероятностной меры в 0)*

Пусть Ω — произвольное множество, $\mathscr{F}$ — алгебра на Ω. Тогда следующие условия эквивалентны:

1. P *— вероятность на* $(\Omega, \mathscr{F})$

2. P *— конечно аддитивная вероятность на* $(\Omega, \mathscr{F})$*, непрерывная сверху:*
$$\forall A_1 \subset A_2 \subset \ldots; \quad \bigcup A_i \in \mathscr{F} \Rightarrow \lim_{n\to\infty} \mathsf{P}(A_n) = \mathsf{P}(\bigcup A_i)$$

3. P *— конечно аддитивная вероятность на* $(\Omega, \mathscr{F})$*, непрерывная снизу:*
$$\forall A_1 \supset A_2 \supset \ldots; \quad \bigcap A_i \in \mathscr{F} \Rightarrow \lim_{n\to\infty} \mathsf{P}(A_n) = \mathsf{P}(\bigcap A_i)$$

4. P *— конечно аддитивная вероятность на* $(\Omega, \mathscr{F})$*, непрерывная в нуле:*
$$\forall A_1 \supset A_2 \supset \ldots; \quad \bigcap A_i = \varnothing \Rightarrow \lim_{n\to\infty} \mathsf{P}(A_n) = 0$$

Доказательство.

1) $\Rightarrow$ 2):

Рассмотрим множества A_i, удовлетворяющие условию 2.
$$\bigcup_{n=1}^{\infty} A_n = A_1 + (A_2 \setminus A_1) + (A_3 \setminus A_2) + \ldots$$
тогда имеем
$$\mathsf{P}(\bigcup_{n=1}^{\infty} A_n) = \mathsf{P}(A_1) + \mathsf{P}(A_2 \setminus A_1) + \mathsf{P}(A_3 \setminus A_2) + \ldots =$$
$$= \mathsf{P}(A_1) + \mathsf{P}(A_2) - \mathsf{P}(A_1) + \mathsf{P}(A_3) - \mathsf{P}(A_2) + \ldots = \lim_{n\to\infty} \mathsf{P}(A_n)$$

2) $\Rightarrow$ 3):

Пусть $n \ge 1$, тогда
$$\mathsf{P}(A_n) = \mathsf{P}(A_1 \setminus (A_1 \setminus A_n)) = \mathsf{P}(A_1) - \mathsf{P}(A_1 \setminus A_n)$$
Последовательность множеств $\{A_1 \setminus A_n\}_{n\ge1}$ является неубывающей ($B_i \subset B_{i+1}$) и
$$\bigcup_{n=1}^{\infty} (A_1 \setminus A_n) = A_1 \setminus \bigcap_{n=1}^{\infty} A_n$$
Имеем, в силу 2):
$$\lim_{n\to\infty} \mathsf{P}(A_1 \setminus A_n) = \mathsf{P}(\bigcup_{n=1}^{\infty} (A_1 \setminus A_n))$$

и, значит,

$$\lim_{n\to\infty}\mathsf{P}(A_n)=\mathsf{P}(A_1)-\lim_{n\to\infty}\mathsf{P}(A_1\setminus A_n)=$$
$$=\mathsf{P}(A_1)-\mathsf{P}(\bigcup_{n=1}^{\infty}(A_1\setminus A_n))=\mathsf{P}(A_1)-\mathsf{P}(A_1\setminus\bigcap_{n=1}^{\infty}A_n)=$$
$$=\mathsf{P}(A_1)-\mathsf{P}(A_1)+\mathsf{P}(\bigcap_{n=1}^{\infty}A_n)=\mathsf{P}(\bigcap_{n=1}^{\infty}A_n).$$

3) $\Rightarrow$ 4):

Очевидно.

4) $\Rightarrow$ 1):

Пусть множества $A_1,\ A_2,\ \ldots\in\mathscr{F}$ попарно не пересекаются и $\bigcup\limits_{n=1}^{\infty}A_n\in\mathscr{F}$. Тогда

$$\mathsf{P}(\bigcup_{n=1}^{\infty}A_n)=\mathsf{P}(\bigcup_{i=1}^{n}A_i)+\mathsf{P}(\bigcup_{i=n+1}^{\infty}A_i)$$

и поскольку $\bigcup\limits_{i=n+1}^{\infty}A_i\downarrow\varnothing\quad n\to\infty$, то

$$\sum_{i=1}^{\infty}\mathsf{P}(A_i)=\lim_{n\to\infty}\sum_{i=1}^{n}\mathsf{P}(A_i)=\lim_{n\to\infty}\mathsf{P}(\bigcup_{i=1}^{n}A_i)=$$
$$=\lim_{n\to\infty}\left[\mathsf{P}(\bigcup_{i=1}^{\infty}A_i)-\mathsf{P}(\bigcup_{i=n+1}^{\infty}A_i)\right]=$$
$$=\mathsf{P}(\bigcup_{i=1}^{\infty}A_i)-\lim_{n\to\infty}\mathsf{P}(\bigcup_{i=n+1}^{\infty}A_i)=\mathsf{P}(\bigcup_{i=1}^{\infty}A_i)$$

$\square$

Определение 1.6. Набор объектов

$$(\Omega,\ \mathscr{F},\ \mathsf{P}),$$

где

a) Ω — *множество* точек ω,

b) $\mathscr{F}$ — *σ-алгебра* подмножеств Ω,

c) P — *вероятность* на $(\Omega,\mathscr{F})$,

называется *вероятностным пространством* или *вероятностной моделью* (эксперимента). При этом *пространство исходов* Ω называется пространством *элементарных исходов* (или *элементарных событий*), множества A из $\mathscr{F}$ — *событиями*, а $\mathsf{P}(A)$ — *вероятностью* события A.

Пример 1.1. *Классическая вероятность:*

Ω конечно, $\mathscr{F}=2^{\Omega}$, $\mathsf{P}(A)=\frac{|A|}{|\Omega|}$.

Например, пространство элементарных событий $\Omega=\{\omega:\omega=(a_1,\ a_2,\ \ldots,\ a_n),\ a_i\in\{\text{О},\ \text{Р}\}\}$ с алгеброй $\mathscr{F}=2^{\Omega}$ соответствует n-кратному подбрасыванию монеты.

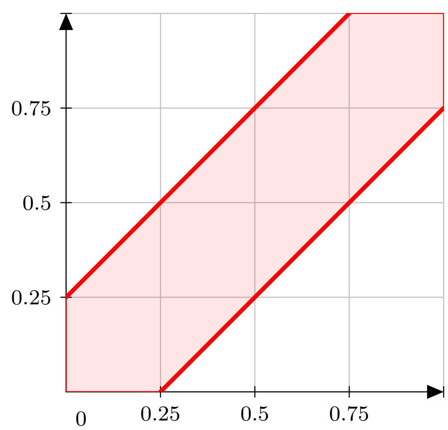
Пример 1.2. *Геометрическая вероятность*

$\Omega \subset \mathbb{R}^k$, т.ч. мера Жордана $\mu(\Omega) < \infty$ (иными словами, Ω — измеримое по Жордану подмножество $\mathbb{R}^k$). $\mathscr{F} \subseteq 2^\Omega$ — множество измеримых по Жордану подмножеств Ω. $\mathsf{P}(A) = \frac{\mu(A)}{\mu(\Omega)}$

Рассмотрим следующую задачу: Петя и Ваня приходят в столовую с 12 до 13 часов. Если Петя пришел раньше Вани, то он ждет его 15 минут и уходит. Аналогично поступает Ваня. Нужно найти вероятность того, что Петя и Ваня встретятся в столовой.

Пространство элементарных исходов составляет $\Omega = [0,1]\times[0,1] = [0,1]^2$, а событие $A = \{\text{П и В встретились}\} = \{(u,\, v) :\ |u - v| \leq \frac{1}{4}\}$. Это множество на рисунке выделено красным цветом. Тогда

$$\mathsf{P}(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{1 - 2 \cdot \frac{1}{2} \cdot \left(\frac{3}{4}\right)^2}{1} = \frac{7}{16}$$



1.2 Наименьшие сигма-алгебры

Определение 1.7. Пусть M — система подмножеств Ω. *Наименьшая σ-алгебра*, порожденая M, есть $\mathscr{F}$ — σ-алгебра, т.ч. не существует σ-алгебры, содержащей M и содержащаяся в $\mathscr{F}$ как собственное подмножество. Обозначение $\sigma(M)$.

Пример 1.3. $\mathscr{F} = \{\varnothing,\, \Omega\}$ — тривиальная σ-алгебра.

Пример 1.4. $\mathscr{F} = 2^\Omega$

Пример 1.5. $\mathscr{F}_A = \{\varnothing,\, A,\, \overline{A},\, \Omega\}$ — σ-алгебра, порожденная событием A.

Пример 1.6. $D_1,\, D_2,\, \ldots \subseteq \Omega,\ D_i \cap D_j = \varnothing,\ \bigcup D_i = \Omega$
$\mathscr{F} = \sigma(\{D_1,\, D_2,\, \ldots\})$ — наименьшая σ-алгебра, порожденная разбиением $D_1,\, D_2,\, \ldots$ — все возможные объединения конечного и бесконечного числа множеств из разбиения.

Утверждение 1.2.1. Если $M \subset 2^\Omega$, то $\sigma(M)$ существует.

Доказательство. Пусть X — мн-во всех σ-алгебр, содержащих M (очевидно, что оно не пусто). $\mathscr{F} = \bigcap\limits_{\varepsilon \in X} \varepsilon$ Покажем, что $\mathscr{F} = \sigma(M)$:

1. $\Omega \in \mathscr{F}$

2. Пусть $A \in \mathscr{F} \Rightarrow \ \forall \varepsilon \in X \quad A \in \varepsilon \Rightarrow \overline{A} \in \varepsilon \Rightarrow \overline{A} \in \mathscr{F}$

3. Для объединения счетного числа множеств доказательство аналогично 2.

Заметим, что $\mathscr{F}$ — минимальная σ-алгебра. Действительно, предположим, что $\exists \mathscr{F}_0$ - меньше $\mathscr{F}$. Тогда:

$$\mathscr{F}_0 \subset \mathscr{F},\ \mathscr{F}_0 \in X \Rightarrow \mathscr{F} \subset \mathscr{F}_0.$$

Противоречие. $\square$

2 Условная вероятность. Независимость событий

2.1 Условная вероятность

Определение 2.1. Пусть $(\Omega, \mathscr{F}, \mathsf{P})$ — вероятностное пространство. $A, B \in \mathscr{F}$ — события. Вероятностью события A *при условии* B называется величина

$$\mathsf{P}(A \mid B) = \begin{cases} \frac{\mathsf{P}(A\cap B)}{\mathsf{P}(B)}, & \mathsf{P}(B) \neq 0 \\ 0, & \mathsf{P}(B) = 0 \end{cases}$$

Замечание 1: Если $\mathsf{P}(B) = 0$, то $\mathsf{P}(A \mid B) = 0$.

Замечание 2: В случае классической вероятности $\mathsf{P}(A) = \frac{|A|}{|\Omega|}$, $\mathsf{P}(A \cap B) = \frac{|A\cap B|}{|\Omega|}$, а значит

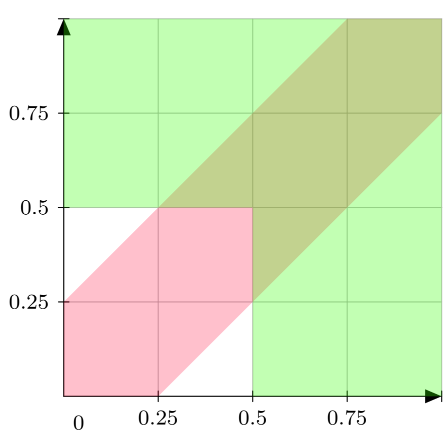
$$\mathsf{P}(B \mid A) = \frac{|A\cap B|}{|A|}.$$

Установим некоторые очевидные свойства условных вероятностей:

1. $\mathsf{P}(A \mid A) = 1$,
2. $\mathsf{P}(\varnothing \mid A) = 0$,
3. $\mathsf{P}(B \mid A) = 1, B \supseteq A$,
4. $\mathsf{P}(B_1 \sqcup B_2 \mid A) = \mathsf{P}(B_1 \mid A) + \mathsf{P}(B_2 \mid A)$,
5. $\mathsf{P}(B \mid A) + \mathsf{P}(\overline{B} \mid A) = 1$.

Пример 2.1. Найдем вероятность встречи Васи и Пети при условии, что хотя бы один из них приходит во вторую половину часа. Как и прежде, множество точек, удовлетворяющих событию $A = \{$Вася и Петя встретились$\}$ выделено красным цветом, а множество точек, удовлетворяющих событию $B = \{$хотя бы один пришел во вторую половину часа$\}$ — зелеными. Тогда ответ на задачу

$$\mathsf{P}(A \mid B) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)} = \frac{1 - \frac{1}{2} - \frac{1}{4}}{1 - \frac{1}{4}} = \frac{1}{3}$$



Теорема 2.1. *Формула полной вероятности:*

Пусть $\{D_1,\ D_2,\ \dots,\ \}$ *— некоторое разбиение множества* Ω*, а* A *— некоторое событие. Тогда:*

$$\mathsf{P}(A) = \sum_{n=1}^{\infty} \mathsf{P}(A \mid D_n) P(D_n).$$

Доказательство. Ясно, что

$$A = (A \cap D_1) \sqcup (A \cap D_2) \sqcup \dots$$

и, значит,

$$\mathsf{P}(A) = \sum_{i=1}^{\infty} \mathsf{P}(A \cap D_i).$$

но

$$\mathsf{P}(A \cap D_i) = \mathsf{P}(A \mid D_i)\mathsf{P}(D_i).$$

□

Пример 2.2. В ящике находиться n шаров: k черных $n-k$ белых. Случайным образом без возвращения из ящика вытягиваются шары. Какова вероятность события $A = \{$ при j-том вытягивании достали черный шар$\}$ для $j \in \{1,\ 2,\ \dots,\ n\}$.

Решение: $\Omega = \{\omega \in \sigma(1,\ 2,\ \dots,\ n)$ — перестановки $\}$. Пусть черным шарам соотвествует цифра 1, а белым — 0.

Первый способ: Рассмотрим два события: $A_1 = \{$на 1 вытягивании достали черный шар$\} = \{(1, k_2, \dots, k_n)\}$, а $A_j = \{$на j вытягивании достали черный шар$\} = \{(k_1, \dots, k_{j-1}, 1, k_{j+1}, \dots, k_n)\}$. Очевидно, что между A_1 и A_j есть биекция, а значит $|A_1| = |A_j| \Rightarrow \mathsf{P}(A_j) = \mathsf{P}(A_1) = \frac{k}{n}$

Второй способ: По индукции. $\mathsf{P}(A_1) = \frac{k}{n}$. Предположение: пусть при изначальном количестве черных шаров x и всех шаров m вероятность $\mathsf{P}(A_{j-1}) = \frac{x}{m}$.

Тогда, если всего шаров n, а черных шаров k, то

$$\mathsf{P}(A_j) = \mathsf{P}(A_j \mid A_1)\mathsf{P}(A_1) + \mathsf{P}(A_j \mid \overline{A}_1)\mathsf{P}(\overline{A}_1) = \frac{k-1}{n-1} \cdot \frac{k}{n} + \frac{k}{n-1} \cdot \frac{n-k}{n} = \frac{k}{n}$$

□

Теорема 2.2. *Формула Байеса:*

$\{D_1,\ D_2,\ \cdots\}$ *— разбиение* Ω*.* A *— событие. Тогда справедлива формула Байеса:*

$$\mathsf{P}(D_n \mid A) = \frac{\mathsf{P}(A \mid D_n)\mathsf{P}(D_n)}{\sum\limits_{i=1}^{\infty} \mathsf{P}(A \mid D_i)\mathsf{P}(D_i)}$$

Доказательство. Следует из формулы полной вероятности $\mathsf{P}(A) = \sum\limits_{n=1}^{\infty} \mathsf{P}(A \mid D_n)\mathsf{P}(D_n)$ и того, что $\mathsf{P}(D_n \mid A) = \frac{\mathsf{P}(D_n \cap A)}{\mathsf{P}(A)}, \quad \mathsf{P}(D_n \cap A) = \mathsf{P}(A \mid D_n)\mathsf{P}(D_n)$ □

2.2 Независимость событий

Определение 2.2. События A и B называются *независимыми*, если $\mathsf{P}(A \cap B) = \mathsf{P}(A)\mathsf{P}(B)$. Если $\mathsf{P}(B) \neq 0$, то

$$\mathsf{P}(A) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)} = \mathsf{P}(A \mid B)$$

Определение 2.3. События $A_1,\ A_2,\ \ldots\ A_n$ называются *независимыми в совокупности*, если

$$\forall n_1,\ n_2,\ \ldots\ ,\ n_k\ (k \geq 2,\ k \leq n) \quad \mathsf{P}(A_{n_1} \cap A_{n_2} \cap \ldots \cap A_{n_k}) = \prod_{i=1}^{k} \mathsf{P}(A_{n_i})$$

Утверждение 2.2.1. Попарно независимые события не обязательно независимы в совокупности.

Доказательство. Приведем контрпример (пример Бернштейна): покрасим грани тетраэдра в 3 цвета: одну во все 3 цвета (красный, зеленый и синий) и 3 других в различные.

Рассмотрим следующие события:

- $A_R =$ 'На грани тетраэдра есть красный цвет'
- $A_G =$ 'На грани тетраэдра есть зеленый цвет'
- $A_B =$ 'На грани тетраэдра есть синий цвет'

тогда:

- $\mathsf{P}(A_R \cap A_G \cap A_B) = \frac{1}{4}$
- $\mathsf{P}(A_R \cap A_G) = \mathsf{P}(A_R \cap A_B) = \mathsf{P}(A_G \cap A_B) = \frac{1}{4}$
- $\mathsf{P}(A_R) = P(A_G) = \mathsf{P}(A_B) = \frac{1}{2}$

Приведенные события попарно независимы, но не независимы в совокупности. □

Определение 2.4. Пусть дано мн-во $\{A_\alpha\}_{\alpha \in \Gamma}$. A_α называется *независимым в совокупности*, если

$$\forall n \in \mathbb{N}\, \forall \text{ различных } \alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n \in \Gamma :\ A_{\alpha_1},\ \ldots,\ A_{\alpha_n} \text{ независимы в совокупности.}$$

Определение 2.5. $\mathscr{M}_1,\ \ldots,\ \mathscr{M}_n \subset \mathscr{F}$ называются *независимыми в совокупности*, если

$$\forall A_i \in \mathscr{M}_i \quad A_1,\ \ldots,\ A_n \text{ независимы в совокупности}$$

Утверждение 2.2.2. Пусть $A_1,\ A_2,\ \ldots,\ A_n$ — независимые в совокупности события. Тогда $\mathscr{F}_{A_i} = \{\varnothing,\ \Omega,\ A_i,\ \overline{A}_i\}$ — независимы в совокупности.

Доказательство. Пусть $B_i \in F_i$. Докажем, что $P(B_1 \cap \ldots \cap B_n) = \prod\limits_{i=1}^{n} P(B_i)$.

Случаи, когда одно из B_i равно $\varnothing$ очевиден. Если же одно из B_i равно Ω, то это очевидным образом сводиться к случаю, когда такого B_i нет (множитель $\mathsf{P}(B_i)$ равен 1, а на пересечение событий множество $\mathscr{B}_i$ не влияет). Предположим теперь, что $\forall i\ B_i \neq \varnothing,\ \Omega$. Пусть k это число множеств из B_i вида $\overline{A}_i$. Докажем утверждение индукцией по k.

База индукции при $k = 0$ следует из условия независимости событий. Покажем переход индукции, заметив, что $\overline{A}_1 = \Omega \setminus A_1$.

$$\begin{aligned}\mathsf{P}(\overline{A}_1 \cap \ldots \cap \overline{A}_k \cap A_{k+1} \cap \ldots \cap A_n) &= \mathsf{P}(\overline{A}_2 \cap \ldots \cap \overline{A}_k \cap A_{k+1} \cap \ldots \cap A_n) - \mathsf{P}(A_1 \cap \ldots \cap \overline{A}_k \cap A_{k+1} \cap \ldots \cap A_n) = \\ &= \mathsf{P}(\overline{A}_2) \ldots \mathsf{P}(\overline{A}_k)\mathsf{P}(A_{k+1}) \ldots \mathsf{P}(A_n) - \mathsf{P}(A_1)\mathsf{P}(\overline{A}_2) \ldots \mathsf{P}(\overline{A}_k)\mathsf{P}(A_{k+1}) \ldots \mathsf{P}(A_n) = \\ &= \mathsf{P}(\overline{A}_2) \ldots \mathsf{P}(\overline{A}_k)\mathsf{P}(A_{k+1}) \ldots \mathsf{P}(A_n)(1 - \mathsf{P}(A_1)) = \mathsf{P}(\overline{A}_1) \ldots \mathsf{P}(\overline{A}_k)\mathsf{P}(A_{k+1}) \ldots \mathsf{P}(A_n)\end{aligned}$$

□

3 Распределение вероятностей

3.1 Борелевские сигма-алгебры. Функция распределения

Определение 3.1. Пусть $\mathbb{R}$ действительная прямая и

$$(a, b] = \{x \in \mathbb{R} : \ a < x \leq b\}$$

Обозначим через $\mathscr{A}$ систему множеств в $\mathbb{R}$, состоящую из *конечных* объединений непересекающихся интервалов вида $(a,\ b]$:

$$A \in \mathscr{A}, \quad \text{если } A = \bigcup_{i=1}^{n} (a_i,\ b_i], \quad n < \infty$$

Нетрудно проверить, что данная система множеств в объединении с пустым образует алгебру, которая, однако, не является σ-алгеброй, поскольку если положить $A_n = (0, 1-\frac{1}{n}] \in \mathscr{A}$, то $\bigcup\limits_n A_n = (0, 1) \notin \mathscr{A}$

Пусть $\mathscr{B}(\mathbb{R})$ — наименьшая σ-алгебра $\sigma(\mathscr{A})$, содержащая систему $\mathscr{A}$. Полученная σ-алгебра называется *борелевской* алгеброй на числовой прямой, а ее множества — *борелевскими.*

Замечание. 1: Заметим, что

$$(a, b) = \bigcup_{n=1}^{\infty} (a,\ b - \frac{1}{n}], \quad a < b, \qquad [a, b] = \bigcap_{n=1}^{\infty} (a - \frac{1}{n},\ b], \quad a < b,$$

$$\{a\} = \bigcap_{n=1}^{\infty} (a - \frac{1}{n},\ a]$$

Таким образом, в борелевскую σ-алгебру наряду с интервалами $(a,\ b]$ входят одноточечные множества $\{a\}$ а так же каждое из семи множеств вида

$$(a,\ b), \quad [a,\ b], \quad [a,\ b), \quad (-\infty,\ b), \quad (-\infty,\ b], \quad (a,\ +\infty), \quad [a,\ -\infty).$$

Замечание. 2: В общем случае борелевской σ-алгеброй называется минимальная σ-алгебра, содержащая все открытые множества топологического пространства.

Определение 3.2. Пара $(\mathbb{R},\ \mathscr{B}(\mathbb{R}))$ называется *измеримым пространством.*

Определение 3.3. Любая вероятностная мера на $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ называется *распределением вероятностей.*

Определение 3.4. *Функцией распределения*, соответствующей распределению вероятностей P, называется функция $F : \mathbb{R} \to [0,\ 1]$, такая, что:

$$F(x) = \mathsf{P}((-\infty,\ x]).$$

Теорема 3.1. *Любая функция распределения обладает следующими свойствами:*

1. $\lim\limits_{x\to-\infty} F(x) = 0$ *и* $\lim\limits_{x\to+\infty} F(x) = 1$

2. F *—непрерывна справа:* $\lim\limits_{x\to x_0+0} F(x) = F(x_0)$

3. F *— неубывает*

Доказательство. 1. $\lim\limits_{x\to-\infty} F(x) = \lim\limits_{x\to-\infty} \mathsf{P}((-\infty,\, x])$

Т.к. $(-\infty,\, x] \downarrow \varnothing$, то по теореме о непрерывности вероятностной меры в 0:

$$\lim_{x\to-\infty} \mathsf{P}(-\infty,\, x]) = \mathsf{P}(\varnothing) = 0.$$

Далее, по той же теореме о непрерывности вероятностной меры в 0, $\lim\limits_{x\to+\infty} F(x) = \lim\limits_{x\to+\infty} \mathsf{P}((-\infty, x]) = 1$, т.к. $(-\infty,\, x] \uparrow \mathbb{R}$.

2. Пусть $x \to x_0 + 0$. Тогда $(-\infty,\, x] \downarrow (-\infty,\, x_0]$ и по теореме о непрерывности вероятностной меры в 0 $\mathsf{P}((-\infty,\, x]) \to \mathsf{P}((\infty,\, x_0])$

3. Пусть $x > y$. Тогда $(-\infty,\, x] \supset (-\infty,\, y] \Rightarrow \mathsf{P}((-\infty,\, x]) \geq \mathsf{P}(-\infty,\, y]) \Rightarrow F(x) \geq F(y)$

$\square$

Замечание: Функция распределения не обязательно непрерывна слева, т.к $(-\infty, x] \xrightarrow[x\to x_0-0]{} (-\infty, x_0)$,а значит $\mathsf{P}((-\infty,\, x]) \to \mathsf{P}((-\infty,\, x_0]) - \mathsf{P}(\{x_0\})$

Теорема 3.2. *(б/д)*

Пусть $F : \mathbb{R} \to [0,\, 1]$ *обладает свойствами* $1) - 3)$*. Тогда существует единственное распределение вероятностей, т.ч.* F *— его функция распределения.*

Любую функцию, обладающую этими свойствами, будем называть функцией распределения.

Пример 3.1. $F(x) = \begin{cases} 1, & x \geq a \\ 0, & x < a \end{cases}$. $\mathsf{P}(A) = I(a \in A) = \begin{cases} 1, & a \in A \\ 0, & a \notin A \end{cases}$

Пример 3.2. $F(x) = \begin{cases} 0, & x < 0 \\ x, & x \in [0,\, 1) \\ 1, & x \geq 1 \end{cases}$. $0 \leq a < b \leq 1 \quad \Rightarrow \mathsf{P}((a,\, b)) = b - a$

P — мера Лебега на отрезке $[0,\, 1]$

3.2 Виды распределений

1. *Дискретные*

   Пусть $X \subset \mathbb{R}$ — не более чем счетный набор точек, т.ч. $\mathsf{P}(X) = 1$, $\mathsf{P}(\mathbb{R} \setminus X) = 0$.

   Тогда P называется дискретным распределением.

   $X = \{x_n \mid n \in \mathbb{N}\}$.

$$F(x) = \sum_{\substack{n \in \mathbb{N} \\ x_n \le x}} \mathsf{P}(\{x_n\})$$

Обозн: $\mathsf{P}(\{x_n\}) = p_n$. Часто последовательность $\{p_n\}$ называется распределением вероятностей, т.к. распределение вероятностей в стандартном понимании однозначно восстанавливается по ней.

2. *Абсолютно непрерывные*

   Пусть F — функция распределения. Если существует функция $p : \mathbb{R} \to \mathbb{R}_+$, такая что
   $$\forall x\; F(x) = \int\limits_{-\infty}^{x} p(t)dt$$
   то распределение вероятностей и F называются абсолютно непрерывными, а p — плотностью этого распределения.

Замечание: Приведены **не все** распределения из существующих.

Утверждение 3.2.1. Пусть $p : \mathbb{R} \to \mathbb{R}_+$ и $\int\limits_{-\infty}^{+\infty} p(t)dt = 1$. Тогда функция
$$F(x) = \int\limits_{-\infty}^{x} p(t)dt$$
является функцией распределения.

Доказательство. Покажем, что определенная таким образом функция F удовлетворяет всем свойствам функции распределения:

1. $$\lim_{x\to-\infty} F(x) = \lim_{x\to-\infty} \int\limits_{-\infty}^{x} p(t)dt = \lim_{x\to-\infty} \int\limits_{-\infty}^{\infty} p(t)I(t \le x)dt = \int\limits_{-\infty}^{\infty} \lim_{x\to-\infty} p(t)I(t \le x)dt = 0$$
   $$\lim_{x\to+\infty} F(x) = \int\limits_{-\infty}^{+\infty} \lim_{x\to+\infty} p(t)I(t \in (-\infty,\, x])dt = \int\limits_{-\infty}^{+\infty} p(t)dt = 1$$

2. Непрерывность справа:
   $$\lim_{x\to x_0+0} F(x) = \lim_{x\to x_0+0} \int\limits_{-\infty}^{x} p(t)dt = \int\limits_{-\infty}^{x_0} p(t)dt = F(x_0)$$

3. Неубывание:
   $$x > y \Rightarrow F(x) = \int\limits_{-\infty}^{x} p(t)dt = \int\limits_{-\infty}^{y} p(t)dt + \int\limits_{y}^{x} p(t)dt = F(y) + \int\limits_{y}^{x} p(t)dt \Rightarrow$$
   неубывание функции F, т.к. p принимает только неотрицательные значения.

□

Определение 3.5. Любую функцию, удовлетворяющую свойствам:

a) $p : \mathbb{R} \to \mathbb{R}_+$

b) $\int\limits_{-\infty}^{+\infty} p(t)dt = 1$

будем называть плотностью вероятности.

Утверждение 3.2.2. $\forall B \in \mathscr{B}(\mathbb{R}) \quad \mathsf{P}(B) = \int\limits_B p(x)dx$, (где интеграл понимается в смысле интеграла Лебега.)

Доказательство (идея).

1. Доказать, что определенная таким образом функция является распределением вероятностей на $\mathscr{B}(\mathbb{R})$
2. Понять, что $\forall x\, F(x) = \int\limits_{-\infty}^{x} p(t)dt$, где F — функция распределения P.
3. из теоремы о единственности распределения вероятностей следует утверждение

□

Упражнение: Если F — дифференцируемая функция, то $p(x) = F'(x)$.

3.3 Примеры распределений

Пример 3.3. $Bern(p)$, $(p \in (0,\,1)\,)$ — распределение Бернулли.
$\mathsf{P}(0) = 1 - p =: q, \quad \mathsf{P}(1) = p$. Дискретное распределение,соответствующее однократному подбрасыванию монеты.

Пример 3.4. $Bin(n, p)$ — биномиальное распределение. $\mathsf{P}(\{k\}) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k \in \{0,\,\dots,\,n\}$ — дискретное распределение, соответствующее n-кратному подбрасыванию монеты.

Пример 3.5. $Pois(\lambda)$, $\mathsf{P}(\{k\}) = \frac{\lambda^k e^{-\lambda}}{k!}, k \in \mathbb{Z}_+,\, \lambda \in \mathbb{R}_+$ — распределение Пуасонна. Дискретное распределения числа успехов в испытаниях Бернулли, где число испытаний n достаточно велико, а вероятность успеха равна $\frac{\lambda}{n}$.

Пример 3.6. $U(\{1,\,2,\,\dots,\,n\})$ — Дискретное равномерное распределение, $\mathsf{P}(\{k\}) = \frac{1}{n}$

Пример 3.7. $U([a,\,b])$ — Непрерывное равномерное распределение на отрезке $[a,\,b]$. $p(x) = \frac{1}{b-a} I(a \leq x \leq b)$

Пример 3.8. $N(\mu,\,\sigma^2)$ — Нормальное распределение. $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

Пример 3.9. $exp(\lambda)$, $\lambda > 0$ — экспоненциальное распределение, где $p(x) = \lambda e^{-\lambda x} I(x \geq 0)$.

Пример 3.10. $\Gamma(\alpha,\ \beta)$, $\quad \alpha,\ \beta > 0$ - гамма распределение. $p(x) = \frac{x^{\beta-1}e^{-\alpha x}}{\Gamma(\beta)} \cdot I(x \geq 0)$, где $\Gamma(\beta)$ — гамма-функция.

Пример 3.11. $C(\theta)$, $\theta > 0$ — распределение Коши. $p(x) = \frac{\theta}{\pi(x^2+\theta^2)}$

3.4 Многомерные распределения

Определение 3.6. $\mathscr{B}(\mathbb{R}^n)$ — наименьшая σ-алгебра, такая, что:

$$\mathscr{B}(\mathbb{R}^n) = \sigma(\{(-\infty,\ a_1] \times (-\infty,\ a_2] \times\ \dots\ \times\ (-\infty,\ a_n];\ a_i \in \mathbb{R}\}) =$$
$$= \sigma(\{B_1 \times B_2 \times\ \dots\ \times B_n;\ \forall i\ B_i \in \mathscr{B}(\mathbb{R})\})$$

Определение 3.7. Если P — вероятностная мера на $\mathscr{B}(\mathbb{R}^n)$, то P называется n-мерным распределением вероятностей.

Определение 3.8. $F : \mathbb{R}^n \to [0,\ 1]$ называется функцией распределения, соответствующей распределению P, если

$$F(x_1,\ \dots\ x_n) = \mathsf{P}((-\infty, x_1]\ \times\ \dots\ \times\ (-\infty,\ x_n])$$

Обозначим

$$F(x_1,\ \dots\ x_n) = F_n(x) = \mathsf{P}((-\infty,\ x]),$$

где $x = (x_1,\ \dots,\ x_n)$, $(-\infty,\ x] = (-\infty, x_1]\ \times\ \dots\ \times\ (-\infty,\ x_n]$.

Введем разностный оператор $\Delta_{a_i b_i} : \mathbb{R}^n \to \mathbb{R}$, действующий по формуле $(a_i \leq b_i)$

$$\Delta_{a_i b_i} F_n(x_1,\ \dots,\ x_n) = F_n(x_1,\ \dots,\ x_{i-1},\ b_i, x_{i+1},\ \dots,\ x_n) -$$
$$- F_n(x_1,\ \dots,\ x_{i-1},\ a_i,\ x_{i+1},\ \dots,\ x_n).$$

По индукции можно показать, что

$$\Delta_{a_1 b_1} \dots \Delta_{a_n b_n} F_n(x) = \mathsf{P}((a,\ b]),$$

где $(a,\ b] = (a_1,\ b_1] \times \dots \times (a_n,\ b_n]$. В частности, отсюда видно, что, *в отличие* от одномерного случая, вероятность $\mathsf{P}(a,\ b]$, вообще говоря, *не равна* разности $F_n(b) - F_n(a)$.

Теорема 3.3. *Функция распределения, соответствующая распределению* P, *обладает следующими свойствами:*

1. $\lim\limits_{\substack{x_1 \to +\infty \\ \dots \\ x_n \to +\infty}} F(x) = 1$ *и* $\forall i \in \{1,\ \dots\ n\}$ $\lim\limits_{x_i \to -\infty} F(x) = 0$
2. *Пусть* $x = (x_1,\ \dots,\ x_n) \geq (y_1,\ \dots,\ y_n) = y \iff \forall i \in \{1,\ \dots,\ n\}\ x_i \geq y_i$,
   $x^{(k)} \downarrow x \iff (x^{(k+1)} \leq x^{(k)}\ \wedge\ \forall i \in \{1,\ \dots,\ n\}\ x_i^{(k)} \to x_i)$. *Тогда:*
   $$\lim_{k \to +\infty} F(x^{(k)}) = F(x).$$

3. $\Delta_{a_1 b_1} \dots \Delta_{a_n b_n} F_n(x) \geq 0$

Доказательство. 1. Зафиксируем $\varepsilon > 0$, $B_k = (-\infty,\, k]^n$. Тогда очевидно, что:

$$B_{k+1} \supset B_k,\ \bigcup_{k=1}^{\infty} B_k = \mathbb{R}^n.$$

По теореме о непрерывности вероятностной меры в 0 $\lim\limits_{k\to+\infty} \mathsf{P}(B_k) = 1 \Rightarrow$
$\Rightarrow \exists N\, \forall k \geq N:\ \mathsf{P}(B_k) > 1-\varepsilon.$

Пусть $x \in \mathbb{R}^n$ такое, что: $x_1 \geq N,\, \dots,\, x_n \geq N$. Тогда $(-\infty,\, x] \supset (-\infty,\, N]^n$ и

$$F_n(x) = \mathsf{P}((-\infty,\, x]) \geq \mathsf{P}(B_k) > 1-\varepsilon.$$

Для доказательства второй части, без ограничения общности предположим, что $x_1 \to -\infty$ и зафиксируем $x_2,\, \dots,\, x_n$. Пусть $B_k = (-\infty,\, -k] \times (-\infty,\, x_2] \times \dots \times (-\infty,\, x_n]$. Тогда:

$$B_{k+1} \subset B_k,\ \bigcap_{k=1}^{\infty} B_k = \varnothing.$$

$\lim\limits_{k\to+\infty} F(B_k) = 0 \Rightarrow \exists N\, \forall k \geq N\ \mathsf{P}(B_k) < \varepsilon$. Пусть $x_1 < N$. Тогда:

$$F(x) = \mathsf{P}((-\infty,\, x]) \leq \mathsf{P}((-\infty,\, N] \times (-\infty,\, x_2] \times \dots \times (-\infty,\, x_n]) < \varepsilon$$

2. Следствие теоремы о непрерывности вероятностной меры в 0.

3. $\Delta_{a_1 b_1} \dots \Delta_{a_n b_n} F_n(x) = \mathsf{P}((a,\, b]) \geq 0$

□

Теорема 3.4. *(б/д)*

Если $F : \mathbb{R}^n \to [0,\, 1]$ обладает свойствами 1)-3), то F является функцией распределения для некоторого распределения вероятностей, причем такое распределение единственное.

Пример 3.12. Пусть $F^1,\, \dots,\, F^n$ — одномерные функция распределения на $\mathbb{R}$ и

$$F_n(x_1,\, \dots,\, x_n) = F^1(x_1) F^2(x_2) \dots F^n(x_n).$$

Нетрудно проверить, что такая функция удовлетворяет условиям *1)-3)*, а значит является некоторой функцией распределения.

Особо важен случай, когда

$$F^k(x_k) = \begin{cases} 0, & x_k < 0, \\ x_k, & 0 \leq x_k \leq 1 \\ 1, & x_k > 1. \end{cases}$$

В этом случае для всех $0 \leq x_k \leq 1$, $k = 1,\, \dots,\, n$.

$$F_n(x_1,\, \dots,\, x_n) = x_1 x_2 \dots x_n$$

Такая F соответствует мере Лебега в $[0,\, 1]^n$.

Определение 3.9. Если существует $p : \mathbb{R}^n \to \mathbb{R}_+$ такая, что

$$\forall x_1,\, \dots,\, x_n \in \mathbb{R}\, F(x_1,\, \dots,\, x_n) = \int\limits_{-\infty}^{x_1} \dots \int\limits_{-\infty}^{x_n} p(t_1,\, \dots,\, t_n) dt_1 dt_2 \dots dt_n$$

то распределение вероятностей и функция распределения называются абсолютно непрерывными, а p называется плотностью этого распределения.

Утверждение 3.4.1. $\forall B \in \mathscr{B}(\mathbb{R})\ \mathsf{P}(B) = \int\limits_B p(t_1,\ \ldots,\ t_n)dt_1dt_2\ldots dt_n$

Утверждение 3.4.2. Пусть функция $p : \mathbb{R}^n \to \mathbb{R}_+$ такая, что

$$\int\limits_{\mathbb{R}^n} p(t_1,\ \ldots,\ t_n)dt_1dt_2\ldots dt_n = 1.$$

Тогда p — плотность некоторого абсолютно непрерывного распределения.

Утверждение 3.4.3. Если существует $\dfrac{\partial^n}{\partial x_1 \partial x_2 \ldots \partial x_n} F(x_1,\ x_2,\ \ldots,\ x_n)$, то она является плотностью распределения F.

Определение 3.10. Пусть $X \subset \mathbb{R}^n$ не более чем счетное множество точек n-мерного пространства. Если $\mathsf{P}(X) = 1$, $\mathsf{P}(\mathbb{R}^n \setminus X) = 0$, то распределение P называется дискретным. Аналогичено одномерному случаю, последовательность $\{p_n\}$, $p_n = P(x^{(n)})$, $X = \{x^{(1)},\ \ldots,\ x^{(n)},\ \ldots\}$ иногда называется распределением вероятностей, т.к. распределение вероятностей в стандратном смысле однозначно восстанавливается по ней.

4 Случайные величины

4.1 Измеримые функции. Случайные величины.

Определение 4.1. Пусть $(\Omega,\, \mathscr{F})$, $(E,\, \mathscr{E})$ — измеримые пространства и $f : \Omega \to E$. Функция f называется $(\mathscr{F}|\mathscr{E})$-измеримой, если

$$\forall A \in \mathscr{E} :\ f^{-1}(A) \in \mathscr{F}.$$

Если $E = \mathbb{R}$, $\mathscr{E} = \mathscr{B}(\mathbb{R})$, то измеримая функция f называется *случайной величиной*.

Если $E = \mathbb{R}^n$, $\mathscr{E} = \mathscr{B}(\mathbb{R}^n)$, то измеримая функция f называется *случайным вектором*.

В случае, когда $(\Omega,\, \mathscr{F}) = (\mathbb{R}^n,\, \mathscr{B}(\mathbb{R}^n))$, а $(E,\, \mathscr{E}) = (\mathbb{R}^k,\, \mathscr{B}(\mathbb{R}^k))$, то $(\mathscr{F}|\mathscr{E})$-измеримые функции называются *борелевскими*.

Случайные величины(и векторы) принято обозначать греческими буквами: $\xi,\, \eta,\, \ldots$. Введем также следующие обозначения:

- $\mathsf{P}(\xi \in B) = \mathsf{P}(\{\omega :\ \xi(\omega) \in B\})$
- $\mathsf{P}(\xi = x) = \mathsf{P}(\{\omega :\ \xi(\omega) = x\})$
- $\mathsf{P}(\xi > x) = \mathsf{P}(\{\omega :\ \xi(\omega) > x\})$
- $\mathsf{P}(\xi < x) = \mathsf{P}(\{\omega :\ \xi(\omega) < x\})$
- $\mathsf{P}(\xi \geq x) = \mathsf{P}(\{\omega :\ \xi(\omega) \geq x\})$
- $\mathsf{P}(\xi \leq x) = \mathsf{P}(\{\omega :\ \xi(\omega) \leq x\})$

Утверждение 4.1.1. Пусть $\mathscr{F}_\xi = \{A \in \mathscr{F} :\ \exists B \in \mathscr{B}(\mathbb{R}^k) :\ A = \xi^{-1}(B)\}$. Тогда F_ξ — σ-алгебра. Эта σ-алгебра называется *σ-алгеброй, порожденной случайной величиной* ξ. (Иногда она обозначается $\sigma(\xi)$)

Доказательство. Покажем, что система множеств $\mathscr{F}_\xi$ удовлетворяет определению σ-алгебры:

1. $\Omega \in \mathscr{F}_\xi$, т.к. $\Omega = \xi^{-1}(\mathbb{R}^k)$
2. $A \in \mathscr{F}_\xi \Rightarrow \exists B \in \mathscr{B}(\mathbb{R}^k) :\ A = \xi^{-1}(B) \Rightarrow \overline{B} \in \mathscr{B}(\mathbb{R}^k) \Rightarrow \xi^{-1}(\overline{B}) = \overline{A} \in \mathscr{F}$
3. $A_1,\, A_2,\, \ldots \in \mathscr{F}_\xi \Rightarrow \exists B_1,\, B_2,\, \ldots \in \mathscr{B}(\mathbb{R}^k) :\ \xi(A_i) = B_i \Rightarrow \xi^{-1}(\bigcup B_i) = \bigcup A_i \in \mathscr{F}$

$\square$

Утверждение 4.1.2. Пусть ξ — случайный вектор. η является $\mathscr{F}_\xi$-измеримым случайным вектором $\iff \mathscr{F}_\eta \subset \mathscr{F}_\xi$.

Доказательство.

$\Rightarrow$:

Пусть η — $\mathscr{F}_\xi$ измерима. Тогда $\forall B \in \mathscr{B}(\mathbb{R}^k) :\ \eta^{-1}(B) \in \mathscr{F}_\xi \iff \mathscr{F}_\eta = \{\eta^{-1}(B)\} \subset \mathscr{F}_\xi$.

$\Leftarrow$:

Пусть $\mathscr{F}_\eta \subset \mathscr{F}_\xi$. Возьмем произвольное $B \in \mathscr{B}(\mathbb{R}^k)$. Тогда $\eta^{-1}(B) \in \mathscr{F}_\eta \Rightarrow \eta^{-1}(B) \in \mathscr{F}_\xi$ $\square$

Утверждение 4.1.3. Пусть ξ — n-мерный случайный вектор. η является $\mathscr{F}_\xi$ измеримым k-мерным случайным вектором $\iff$ существует борелевская функция $g : \mathbb{R}^n \to \mathbb{R}^k$, т.ч $\eta = g(\xi)$.

Доказательство. Доказательство **только** в $\Leftarrow$.

Пусть $\eta = g(\xi)$. Покажем, что в таком случае $\mathscr{F}_\eta \subset \mathscr{F}_\xi$. Действительно, рассмотрим произвольное $A \in \mathscr{F}_\eta$. Для него существует $B \in \mathscr{B}(\mathbb{R}^k):\ A = \eta^{-1}(B) = (g(\xi))^{-1}(B) = \xi^{-1}(g^{-1}(B)),\ B_g := g^{-1}(B) \in \mathscr{B}(\mathbb{R}^n),\ \xi^{-1}(B_g) \in \mathscr{F}_\xi \Rightarrow \mathscr{F}_\eta \subset \mathscr{F}_\xi$. $\square$

Теорема 4.1. *Критерий измеримости*

Пусть даны $(\Omega,\ \mathscr{F})$, $(E,\ \mathscr{E})$ — измеримые пространства, $M \subset \mathscr{E}$, $\sigma(M) = \mathscr{E}$. Тогда для $(\mathscr{F} \mid \mathscr{E})$-измеримости функции $f : \Omega \to E$ необходимо и достаточно, чтобы $\forall B \in M\ f^{-1}(B) \in \mathscr{F}$.

Доказательство. Необходимость очевидна по определению. Для доказательства достаточности воспользуемся *принципом подходящих множеств.*

Рассмотрим множество $\mathscr{A} = \{B \in \mathscr{E} \mid f^{-1}(B) \in \mathscr{F}\}$. Из условия теоремы следует, что $M \subset \mathscr{A}$. Докажем, что система множеств $\mathscr{A}$ является σ-алгеброй. Действительно:

1. $E \in \mathscr{A}$, т.к. $f^{-1}(E) = \Omega$
2. Пусть $A \in \mathscr{A}$. Тогда $f^{-1}(\overline{A}) = \overline{f^{-1}(A)} \in \mathscr{F} \Rightarrow \overline{A} \in \mathscr{A}$
3. $A_1,\ A_2,\ \dots \in \mathscr{A}$. Тогда $f^{-1}(A_i) \in \mathscr{F}$, откуда:
$$f^{-1}(\bigcup A_i) = \bigcup f^{-1}(A_i) \Rightarrow \bigcup A_i \in \mathscr{A}$$

Мы доказали, что $\mathscr{A}$ — σ-алгебра, причем $M \subset \mathscr{A} \subset \mathscr{E}$. Но $\sigma(M) = \mathscr{E} \Rightarrow \mathscr{A} = \mathscr{E}$ $\square$

Следствие 4.1.1. *Для того чтобы функция $\xi : \Omega \to \mathbb{R}$ была случайной величиной, необходимо и достаточно, чтобы $\forall x \in \mathbb{R}$*
$$\{\omega :\ \xi(\omega) < x\} \in \mathscr{F}$$
или
$$\{\omega :\ \xi(\omega) \leq x\} \in \mathscr{F}.$$

Доказательство следует из того, что каждая из систем множеств
$$\mathscr{E}_1 = \{x :\ x < c,\ c \in \mathbb{R}\},$$
$$\mathscr{E}_2 = \{x :\ x \leq c,\ c \in \mathbb{R}\}$$
порождают σ-алгебру $\mathscr{B}(\mathbb{R})$.

Следствие 4.1.2. $\xi = (\xi_1,\ \xi_2,\ \dots,\ \xi_n)$ *— случайный вектор* $\iff$ $\xi_1,\ \dots,\ \xi_n$ *— случайные величины.*

Доказательство.

$\Rightarrow$:

$pr_n^i(\xi) = \xi_i$, где ξ — случайный вектор. Функция проектор — борелевская, и по утверждению 4.1.3 ξ_i — случайная величина.

$\Leftarrow$:

Пусть $B_1, \ldots, B_n \in \mathscr{B}(\mathbb{R})$, $B = B_1 \times B_2 \times \ldots \times B_n$, $\xi^{-1}(B) = \bigcap\limits_{i=1}^{n} \xi_i^{-1}(B_i) \in \mathscr{F}$.

Рассмотрим $M = \{B_1 \times B_2 \times \ldots \times B_n : B_i \in \mathscr{B}(\mathbb{R})\}$. Тогда $\sigma(M) = \mathscr{B}(\mathbb{R}^n) \Rightarrow$ по критерию измеримости ξ — случайная величина. □

Утверждение 4.1.4. (б/д)

Любая непрерывная(кусочно-непрерывная) функция является борелевской.

Следствие 4.1.3. *Пусть ξ, η — случайные величины. Тогда*

$$\xi + \eta,\ \xi - \eta,\ \xi \cdot \eta,\ \frac{\xi}{\eta} I(\eta \neq 0),\ \xi^n, \xi^+ = max(\xi,\ 0),\ \xi^- = -min(\xi,\ 0),\ |\xi|$$

также являются случайными величинами.

Доказательство. Из следствия 4.1.2 — $(\xi,\ \eta)$ — случайный вектор. Из утверждения 4.1.4 функции $x + y,\ x - y,\ xy,\ x^n,\ |x|,\ x^+,\ x^-,\ \frac{x}{y} I(y \neq 0)$ — борелевские. По утверждению 4.1.3 — все доказано. □

Утверждение 4.1.5. Пусть $\xi_1, \xi_2, \ldots$ — случайные величины. Тогда следующие четыре функции являются расширенными случайными величинами (т.е. случайными величинами, принимающими значения в $\mathbb{R} \cup \{\pm\infty\}$):

$$\inf_n \xi_n;\ \sup_n \xi_n;\ \varliminf_{n \to +\infty} \xi_n;\ \varlimsup_{n \to +\infty} \xi_n.$$

Доказательство. Простое следствие того, что

$$\{\omega : \quad \sup \xi_n > x\} = \bigcup_n \{\omega :\ \xi_n > x\} \in \mathscr{F}$$
$$\{\omega : \quad \inf \xi_n < x\} = \bigcup_n \{\omega :\ \xi_n < x\} \in \mathscr{F}$$

и $\varlimsup\limits_{n \to +\infty} \xi_n = \inf\limits_n \sup\limits_{m \geq n} \xi_m$, $\varliminf\limits_{n \to +\infty} \xi_n = \sup\limits_n \inf\limits_{m \geq n} \xi_m$ □

Определение 4.2. Распределением случайного n-мерного вектора ξ называется такая функция $\mathsf{P}_\xi : \mathscr{B}(\mathbb{R}^n) \to [0,\ 1]$, что $\mathsf{P}_\xi(B) = \mathsf{P}(\xi \in B)$

4.2 Примеры

Пример 4.1. Рассмотрим $\Omega = \{\omega_1,\ \omega_2\}$, $\mathscr{F} = 2^\Omega$, $\mathsf{P}(\{\omega_1\}) = p \quad \mathsf{P}(\{\omega_2\}) = 1 - p =: q$. Тогда случайной величиной является функция $\xi : \Omega \to \{0,\ 1\}$, такая что $\xi(\omega_1) = 1$, $\xi(\omega_2) = 0$. При этом, $\mathsf{P}(\xi = 1) = p$, $\mathsf{P}(\xi = 0) = q$

Пример 4.2. Рассмотрим $(\Omega, \mathscr{F}, \mathsf{P})$, $\quad A \in \mathscr{F}$. Тогда случайной величиной является функция $\xi = I_A = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A \end{cases}$. При этом, $\mathsf{P}(\xi = 1) = \mathsf{P}(A)$, $\mathsf{P}(\xi = 0) = 1 - \mathsf{P}(A)$

Пример 4.3. Если $\mathscr{F} = 2^\Omega$, то любая функция является случайной величиной. Если $\exists A \in 2^\Omega \wedge A \notin \mathscr{F}$, то I_A не является случайной величиной.

4.3 Независимость случайных величин

Определение 4.3. Пусть ξ — n-мерный случайный вектор, η — k-мерный. Тогда они называются *независимыми* (Обозначение: $\xi \perp\!\!\!\perp \eta$) , если $\forall B_1 \in \mathscr{B}(\mathbb{R}^n),\ B_2 \in \mathscr{B}(\mathbb{R}^k)$: $\mathsf{P}(\xi \in B_1, \eta \in B_2) = \mathsf{P}(\xi \in B_1)\mathsf{P}(\eta \in B_2)$

Случайные вектора $\xi_1,\ \xi_2,\ \dots,\ \xi_n$ называются попарно независимыми, если $\forall i \neq j\ \xi_i \perp\!\!\!\perp \xi_j$

Если для любых борелевских множеств $B_1,\ B_2,\ \dots,\ B_n$ выполнено, что $\mathsf{P}(\xi_1 \in B_1,\ \dots,\ \xi_n \in B_n) = \mathsf{P}(\xi_1 \in B_1)\dots\mathsf{P}(\xi_n \in B_n)$, то случайные вектора называются независимыми в совокупности.

Замечание: Взяв в определении независимости в совокупности $B_i = \mathbb{R}^k$ можно показать, что аналогичное равенство следует для любого поднабора случайных величин.

$\{\xi_\alpha\ :\ \alpha \in A\}$ независимы в совокупности, если $\forall n \in \mathbb{N}\ \forall$ попарно различных $\alpha_1,\ \dots,\ \alpha_n \in A$: $\xi_{\alpha_1},\ \dots,\ \xi_{\alpha_n}$ независимы в совокупности.

Утверждение 4.3.1. Пусть $\xi_1, \dots, \xi_n$ — дискретные случайные величины, $X_1, \dots, X_n$ — множества их значений. Тогда $\xi_1, \dots, \xi_n$ независимы в совокупности $\iff \forall x_1 \in X_1, \dots, x_n \in X_n\ \mathsf{P}(\xi_1 = x_1, \dots, \xi_n = x_n) = \mathsf{P}(\xi_1 = x_1) \dots\ \mathsf{P}(\xi_n = x_n)$

Доказательство.

$\Rightarrow$:

Очевидно из определения.

$\Leftarrow$:

Пусть $B_i \in \mathscr{F}_{\xi_i}$. Положим $\{x^{(i)}\} := \xi_i^{-1}(B_i)$

$$\mathsf{P}(\xi_1 \in B_1,\ \dots,\ x_n \in B_n) = \mathsf{P}\Big(\bigcap_{j=1}^{n} \bigcup_{i=1}^{\infty} \{\xi_j = x_i^{(j)}\}\Big) = \mathsf{P}\Big(\bigcup_{i_1=1}^{\infty} \dots \bigcup_{i_n=1}^{\infty} \{\xi_1 = x_{i_1}^{(1)}\} \cap \dots \cap \{\xi_n = x_{i_n}^{(n)}\}\Big) = \lim_{\substack{k_1 \to \infty \\ \dots \\ k_n \to \infty}} \mathsf{P}\Big(\bigcup_{i_1=1}^{k_1} \dots \bigcup_{i_n=1}^{k_n} \{\xi_1 = x_{i_1}^{(1)}\} \cap \dots \cap \{\xi_n = x_{i_n}^{(n)}\}\Big) = \lim_{\substack{k_1 \to \infty \\ \dots \\ k_n \to \infty}} \Bigg[\sum_{i_1=1}^{k_1} \dots \sum_{i_n=1}^{k_n} \mathsf{P}(\xi_1 = x_{i_1}^{(1)}) \dots \mathsf{P}(\xi_n = x_{i_n}^{(n)})\Bigg] = \lim_{\substack{k_1 \to \infty \\ \dots \\ k_n \to \infty}} \Big[\sum_{i_1=1}^{k_1} P(\xi_1 = x_{i_1}^{(1)}) \dots \sum_{i_n=1}^{k_n} \mathsf{P}(\xi_n = x_{i_n}^{(n)})\Big] = \mathsf{P}(\xi_1 \in B_1) \dots \mathsf{P}(\xi_n \in B_n)$$

□

Определение 4.4. Пусть Ω — некоторое пространство. Система $\mathscr{P}$ подмножеств Ω называется *π-системой*, если она *замкнута* относительно взятия конечных пересечений: $\forall A_1, \dots, A_n \in \mathscr{P} : \bigcap\limits_{1 \le k \le n} A_k \in \mathscr{P},\ n \geq 1$.

Система $\mathscr{L}$ подмножеств Ω называется *λ-системой*, если

1. $\Omega \in \mathscr{L}$
2. $(A,\ B \in \mathscr{L}$ и $A \subseteq B) \Rightarrow (B \setminus A \in \mathscr{L}\ )$

3. $(A_n \in \mathscr{L},\ n \geq 1,$ и $A_n \uparrow A) \Rightarrow (A \in \mathscr{L})$

Система $\mathscr{D}$ подмножеств Ω, являющаяся *одновременно* π-системой и λ-системой, называется *π-λ-системой.*

Если $\mathscr{E}$ — некоторая система множеств, то через $\pi(\mathscr{E})$, $\lambda(\mathscr{E})$ и $d(\mathscr{E})$ будем обозначать соответственно наименьшие π-, λ- и π-λ-системы, содержащие $\mathscr{E}$.

Заметим, что каждая σ-алгебра является λ-системой. Обратное же неверно. Например, если $\Omega = \{1,\ 2,\ 3,\ 4\}$, то система

$$\mathscr{L} = \{\varnothing,\ \Omega,\ (1,\ 2),\ (1,\ 3),\ (1,\ 4),\ (2,\ 3),\ (2,\ 4),\ (3,\ 4)\}$$

является λ-системой, но не σ-алгеброй.

Теорема 4.2. *(О π-λ-системах, на лекции б/д)*

Верны следующие утверждения:

a) Всякая π-λ-система является σ-алгеброй

b) Если $\mathscr{E}$ — некоторая π-система множеств, то $\lambda(\mathscr{E}) = d(\mathscr{E}) = \sigma(\mathscr{E})$

Доказательство. Сначала заметим, что следующее определение λ-системы $\mathscr{L}$ эквивалентно данному выше.

1. $\Omega \in \mathscr{L}$
2. $A \in \mathscr{L} \Rightarrow \overline{A} \in \mathscr{L}$
3. если $A_n \in \mathscr{L}$, $n \geq 1$ и $\forall i \neq j\ A_i \cap A_j = \varnothing$, то $\bigcup A_n \in \mathscr{L}$

a) Система $\mathscr{E}$ содержит Ω т.к. это λ-система. Из определения λ-систем, $\mathscr{E}$ замкнуто относительно взятия дополнения множества, а из определения π-систем — конечного пересечения множеств, а значит, что $\mathscr{E}$ является алгеброй. Покажем, что это σ-алгебра. Для этого докажем, что система $\mathscr{E}$ замкнута относительно взятия счетного объединения множеств $B_1,\ B_2,\ \ldots$ из $\mathscr{E}$.

Положим $A_1 = B_1$ и $A_n = B_n \cap \overline{A_1} \cap \ldots \cap \overline{A}_{n-1}$. Тогда $\forall i \neq j\ A_i \cap A_j = \varnothing$, а значит $\bigcup A_n \in \mathscr{E}$. Но $\bigcup A_n = \bigcup B_n \Rightarrow \bigcup B_n \in \mathscr{E}$, откуда следует а).

b) Рассмотрим λ-систему $\lambda(\mathscr{E})$ и σ-алгебру $\sigma(\mathscr{E})$. Как отмечалось, всякая σ-алгебра является λ-системой. Поскольку $\sigma(\mathscr{E}) \supseteq \mathscr{E}$, то $\sigma(\mathscr{E}) = \lambda(\sigma(\mathscr{E})) \supseteq \lambda(\mathscr{E})$. Тем самым $\lambda(\mathscr{E}) \subseteq \sigma(\mathscr{E})$

Теперь, если показать, что система $\lambda(\mathscr{E})$ является π-системой, то по утверждению а) она является и σ-алгеброй, содержащей $\mathscr{E}$. А т.к. $\sigma(\mathscr{E})$ — минимальная σ-алгебра, то $\lambda(\mathscr{E}) = \sigma(\mathscr{E})$

Воспользуемся принципом подходящих множеств. Пусть

$$\mathscr{E}_1 = \{B \in \lambda(\mathscr{E}) :\ B \cap A \in \lambda(\mathscr{E}) \text{ для всех } A \in \mathscr{E}\}$$

Если $B \in \mathscr{E}$, то $B \cap A \in \mathscr{E}$ (т.к. $\mathscr{E}$ — π-система). Значит, $\mathscr{E} \subseteq \mathscr{E}_1$. Но система $\mathscr{E}_1$ есть λ-система (в силу своего определения). Поэтому $\lambda(\mathscr{E}) \subseteq \lambda(\mathscr{E}_1) = \mathscr{E}_1$. С другой стороны, по определнию системы $\mathscr{E}_1$ имеет место включение $\mathscr{E}_1 \subseteq \lambda(\mathscr{E})$.

Таким образом $\mathscr{E}_1 = \lambda(\mathscr{E})$.

Пусть теперь

$$\mathscr{E}_2 = \{B \in \lambda(\mathscr{E}) : \ B \cap A \in \lambda(\mathscr{E}) \text{ для всех } A \in \lambda(\mathscr{E})\} \qquad .$$

Как и $\mathscr{E}_1$, $\mathscr{E}_2$ является λ-системой.

Возьмем множество $B \in \mathscr{E}$. Тогда по опредлению системы $\mathscr{E}_1$ для всех $A \in \mathscr{E}_1 = \lambda(\mathscr{E})$ находим, что $B \cap A \in \lambda(\mathscr{E})$. Следовательно, из определния системы $\mathscr{E}_2$ видим, что $\mathscr{E} \subseteq \mathscr{E}_2$ и $\lambda(\mathscr{E}) \subseteq \lambda(\mathscr{E}_2) = \mathscr{E}_2$. Но $\lambda(\mathscr{E}) \supseteq \mathscr{E}_2$. Поэтому $\lambda(\mathscr{E}) = \mathscr{E}_2$, и, значит, для любых A и B из $\lambda(\mathscr{E})$ множество $A \cap B \in \lambda(\mathscr{E})$, т.е. система $\lambda(\mathscr{E})$ является π-системой. Значит, система $\lambda(\mathscr{E})$ является π-λ-системой (а, значит, $\lambda(\mathscr{E}) = d(\mathscr{E})$), а, как отмечено выше, отсюда следует, что $\lambda(\mathscr{E}) = \sigma(\mathscr{E})$. □

Теорема 4.3. *Критерий независимости.*

Пусть $\xi_1, \ldots, \xi_n$ — случайные величины, и $\xi = (\xi_1, \ldots, \xi_n)$. Пусть $F_\xi(x_1, \ldots, x_n) = \mathsf{P}(\xi_1 \le x_1, \ldots, \xi_n \le x_n)$ — функция распределения случайного вектора ξ, а $F_{\xi_i}(x)$ — функция распределения ξ_i.

Тогда для того, чтобы случайные величины $\xi_1, \ldots, \xi_n$ были независимы, необходимо и достаточно, чтобы для всех векторов $(x_1, \ldots, x_n) \in \mathbb{R}^n$

$$F_\xi(x_1, \ldots, x_n) = F_{\xi_1}(x_1) \ldots F_{\xi_n}(x_n)$$

Доказательство. Необходимость очевидна.

Для доказательства достаточности положим $M = \{(-\infty, x], : \ x \in \mathbb{R}\}$. Очевидно, что M это π-система. Докажем индукцией по k утверждение: $\forall k \in \{1, \ldots, n\}, \forall B_1, \ldots, B_k \in \mathscr{B}(\mathbb{R})$ и $\forall x_{k+1}, \ldots, x_n \in \mathbb{R}$ верно:

$$\mathsf{P}(\xi_1 \in B_1, \ldots, \xi_k \in B_k, \xi_{k+1} \le x_{k+1}, \ldots, \xi_n \le x_n) = \\ = \mathsf{P}(\xi_1 \in B_1) \ldots \mathsf{P}(\xi_k \in B_k) \mathsf{P}(\xi_{k+1} \le x_{k+1}) \ldots \mathsf{P}(\xi_n \le x_n)$$

База индукции:

Воспользуемся методом подходящих множеств.

Положим $\mathscr{E} = \{B_1 \in \mathscr{B}(\mathbb{R}) : \ \mathsf{P}(\xi_1 \in B_1, \ \xi_2 \le x_2, \ldots, \xi_n \le x_n) = \mathsf{P}(\xi_1 \in B_1)\mathsf{P}(\xi_2 \le x_2) \ldots \mathsf{P}(\xi_n \le x_n)\}$

Заметим следующее:

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

а значит, $M \subset \mathscr{E} \subseteq \mathscr{B}(\mathbb{R})$. Покажем, что $\mathscr{E}$ является λ-системой.

1. $\mathsf{P}(\xi_1 \in \mathbb{R}, \xi_2 \le x_2, \ldots, \xi_n \le x_n) = \lim\limits_{x_1 \to +\infty} \mathsf{P}(\xi_1 \le x_1, \ \xi_2 \le x_2, \ldots, \xi_n \le x_n) = \lim\limits_{x_1 \to +\infty} \mathsf{P}(\xi_1 \le x_1) \ldots \mathsf{P}(\xi_n \le x_n) = \mathsf{P}(\xi_1 \in \mathbb{R}) \ldots \mathsf{P}(\xi_n \le x_n) \Rightarrow \mathbb{R} \in \mathscr{E}$ (по теореме о непрерывности вероятностной меры в 0).

2. Рассмотрим $A,\ B \in \mathscr{E},\ A \subset B$.

$$\begin{aligned}
&\mathsf{P}(\xi_1 \in B \setminus A, \xi_2 \le x_2, \ldots, \xi_n \le x_n) = \\
&\qquad = \mathsf{P}(\xi_1 \in B, \xi_2 \le x_2, \ldots, \xi_n \le x_n) - \mathsf{P}(\xi_1 \in A, \xi_2 \le x_2, \ldots, \xi_n \le x_n) = \\
&\qquad = \mathsf{P}(\xi_1 \in B)\mathsf{P}(\xi_2 \le x_2)\ldots\mathsf{P}(\xi_n \le x_n) - \mathsf{P}(\xi_1 \in A)\mathsf{P}(\xi_2 \le x_2)\ldots\mathsf{P}(\xi_n \le x_n) = \\
&\qquad = (\mathsf{P}(\xi_1 \in B) - \mathsf{P}(\xi_1 \in A))\mathsf{P}(\xi_2 \le x_2)\ldots\mathsf{P}(\xi_n \le x_n) = \mathsf{P}(\xi_1 \in B \setminus A)\mathsf{P}(\xi_2 \le x_2)\ldots\mathsf{P}(\xi_n \le x_n)
\end{aligned}$$

Откуда следует, что $B \setminus A \in \mathscr{E}$

3. Пусть $A_n \uparrow A$, $\forall i\ A_i \in \mathscr{E}$. Тогда (по теореме о непрерывности вероятностной меры в 0).:

$$\begin{aligned}
\mathsf{P}(\xi_1 \in A, \xi_2 \le x_2, \ldots, \xi_n \le x_n) &= \lim_{k \to +\infty} \mathsf{P}(\xi_1 \in A_k, \xi_2 \le x_2, \ldots, \xi_n \le x_n) = \\
&= \lim_{k \to +\infty} \mathsf{P}(\xi_1 \in A_k)\mathsf{P}(\xi_2 \le x_2)\ldots\mathsf{P}(\xi_n \le x_n) = \\
&= \mathsf{P}(\xi_1 \in A)\mathsf{P}(\xi_2 \le x_2)\ldots\mathsf{P}(\xi_n \le x_n)
\end{aligned}$$

а значит $A \in \mathscr{E}$

$\Rightarrow \mathscr{B}(\mathbb{R}) \supseteq \mathscr{E} = \lambda(\mathscr{E}) \supseteq \lambda(M) = \sigma(M) = \mathscr{B}(\mathbb{R}) \Rightarrow \mathscr{E} = \mathscr{B}(\mathbb{R})$(по теореме 4.2). База индукции доказана.

Шаг индукции. Доказательство аналогично доказательству базы. □

Следствие 4.3.1. *Пусть $\{\xi_\alpha,\ \alpha \in A\}$ — набор независимых в совокупности случайных величин. Пусть $\alpha_1, \ldots, \alpha_{k_1}, \alpha_{k_1+1}, \ldots, \alpha_{k_1+k_2}, \ldots, \alpha_{k_1+\ldots+k_n}$ — различные индексы из A. Тогда случайные вектора $\eta_1 := (\xi_{\alpha_1}, \ldots, \xi_{\alpha_{k_1}}), \ldots, \eta_n := (\xi_{\alpha_{k_1+\ldots+k_{n-1}+1}}, \ldots, \xi_{\alpha_{k_1+\ldots+k_n}})$ — независимы в совокупности.*

Доказательство. По критерию независимости ($x_i \in \mathbb{R}^{k_i}$):

$$F_{(\eta_1, \ldots, \eta_n)}(x_1, \ldots x_n) = F_{\eta_1}(x_1)\ldots F_{\eta_n}(x_n) = F_{\xi_1}(x_1^1)\ldots F_{\xi_{k_1}}(x_1^{k_1})\ldots F_{\xi_{k_1+\ldots k_n}}(x_n^{k_n}),$$

а это так, т.к. критерий независимости верен для любого поднабора ξ_i □

Утверждение 4.3.2. Пусть $\xi_1, \ldots, \xi_n$ — независимые в совокупности случайные векторы размерности k_i, а $g_1, \ldots, g_n$ — борелевские функции, $g_i : \mathbb{R}^{k_i} \to \mathbb{R}^{m_i}$. Тогда случайные вектора $g_1(\xi_1), \ldots, g_n(\xi_n)$ — независимы в совокупности.

Доказательство. $\mathsf{P}(g_1(\xi_1) \in B_1, \ldots, g_n(\xi_n) \in B_n) = \mathsf{P}(\xi_1 \in g_1^{-1}(B_1), \ldots, \xi_n \in g_n^{-1}(B_n)) = \mathsf{P}(\xi_1 \in g_1^{-1}(B_1))\ldots\mathsf{P}(\xi_n \in g_n^{-1}(B_n)) = \mathsf{P}(g_1(\xi_1) \in B_1)\ldots\mathsf{P}(g_n(\xi_n) \in B_n)$ □

5 Математическое ожидание

5.1 Дискретный случай. Свойства

Определение 5.1. Пусть ξ — *дискретная* случайная величина на пространстве $(\Omega, \mathscr{F}, \mathsf{P})$. X — множество ее значений. Тогда *математическим ожиданием* ξ называется число, равное $\mathsf{E}\xi = \sum\limits_{x \in X} x\mathsf{P}(\xi = x)$, если этот ряд сходится абсолютно.

Утверждение 5.1.1. Пусть $A_1,\ A_2,\ \ldots$ — разбиение Ω, такое что $\forall i\ \xi|_{A_i} = const$. Пусть $\omega_i \in A_i$. Тогда $\mathsf{E}\xi = \sum\limits_i \xi(\omega_i)\mathsf{P}(A_i)$

Доказательство. $\mathsf{E}\xi = \sum\limits_{x \in X} x\mathsf{P}(\xi = x) = \sum\limits_{x \in X} x\mathsf{P}(\bigcup\limits_{i\,:\,\xi(\omega_i) = x} A_i) = \sum x \sum\limits_{i:\xi(\omega_i)=x} \mathsf{P}(A_i) = \sum \xi(\omega_i)\mathsf{P}(A_i)$ □

Замечание: В частном случае, когда $(\Omega,\ \mathscr{F},\ \mathsf{P})$ — дискретное вероятностное пространство, формула принимает вид $\mathsf{E}\xi = \sum\limits_{\omega \in \Omega} \xi(\omega)\mathsf{P}(\{\omega\})$. В случае классической вероятности, $\mathsf{E}\xi = \frac{1}{|\Omega|} \sum\limits_{\omega \in \Omega} \xi(\omega)$

Утверждение 5.1.2. Математическое ожидание **дискретной** случайной величины обладает следующими свойствами:

1. $\xi = c,\ c \in \mathbb{R}$. $\mathsf{E}c = c$
2. $\mathsf{E}I_A = \mathsf{P}(A)$
3. Если $\xi \geq 0$, то $\mathsf{E}\xi \geq 0$
4. Если $\mathsf{P}(\xi = 0) = 1$, то $\mathsf{E}\xi = 0$
5. (Линейность) $\xi,\ \eta$ — случайные величины, $a,\ b \in \mathbb{R}$. Тогда $\mathsf{E}(a\xi + b\eta) = a\mathsf{E}\xi + b\mathsf{E}\eta$
6. Пусть $\xi \geq \eta$. Тогда $\mathsf{E}\xi \geq \mathsf{E}\eta$
7. $|\mathsf{E}\xi| \leq \mathsf{E}|\xi|$
8. Если $\xi \perp\!\!\!\perp \eta$, то $\mathsf{E}\xi\eta = \mathsf{E}\xi\ \mathsf{E}\eta$ (обратное неверно)
9. (неравенство Коши-Буняковского) $(\mathsf{E}\xi\eta)^2 \leq (\mathsf{E}\xi)^2\ (\mathsf{E}\eta)^2$

Доказательство.

1. $\xi = c \Rightarrow \mathsf{P}(\xi = c) = 1,\ \mathsf{E}\xi = c\mathsf{P}(\xi = c) = c$
2. $\xi = I_A \Rightarrow \mathsf{P}(\xi = 1) = \mathsf{P}(A),\ \mathsf{P}(\xi = 0) = 1 - \mathsf{P}(A),\ \mathsf{E}\xi = 1 \cdot \mathsf{P}(A) + 0 \cdot (1 - \mathsf{P}(A)) = \mathsf{P}(A)$
3. $\xi \geq 0, \mathsf{P}(\xi = x) \geq 0 \Rightarrow$ каждое отдельное слагаемое в $\sum\limits_{x \in X} x\mathsf{P}(\xi = x)$ неотрицательно $\Rightarrow \mathsf{E}\xi \geq 0$
4. $\mathsf{P}(\xi = 0) = 1 \Rightarrow \mathsf{E}\xi = 0 \cdot \mathsf{P}(\xi = 0) + \sum\limits_{x \in X,\, x \neq 0} x\mathsf{P}(\xi = x) = 0$

5. $c \in \mathbb{R}$. $\mathsf{E}c\xi = \sum\limits_{x \in X} cx\mathsf{P}(\xi = x) = c\sum\limits_{x \in X} x\mathsf{P}(\xi = x) = c\mathsf{E}\xi$

   $\mathsf{E}(\xi + \eta) = \sum\limits_i (\xi(\omega_i) + \eta(\omega_i))\mathsf{P}(\{\omega_i\}) = \sum\limits_i \xi(\omega_i)\mathsf{P}(\{\omega_i\}) + \sum\limits_i \eta(\omega_i)\mathsf{P}(\{\omega_i\}) = \mathsf{E}\xi + \mathsf{E}\eta$

6. Следует из записи $\mathsf{E}\xi = \sum\limits_{\omega \in \Omega} \xi(\omega)\mathsf{P}(\{\omega\})$

7. $|\mathsf{E}\xi| = |\sum\limits_{x \in X} x\mathsf{P}(\xi = x)| \leq |\sum\limits_{x \in X} |x|\mathsf{P}(\xi = x)| = \sum\limits_{x \in |X|} |x|\mathsf{P}(|\xi| = x) = \mathsf{E}|\xi|$

8. $\mathsf{E}\xi\eta = \sum\limits_{k,\, n} (x_k y_n)\mathsf{P}(\xi = x_k,\, \eta = y_n) = \sum\limits_{k,\, n} (x_k y_n)\mathsf{P}(\xi = x_k)\mathsf{P}(\eta = y_n) = \sum\limits_k x_k\mathsf{P}(\xi = x_k)\sum\limits_n y_n\mathsf{P}(\eta = y_n) = \mathsf{E}\xi\mathsf{E}\eta$

9. Рассмотрим функцию $f(\lambda) = \mathsf{E}(\xi + \lambda\eta)^2 \geq 0$. По линейности мат.ожидания, $f(\lambda) = \mathsf{E}\xi^2 + 2\lambda\mathsf{E}\xi\eta + \lambda^2\mathsf{E}\eta^2$ — это всюду неотрицательный квадратный трехчлен относительно $\lambda \Rightarrow$ его дискриминант меньше либо равен нулю.

   $4(\mathsf{E}\xi\eta)^2 - 4\mathsf{E}\xi^2\mathsf{E}\eta^2 \leq 0 \Rightarrow$ требуемое.

□

Утверждение 5.1.3. Пусть ξ — случайная величина, X — ее множество значений, а φ — борелевская функция. Тогда $\mathsf{E}\varphi(\xi) = \sum\limits_{x \in X} \varphi(x)P(\xi = x)$

Доказательство. Пусть Y — множество значений $\varphi(\xi)$.

$\mathsf{E}\varphi(\xi) = \sum\limits_{y \in Y} y\mathsf{P}(\varphi(\xi) = y) = \sum\limits_{y \in Y} y \sum\limits_{x:\, \varphi(x) = y} \mathsf{P}(\xi = x) = \sum\limits_{y \in Y}\sum\limits_{x:\, \varphi(x) = y} \varphi(x)\mathsf{P}(\xi = x) = \sum\limits_{x \in X} \varphi(x)\mathsf{P}(\xi = x)$ □

Пример 5.1. $\xi \sim Bern(p)$.

$\mathsf{E}\xi = 0\mathsf{P}(\xi = 0) + 1\mathsf{P}(\xi = 1) = p$

Пример 5.2. $\xi \sim Bin(n,\, p)$.

$\mathsf{E}\xi = \mathsf{E}(\xi_1 + \ldots + \xi_n)$, где $\xi_i \sim Bern(p)$ — независимые с.в. Тогда $\mathsf{E}\xi = \sum \mathsf{E}\xi_i = np$

Пример 5.3. $\xi \sim U(\{1,\, \ldots,\, n\})$, $\quad \mathsf{P}(\{k\}) = \frac{1}{n}$

$\mathsf{E}\xi = \frac{1}{n}\sum\limits_{k=1}^{n} k = \frac{n+1}{2}$

Пример 5.4. $\xi \sim Pois(\lambda)$

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

5.2 Абсолютно непрерывный случай

Определение 5.2. Пусть ξ — абсолютно непрерывная случайная величина. F_ξ, p_ξ — ее функция распределения и плотность соответственно. Тогда ее математическим ожиданием называется величина $\mathsf{E}\xi = \int\limits_{-\infty}^{+\infty} xd(F_\xi(x)) = \int\limits_{-\infty}^{+\infty} xp_\xi(x)dx$

Пример 5.5. $\xi \sim U([a,\, b])$. $\quad p_\xi(x) = \frac{1}{b-a}I(x \in [a,\, b])$

$\mathsf{E}\xi = \frac{a+b}{2}$

Пример 5.6. $\xi \sim Exp(\lambda)$. $\quad p_\xi(x) = \lambda e^{-\lambda x} I(x>0)$

$\mathsf{E}\xi = \int\limits_{-\infty}^{+\infty} x\lambda e^{-\lambda x} I(x>0)dx = \int\limits_{0}^{+\infty} x\lambda e^{-\lambda x}dx = \frac{1}{\lambda}$

Пример 5.7. $\xi \sim N(a,\, \sigma^2)$. $\quad p_\xi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}$

$$\begin{aligned} \mathsf{E}\xi &= \int\limits_{-\infty}^{+\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \\ &= \int\limits_{-\infty}^{+\infty} \frac{x-a}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx + a \int\limits_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \\ &= \int\limits_{-\infty}^{+\infty} \frac{x}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx + a \cdot 1 = 0 + a = a \end{aligned}$$

т.к. интегрируемая функция нечетная и абсолютно интегрируема.

5.3 Математическое ожидание произвольной случайной величины

Определение 5.3. Пусть ξ — случайная величина на $(\Omega,\, \mathscr{F},\, \mathsf{P})$. Если существует интеграл Лебега по пространству Ω по вероятностной мере P, то он называется *математическим ожиданием* ξ

$$\mathsf{E}\xi = \int\limits_{\Omega} \xi(\omega)\mathsf{P}(d\omega) = \int\limits_{\Omega} \xi d\mathsf{P}_\xi$$

(интеграл Лебега по вероятностной мере будет определен далее)

Определение 5.4. Случайная величина ξ называется *простой*, если ее множество значений конечно. (частный случай дискретной с.в.)

Определение 5.5. Пусть ξ — с.в. Последовательность случайных величин $\{\xi_n\}$ *монотонно сходится* к ξ (обозначение: $\xi_n \uparrow \xi$), если функц. последовательность ξ_n поточечно сходится к ξ на Ω и $\forall n \in \mathbb{N} \forall \omega \in \Omega :\ \xi_n(\omega) \le \xi_{n+1}(\omega)$

Определение 5.6. Пусть ξ — случайная величина. Тогда $\xi^+ = max(\xi, 0)$, $\xi^- = -min(\xi, 0)$. Нетрудно заметить, что $\xi = \xi^+ - \xi^-$

Теорема 5.1. *О приближении простыми.*

1. *Пусть ξ — неотрицательная случайная величина. Тогда найдется последовательность простых P_ξ-измеримых случайных величин $\{\xi_n\}$, такая что $\xi_n \uparrow \xi$*

2. *Пусть ξ — произвольная случайная величина. Тогда найдется последовательность простых P_ξ-измеримых случайных величин, такая что $\xi_n \to \xi$ и $|\xi_n| \le |\xi_{n+1}|$*

Доказательство.

1. Пример такой последовательности:
$$\xi_n = \begin{cases} \frac{k-1}{2^n}, & \frac{k-1}{2^n} \le \xi < \frac{k}{2^n}, \quad k = 1, \ldots, n2^n \\ n, & \xi \ge n \end{cases}$$
Зафиксируем $\omega \in \{\xi \in \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right]\}$. Тогда для него
$$\xi_n(\omega) = \frac{k-1}{2^n},\ \xi_{n+1}(\omega) = \frac{k-1}{2^n} I\left(\xi(\omega) \in \left[\frac{k-1}{2^n}, \frac{2k-1}{2^n}\right)\right) + \frac{2k-1}{2^{n+1}} I\left(\xi(\omega) \in \left[\frac{2k-1}{2^{n+1}}, \frac{k}{2^n}\right)\right)$$
а значит $\xi_n \le \xi_{n+1}$. Сходимость же следует из того, что $\{\xi \le n\} \uparrow \Omega$ и на этом множестве $|\xi_n - \xi| \le \frac{1}{2^n} \to 0$.

2. Сначала заметим, что $\xi = \xi^+ - \xi^-$, $|\xi| = \xi^+ + \xi^-$, где ξ^+, ξ^- — неотрицательные случайные величины. Пусть $\eta_n \uparrow \xi^+$, $\zeta_n \uparrow \xi^-$. Положим $\xi_n = \eta_n - \zeta_n$. Тогда ξ_n сходится поточечно к ξ, а $|\xi_n| = \eta_n + \zeta_n \le \eta_{n+1} + \zeta_{n+1} = |\xi_{n+1}|$

□

Пусть ξ — неотрицательная случайная величина. Построим последовательность простых $\xi_n \uparrow \xi$. Т.к. $\xi_n \le \xi_{n+1}$, то $\mathsf{E}\xi_n \le \mathsf{E}\xi_{n+1}$, а значит существует $\lim\limits_{n\to\infty} \mathsf{E}\xi_n$

Определение 5.7. *Математическим ожиданием* неотрицательной случайной величины ξ, или ее *интегралом Лебега* называется величина
$$\mathsf{E}\xi = \lim_{n\to\infty} \mathsf{E}\xi_n$$

Лемма 5.1. *Пусть ξ_n, η — простые неотрицательные случайные величины, и $\xi_n \uparrow \xi \ge \eta$. Тогда:*
$$\lim_{n\to+\infty} \mathsf{E}\xi_n \ge \mathsf{E}\eta$$

Доказательство. Зафиксируем $\varepsilon > 0$. Положим
$$A_n := \{\omega \in \Omega : \xi_n \ge \eta - \varepsilon\} \in \mathscr{F}$$
Поскольку $\xi_n \uparrow \xi \ge \eta$, то $A_n \to \Omega \Rightarrow \mathsf{P}(A_n) \to 1$ и
$$\xi_n = \xi_n I_{A_n} + \xi_n I_{\overline{A_n}} \ge \xi_n I_{A_n} \ge (\eta - \varepsilon) I_{A_n}$$
Используя свойста мат.ожидания от простых случайных величин, находим, что
$$\begin{aligned} \mathsf{E}\xi_n \ge \mathsf{E}(\eta-\varepsilon)I_{A_n} = \mathsf{E}\eta I_{A_n} - \varepsilon \mathsf{P}(A_n) = \\ = \mathsf{E}\eta - \mathsf{E}\eta I_{\overline{A_n}} - \varepsilon\mathsf{P}(A_n) \ge \mathsf{E}\eta - C\mathsf{P}(\overline{A}_n) - \varepsilon \end{aligned}$$
где $C = \max \eta$. Переходя к пределу при $n \to \infty$, получаем требуемое неравенство. □

Утверждение 5.3.1. Определение математического ожидание корректно, т.е. не зависит от выбора аппроксимирующей последовательности.

Доказательство. Пусть $\xi_n \uparrow \xi$, $\eta_n \uparrow \xi$. Тогда по по лемме 5.1:

$$\lim_n \mathsf{E}\eta_n \geq \mathsf{E}\xi_n \Rightarrow \lim_n \mathsf{E}\eta_n \geq \lim_n \mathsf{E}\xi_n.$$

Абсолютно аналогично получаем, что $\lim_n \mathsf{E}\xi_n \geq \lim_n \mathsf{E}\eta_n \Rightarrow \lim_n \mathsf{E}\xi_n = \lim_n \mathsf{E}\eta_n$ □

Следствие 5.1.1. *Если $\xi \geq 0$ — случайная величина, то*

$$\mathsf{E}\xi = \sup_{\eta:\eta\leq\xi} \mathsf{E}\eta$$

где η — простые случайные величины.

Определение 5.8. Пусть ξ — произвольная случайная величина на $(\Omega,\ \mathscr{F},\ \mathsf{P})$. Тогда ее *математическим ожиданием* $\mathsf{E}\xi$, или *интегралом Лебега*, называется

1. $\mathsf{E}\xi^+ - \mathsf{E}\xi^-$, если $\mathsf{E}\xi^+$ и $\mathsf{E}\xi^-$ — конечны.

2. $+\infty$, если $\mathsf{E}\xi^+ = \infty$, $\mathsf{E}\xi^-$ — конечно.

3. $-\infty$, если $\mathsf{E}\xi^- = \infty$, $\mathsf{E}\xi^+$ — конечно.

4. неопределено, если $\mathsf{E}\xi^+$ и $\mathsf{E}\xi^-$ — бесконечны.

5.4 Свойства математического ожидания

Утверждение 5.4.1. Пусть $\xi \leq \eta$. $\mathsf{E}\xi$, $\mathsf{E}\eta$ — существуют. Тогда $\mathsf{E}\xi \leq \mathsf{E}\eta$.

Доказательство. Для простых доказано. Пусть $\eta \geq \xi \geq 0$. Тогда по следствию 5.1.1:

$$\mathsf{E}\xi = \sup_{\mu\leq\xi} \mathsf{E}\mu \leq \sup_{\mu\leq\eta} \mathsf{E}\mu = \mathsf{E}\eta.$$

что доказывает утверждение для неотрицательных случайных величин.

Пусть ξ, η — произвольные. Тогда $\xi = \xi^+ - \xi^-$, $\eta = \eta^+ - \eta^- \Rightarrow \xi^+ \leq \eta^+$, $\xi^- \geq \eta^-$ (т.к. $\eta \geq \xi$) $\Rightarrow$

$$\mathsf{E}\xi^+ \leq \mathsf{E}\eta^+,\ \mathsf{E}\xi^- \geq \mathsf{E}\eta^-$$

Разбор случаев, когда одно или оба математических ожидания бесконечны тривиален.

Ежели $|\mathsf{E}\xi| < \infty$, $|\mathsf{E}\eta| < \infty$, то:

$$\mathsf{E}\xi = \mathsf{E}\xi^+ - \mathsf{E}\xi^- \leq \mathsf{E}\eta^+ - \mathsf{E}\eta^- = \mathsf{E}\eta.$$

□

Утверждение 5.4.2. Если $\xi \geq 0$, то $\mathsf{E}\xi \geq 0$. Дополнительно, если $\mathsf{E}\xi = 0$, то $P(\xi = 0) = 1$.

Доказательство. Взяв $\eta = 0$ и применив свойство 5.4.1 получаем требуемое.

Для простых: $\mathsf{E}\xi = \sum_n c_n \mathsf{P}(A_n) = 0$ и $\forall i:\ c_i \geq 0 \Rightarrow \mathsf{P}(\xi = 0) = 1$

Пусть $\xi \geq 0$ — произвольная. Тогда существует последовательность простых $\xi_n \uparrow \xi$. Откуда:

$$0 \leq \mathsf{E}\xi_n \leq \mathsf{E}\xi = 0 \Rightarrow \mathsf{P}(\xi_n = 0) = 1$$

Т.к. $\xi_n \uparrow \xi$, то $\{\xi_n = 0\} \downarrow \{\xi = 0\} \Rightarrow \mathsf{P}(\xi = 0) = 1$ □

Утверждение 5.4.3. Пусть ξ случайная величина на $(\Omega, \mathscr{F}, \mathsf{P})$. Тогда, если существует $\mathsf{E}\xi$, то для любого события $A \in \mathscr{F}$ существует $\mathsf{E}\xi \cdot I_A$. Дополнительно, если $\mathsf{E}\xi$ конечно, то для любого события $A \in \mathscr{F}$ конечно $\mathsf{E}\xi \cdot I_A$.

Доказательство. $\mathsf{E}\xi$ существует $\iff \min(\mathsf{E}\xi^+, \mathsf{E}\xi^-) < \infty$. Тогда имеем:

$$\xi^- I_A \leq \xi^-, \ \xi^+ I_A \leq \xi^+ \Rightarrow \mathsf{E}(\xi^- I_A) \leq \mathsf{E}\xi^-, \ \mathsf{E}(\xi^+ I_A) \leq \mathsf{E}\xi^+ \Rightarrow \min(\mathsf{E}(\xi^+ I_A), \ \mathsf{E}(\xi^- I_A)) < \infty,$$

а значит $\mathsf{E}\xi I_A$ существует.

Если мат. ожидание ξ конечно, то конечно мат.ожидание ξ^+ и ξ^-. А значит:

$$\mathsf{E}(\xi I_A)^+ = \mathsf{E}\xi^+ I_A < \infty, \ \mathsf{E}(\xi I_A)^- = \mathsf{E}\xi^- I_A < \infty \Rightarrow |\mathsf{E}\xi I_A| < \infty$$

□

Утверждение 5.4.4. Если $\mathsf{P}(\xi = 0) = 1$, то $\mathsf{E}\xi = 0$

Доказательство. Для простых доказано. Пусть $\xi = \xi^+ - \xi^-$ — произвольная. Из условия следует, что $\mathsf{P}(\xi^+ = 0) = \mathsf{P}(\xi^- = 0) = 1$ Пусть $\xi_n^+ \uparrow \xi^+$, $\xi_n^- \uparrow \xi^-$ — последовательности неотрицательных простых. Тогда:

$$\mathsf{P}(\xi_n^+ = 0) = \mathsf{P}(\xi_n^- = 0) = 1 \Rightarrow \mathsf{E}\xi_n^+ = \mathsf{E}\xi_n^- = 0 \Rightarrow \mathsf{E}\xi^+ = \mathsf{E}\xi^- = 0 \Rightarrow \mathsf{E}\xi = 0$$

□

Утверждение 5.4.5. Если $\mathsf{E}\xi$ существует, то $|\mathsf{E}\xi| \leq \mathsf{E}|\xi|$

Доказательство. Для простых доказано. Для неотрицательных очевидно.

$|\xi| = \xi^+ + \xi^-$. Разберем случаи:

1. $\mathsf{E}\xi^- = \infty, \ \mathsf{E}\xi^+ < \infty$. Тогда $|\mathsf{E}\xi| = \mathsf{E}|\xi| = +\infty$
2. $\mathsf{E}\xi^+ = \infty, \ \mathsf{E}\xi^- < \infty$. Тогда аналогично $|\mathsf{E}\xi| = \mathsf{E}|\xi| = +\infty$
3. $\mathsf{E}\xi^+ < \infty, \ \mathsf{E}\xi^- < \infty$. Тогда $|\mathsf{E}\xi| = |\mathsf{E}\xi^+ - \mathsf{E}\xi^-| \leq |\mathsf{E}\xi^+| + |\mathsf{E}\xi^-| = \mathsf{E}\xi^+ + \mathsf{E}\xi^- = \mathsf{E}|\xi|$

□

Утверждение 5.4.6. Пусть $c \in \mathbb{R}$. Тогда если $\mathsf{E}\xi$ существует, то $\mathsf{E}c\xi = c\mathsf{E}\xi$

Доказательство. Для простых доказано.

Пусть $\xi \geq 0$. Рассмотрим последовательность простых $\xi_n \uparrow \xi$

1. $c \geq 0 \Rightarrow c\xi_n \uparrow c\xi$ — последовательность неотрицательных простых. Тогда для нее:
$$\mathsf{E}c\xi_n \to \mathsf{E}c\xi, \quad \mathsf{E}c\xi_n = c\mathsf{E}\xi_n \to c\mathsf{E}\xi \Rightarrow \mathsf{E}c\xi = c\mathsf{E}\xi$$
2. $c < 0$ Тогда $\mathsf{E}c\xi = -\mathsf{E}(c\xi)^- = -\mathsf{E}(-c\xi) = -(-c)\mathsf{E}\xi = c\mathsf{E}\xi$

Пусть теперь $\xi = \xi^+ - \xi^-$ — произвольная случайная величина. Случаи, когда одно из мат.ожиданий $\mathsf{E}\xi^+$, $\mathsf{E}\xi^-$ бесконечно разбирается очевидно. Предположим, что $\mathsf{E}\xi^+ < \infty$, $\mathsf{E}\xi^- < \infty$. Пусть $\xi_n^+ \uparrow \xi^+$, $\xi_n^- \uparrow \xi^-$.

1. $c \geq 0$. Тогда $c\xi_n^+ \uparrow c\xi$, $c\xi_n^- \uparrow c\xi^-$. По определению мат. ожидания
$$\mathsf{E}c\xi = \mathsf{E}(c\xi)^+ - \mathsf{E}(c\xi)^- = \lim_n(\mathsf{E}c\xi_n^+ - \mathsf{E}c\xi_n^-) = c\lim_n \mathsf{E}\xi_n = c\mathsf{E}\xi$$

2. $c < 0$. Тогда $-c\xi_n^+ \uparrow -c\xi^+$, $-c\xi_n^- \uparrow -c\xi^-$. По определению мат. ожидания
$$\mathsf{E}c\xi = \mathsf{E}(c\xi)^+ - \mathsf{E}(c\xi)^- = \lim_n \mathsf{E}(-c\xi_n^-) - \mathsf{E}(-c\xi_n^+) = \lim_n -c(\mathsf{E}\xi^- - \mathsf{E}\xi^+) = c\mathsf{E}\xi$$

□

Утверждение 5.4.7. Если ξ, η — неотрицательные случайные величины или такие, что $\mathsf{E}|\xi| < \infty$, $\mathsf{E}|\eta| < \infty$, то
$$\mathsf{E}(\xi+\eta) = \mathsf{E}\xi + \mathsf{E}\eta$$

Доказательство. Для простых уже доказано.

Пусть ξ, η — неотрицательные. Рассмотрим последовательности простых $\xi_n \uparrow \xi$, $\eta_n \uparrow \eta$, сходящиеся к ним. Тогда $\mathsf{E}\xi_n \to \mathsf{E}\xi$, $\mathsf{E}\eta_n \to \mathsf{E}\eta$. А значит
$$\left.\begin{aligned}&\mathsf{E}(\xi_n+\eta_n) = \mathsf{E}\xi_n + \mathsf{E}\eta_n \to \mathsf{E}\xi + \mathsf{E}\eta\\ &\mathsf{E}(\xi_n+\eta_n) \to \mathsf{E}(\xi+\eta)\end{aligned}\right\} \Rightarrow \mathsf{E}(\xi+\eta) = \mathsf{E}\xi + \mathsf{E}\eta$$
(б/д): Случай $\mathsf{E}|\xi| < \infty$, $\mathsf{E}|\eta| < \infty$ сводится к рассмотреному, если воспользоваться тем, что $\xi = \xi^+ - \xi^-$, $\eta = \eta^+ - \eta^-$ и тем, что $\xi^+ \leq |\xi|$, $\xi^- \leq |\xi|$. □

Утверждение 5.4.8. Пусть математические ожидания ξ, η конечны и для всех $A \in \mathscr{F}$: $\mathsf{E}\xi I_A \leq \mathsf{E}\eta I_A$. Тогда $\mathsf{P}(\xi \leq \eta) = 1$.

Доказательство. Действительно, пусть $A = \{\omega : \xi(\omega) > \eta(\omega)\} \in \mathscr{F}$. Тогда $\mathsf{E}\eta I_A \leq \mathsf{E}\xi I_A \leq \mathsf{E}\eta I_A \Rightarrow \mathsf{E}\eta I_A = \mathsf{E}\xi I_A$. По линейности математического ожидания $\mathsf{E}((\xi-\eta)I_A) = 0 \Rightarrow \mathsf{P}(A) = 0 \Rightarrow \mathsf{P}(\xi \leq \eta) = 1$. □

Утверждение 5.4.9. Если $\xi = \eta$ почти наверное (т.е. $\mathsf{P}(\ \{\xi \neq \eta\ \}\ ) = 0$) и $\mathsf{E}|\xi| < \infty$, то $\mathsf{E}|\eta| < \infty$ и $\mathsf{E}\xi = \mathsf{E}\eta$.

Доказательство. Пусть $N = \{\omega : \xi \neq \eta\}$. Тогда $\mathsf{P}(N) = 0$ и $\xi = \xi I_N + \xi I_{\overline{N}}$, $\eta = \eta I_N + \eta I_{\overline{N}}$. По линейности математического ожидания, $\mathsf{E}\xi = \mathsf{E}\xi I_N + \mathsf{E}\xi I_{\overline{N}} = \mathsf{E}\xi I_{\overline{N}} = \mathsf{E}\eta I_{\overline{N}} = \mathsf{E}\eta I_{\overline{N}} + 0 = \mathsf{E}\eta I_{\overline{N}} + \mathsf{E}\eta I_N = \mathsf{E}\eta$ □

Утверждение 5.4.10. (б/д)
Пусть математические ожидания ξ, η конечны, а $\xi \perp\!\!\!\perp \eta$. Тогда $\mathsf{E}\xi\eta = \mathsf{E}\xi\ \mathsf{E}\eta$

5.5 Основные теоремы о предельном переходе под знаком интеграла Лебега

Теорема 5.2. *(Лебега о монотонной сходимости)*

Пусть η, ξ, ξ_1, ξ_2, ... — случайные величины.

а) Если $\xi_n \geq \eta$ для всех $n \geq 1$, $\mathsf{E}\eta > -\infty$ и $\xi_n \uparrow \xi$, то
$$\mathsf{E}\xi_n \uparrow \mathsf{E}\xi$$

b) Если $\xi_n \le \eta$ для всех $n \ge 1$, $\mathsf{E}\eta < +\infty$ и $\xi_n \downarrow \xi$, то

$$\mathsf{E}\xi_n \downarrow \mathsf{E}\xi$$

Доказательство. (На лекции была б/д)

a) Предположим сначала, что $\eta \ge 0$. Определим для каждого $k \ge 1$ последовательность случайных величин $\{\xi_k^{(n)}\}_{n\ge1}$ как последовательность простых случайных величин, сходящуюся к ξ_k (т.е. $\xi_k^{(n)} \uparrow \xi_k$). Обозначим за $\zeta^{(n)} = \max\limits_{1\le k\le n} \xi_k^{(n)}$. Тогда

$$\zeta^{(n-1)} \le \zeta^{(n)} = \max_{1\le k\le n} \xi_k^{(n)} \le \max_{1\le k\le n} \xi_k = \xi_n.$$

Пусть $\zeta = \lim\limits_n \zeta^{(n)}$. Поскольку для $1 \le k \le n$

$$\xi_k^{(n)} \le \zeta^{(n)} \le \xi_n,$$

то, переходя к пределу при $n \to \infty$ получим, что для любого $k \ge 1$ верно

$$\xi_k \le \zeta \le \xi$$

а значит $\xi = \zeta$.

Случайные величины $\zeta^{(n)}$ простые и $\zeta^{(n)} \uparrow \zeta$, поэтому

$$\mathsf{E}\xi = \mathsf{E}\zeta = \lim \mathsf{E}\zeta^{(n)} \le \mathsf{E}\xi_n.$$

С другой стороны, поскольку $\xi_n \le \xi_{n+1} \le \xi$, то $\lim \mathsf{E}\xi_n \le \mathsf{E}\xi$. Тем самым, $\mathsf{E}\xi = \lim \mathsf{E}\xi_n$.

Пусть теперь η — произвольная случайная величина с $\mathsf{E}\eta > -\infty$.

Если $\mathsf{E}\eta = \infty$, то в силу $\mathsf{E}\xi_n = \mathsf{E}\xi = \infty$ и утверждение доказано. Пусть теперь $\mathsf{E}\eta < \infty$. Тогда $\mathsf{E}|\eta| < \infty$. Тогда $0 \le \xi_n - \eta \uparrow \xi - \eta$ поточечно на Ω. Значит, согласно доказанному ранее, $\mathsf{E}(\xi_n - \eta) \uparrow \mathsf{E}(\xi - \eta)$ и, по линейности

$$\mathsf{E}\xi_n - \mathsf{E}\eta \uparrow \mathsf{E}\xi - \mathsf{E}\eta \Rightarrow \mathsf{E}\xi_n \uparrow \mathsf{E}\xi$$

b) Доказательство следует из *a*), если вместо исходных величин рассмотреть величины со знаком минус.

□

Теорема 5.3. *(лемма Фату)*

Пусть η, ξ_1, ξ_2, … — случайные величины.

a) Если $\xi_n \ge \eta$ для всех n и $\mathsf{E}\eta > -\infty$, то

$$\mathsf{E} \varliminf_{n\to\infty} \xi_n \le \varliminf_{n\to\infty} \mathsf{E}\xi_n$$

b) Если $\xi_n \le \eta$ для всех n и $\mathsf{E}\eta < +\infty$, то

$$\varlimsup_{n\to\infty} \mathsf{E}\xi_n \le \mathsf{E} \varlimsup_{n\to\infty} \xi_n$$

c) Если $|\xi_n| \le \eta$ для всех n и $\mathsf{E}\eta < \infty$, то

$$\mathsf{E} \varliminf_{n\to\infty} \xi_n \le \varliminf_{n\to\infty} \mathsf{E}\xi_n \le \varlimsup_{n\to\infty} \mathsf{E}\xi_n \le \mathsf{E} \varlimsup_{n\to\infty} \xi_n$$

Доказательство. Положим $\zeta_n = \inf\limits_{m \geq n} \xi_m$. Тогда, очевидно, что $\zeta_n \leq \zeta_{n+1}$. Кроме этого,

$$\varliminf_{n\to\infty} \xi_n = \lim_{n\to\infty} \inf_{m\geq n} \xi_m = \lim_{n\to\infty} \zeta_n \Rightarrow \zeta_n \uparrow \varliminf_{n\to\infty} \xi_n$$

Т.к. $\zeta_n \geq \eta$ для всех n. Тогда из теоремы 5.2

$$\mathsf{E} \varliminf_{n\to\infty} \xi_n = \mathsf{E} \lim_{n\to\infty} \zeta_n = \lim_{n\to\infty} \mathsf{E}\zeta_n = \varliminf_{n\to\infty} \mathsf{E}\zeta_n \leq \varliminf_{n\to\infty} \mathsf{E}\xi_n$$

что и доказывает утверждение a). Второе утверждение следует из первого, если рассмотреть величины со знаком минус. Третье утверждение это следствие первых двух. □

Теорема 5.4. *(Лебега о мажорируемой сходимости)*

Пусть случайные величины $\eta, \xi, \xi_1, \xi_2, \dots$ *таковы, что* $|\xi_n| \leq \eta$, $\mathsf{E}|\eta| < \infty$ *и* $\xi_n \to \xi$ *почти наверное (т.е.* $P(\ \{\omega : \ \xi_n(\omega) \to \xi(\omega)\ \}\ ) = 1$*). Тогда* $\mathsf{E}|\xi| < \infty$ *и*

$$\mathsf{E}\xi_n \to \mathsf{E}\xi$$

$$\mathsf{E}|\xi_n - \xi| \to 0$$

Доказательство. По предположению, $\varliminf\limits_{n\to\infty} \xi_n = \varlimsup\limits_{n\to\infty} \xi_n = \xi$ (почти наверное). Поэтому, в силу утверждения 5.4.9 и леммы Фату:

$$\mathsf{E}\xi = \mathsf{E} \varliminf_{n\to\infty} \xi_n \leq \varliminf_{n\to\infty} \mathsf{E}\xi_n \leq \varlimsup_{n\to\infty} \mathsf{E}\xi_n \leq \mathsf{E} \varlimsup_{n\to\infty} \xi_n = \mathsf{E}\xi$$

что и доказывает первое утверждение. Ясно также, что $|\xi| \leq \eta \Rightarrow \mathsf{E}|\xi| < \infty$.

Далее заметим, что $|\xi - \xi_n| \leq |\xi| + |\xi_n| \leq 2\eta$ и $\varlimsup\limits_{n\to\infty} |\xi - \xi_n| = 0$. В силу линейности математического ожидания:

$$0 = \lim_{n\to\infty} |\mathsf{E}\xi - \mathsf{E}\xi_n| = \lim_{n\to\infty} |\mathsf{E}(\xi - \xi_n)| \leq \lim_{n\to\infty} \mathsf{E}|\xi - \xi_n|.$$

Тогда, применяя лемму Фату, имеем:

$$\varlimsup_{n\to\infty} \mathsf{E}|\xi - \xi_n| \leq \mathsf{E} \varlimsup_{n\to\infty} |\xi - \xi_n| = 0 \Rightarrow \lim_{n\to\infty} \mathsf{E}|\xi - \xi_n| = 0,$$

что завершает доказательство теоремы. □

Теорема 5.5. *(теорема о замене переменных под знаком интеграла Лебега)*

Пусть ξ *— случайная величина с распределением вероятностей* P_ξ, g *— борелевская функция и существует* $\mathsf{E}g(\xi)$*. Тогда* $\mathsf{E}g(\xi) = \int\limits_\Omega g(\xi(\omega))P(d\omega) = \int\limits_{\mathbb{R}} g(x)P_\xi(dx)$

Доказательство.

1. для индикаторов $g(x) = I(x \in B)$

   $$\mathsf{E}g(\xi) = \mathsf{P}(\xi \in B) = \int\limits_B \mathsf{P}_\xi(dx) = \int\limits_{\mathbb{R}} I(x \in B)\mathsf{P}_\xi(dx) = \int\limits_{\mathbb{R}} g(x)\mathsf{P}_\xi(dx)$$

2. для простых $g(x) = \sum\limits_{i=0}^n c_i I(x \in B_i)$, $B_1 \sqcup \ldots \sqcup B_n = \mathbb{R}$, $B_i \in \mathscr{B}(\mathbb{R})$ имеем (по линейности мат.ожидания):

   $$\mathsf{E}g(\xi) = \mathsf{E}(\sum_{i=0}^n c_i I(x \in B_i)) = \sum_{i=0}^n c_i \mathsf{E} I(x \in B_i) =$$
   $$= \sum_{i=0}^n c_i \int\limits_{\mathbb{R}} I(x \in B_i)\mathsf{P}_\xi(dx) = \int\limits_{\mathbb{R}} \sum_{i=0}^n c_i I(x \in B_i)\mathsf{P}_\xi(dx) = \int\limits_{\mathbb{R}} g(x)\mathsf{P}_\xi(dx)$$

3. Пусть g — произвольная неотрицательная. Тогда существует последовательность простых борелевских функций $g_n \uparrow g$. По пункту 2)$\mathsf{E}g_n(\xi) = \int\limits_{\mathbb{R}} g_n(x)\mathsf{P}_\xi(dx)$. Т.к. $g_n(\xi)$ неотрицательны, то по теореме Лебега о монотонной сходимости:
$$\lim \mathsf{E}g_n(\xi) = \mathsf{E}g(\xi) = \lim_{n\to\infty} \int\limits_{\mathbb{R}} g_n(x)\mathsf{P}_\xi(dx) = \int\limits_{\mathbb{R}} g(x)\mathsf{P}_\xi(dx)$$

4. g — произвольная борелевская. Тогда $g = g^+ - g^-$, где g^+ и g^- — неотрицательные, и

$$\mathsf{E}g(\xi) = \mathsf{E}g^+(\xi) - \mathsf{E}g^-(\xi) = \int\limits_{\mathbb{R}} g^+(x)\mathsf{P}_\xi(dx) - \int\limits_{\mathbb{R}} g^-(x)\mathsf{P}_\xi(dx) =$$
$$= \int\limits_{\mathbb{R}} (g^+(x) - g^-(x))\mathsf{P}_\xi(dx) = \int\limits_{\mathbb{R}} g(x)\mathsf{P}_\xi(dx)$$

Замечание: В случае, когда ξ имеет абсолютно-непрерывное распределение, $\mathsf{P}_\xi(dx) = p_\xi(x)dx$. □

Пример 5.8. Пусть ξ имеет нормальное распределение с параметрами (0, 1). Найдем математическое ожидание величины ξ^2.

По теореме о замене переменных под знаком интеграла Лебега:

$$\mathsf{E}\xi^2 = \int\limits_{\mathbb{R}} x^2\mathsf{P}_\xi(dx) = \int\limits_{\mathbb{R}} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx =$$
$$= -\int\limits_{\mathbb{R}} x \frac{1}{\sqrt{2\pi}} d(e^{-\frac{x^2}{2}}) = \int\limits_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$$

Пример 5.9. Пусть ξ имеет экспоненциальное распределение с параметром λ. Найдем математическое ожидание случайной величины e^ξ:

$$\mathsf{E}e^\xi = \int\limits_{\mathbb{R}} e^x \mathsf{P}_\xi(dx) = \int\limits_{\mathbb{R}} e^x(\lambda e^{-\lambda x}) dx I(x \geq 0) =$$
$$= \int\limits_0^{+\infty} \lambda e^{x(1-\lambda)} dx = \frac{\lambda e^{x(1-\lambda)}}{1-\lambda}\Big|_0^\infty = \begin{cases} \infty, & \lambda \leq 1 \\ \frac{\lambda}{\lambda - 1}, & \lambda > 1 \end{cases}$$

6 Дисперсия и ковариация

Определение 6.1. Пусть ξ — случайная величина на пространстве $(\Omega,\ \mathscr{F},\ \mathsf{P})$. Тогда ее *дисперсией* называется величина $\mathsf{D}\xi = \mathsf{E}(\xi - \mathsf{E}\xi)^2$.

Определение 6.2. Пусть $\xi,\ \eta$ — две случайных величины на пространстве $(\Omega,\ \mathscr{F},\ \mathsf{P})$. Тогда их *ковариацией* называется величина $\mathrm{cov}(\xi,\ \eta) = \mathsf{E}(\xi - \mathsf{E}\xi)(\eta - \mathsf{E}\eta)$.

Определение 6.3. Пусть $\xi,\ \eta$ — две случайных величины на пространстве $(\Omega,\ \mathscr{F},\ \mathsf{P})$. Тогда их *коэффицентом корреляции* называется $\rho(\xi,\ \eta) = \dfrac{\mathrm{cov}(\xi,\ \eta)}{\sqrt{\mathsf{D}\xi\,\mathsf{D}\eta}}$.

Установим свойства дисперсии и ковариации:

Утверждение 6.1.1. $\mathsf{D}\xi \geq 0$ — очевидно из определения.

Утверждение 6.1.2. $\mathsf{D}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2$.

Доказательство.

$$\mathsf{D}\xi = \mathsf{E}(\xi - \mathsf{E}\xi)^2 = \mathsf{E}(\xi^2 - 2\xi\mathsf{E}\xi + (\mathsf{E}\xi)^2) = \mathsf{E}\xi^2 - 2\mathsf{E}\xi\,\mathsf{E}\xi + \mathsf{E}(\mathsf{E}\xi)^2$$
$$= \mathsf{E}\xi^2 - 2(\mathsf{E}\xi)^2 + (\mathsf{E}\xi)^2 = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2$$

□

Утверждение 6.1.3. $\mathrm{cov}(\xi,\ \eta) = \mathsf{E}\xi\eta - \mathsf{E}\xi\,\mathsf{E}\eta$

Доказательство. $\mathrm{cov}(\xi,\ \eta) = \mathsf{E}(\xi - \mathsf{E}\xi)(\eta - \mathsf{E}\eta) = \mathsf{E}(\xi\eta - \eta\mathsf{E}\xi - \xi\mathsf{E}\eta + \mathsf{E}\eta\mathsf{E}\xi) = \mathsf{E}\xi\eta - \mathsf{E}\xi\mathsf{E}\eta$ □

Утверждение 6.1.4. Если $\xi \perp\!\!\!\perp \eta$, то $\mathrm{cov}(\xi,\ \eta) = 0$ (обратное неверно).

Утверждение 6.1.5. $\mathsf{D}\xi = \mathrm{cov}(\xi,\ \xi)$

Утверждение 6.1.6. Ковариация — билинейна ($\mathrm{cov}(a\xi_1 + b\xi_2,\ \eta) = a\mathrm{cov}(\xi_1,\ \eta) + b\mathrm{cov}(\xi_2,\ \eta)$)

Доказательство. Симметричность следует из определения.

$$\mathrm{cov}(a\xi_1 + b\xi_2,\ \eta) = \mathsf{E}((a\xi_1 + b\xi_2)\eta) - \mathsf{E}\eta\mathsf{E}(a\xi_1 + b\xi_2) =$$
$$= a\mathsf{E}\xi_1\eta + b\mathsf{E}\xi_2\eta - a\mathsf{E}\xi_1\mathsf{E}\eta - b\mathsf{E}\xi_2\mathsf{E}\eta = a\mathrm{cov}(\xi_1,\ \eta) + b\mathrm{cov}(\xi_2,\ \eta)$$

□

Утверждение 6.1.7. $\mathsf{D}(\xi_1 + \ldots + \xi_n) = \sum\limits_{i=1}^{n} \mathsf{D}\xi_i + 2\sum\limits_{1 \leq i < j \leq n} \mathrm{cov}(\xi_i,\ \xi_j)$

Доказательство.

$$\mathsf{D}(\xi_1 + \ldots + \xi_n) = \mathrm{cov}((\xi_1 + \ldots + \xi_n),\ (\xi_1 + \ldots + \xi_n)) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{cov}(\xi_i,\ \xi_j) = \sum_{i=1}^{n} \mathsf{D}\xi_i + 2\sum_{1 \leq i < j \leq n} \mathrm{cov}(\xi_i,\ \xi_j)$$

□

Утверждение 6.1.8. Если $\xi_1, \ldots, \xi_n$ — попарно независимы, то $\mathsf{D}(\xi_1 + \ldots + \xi_n) = \sum\limits_{i=1}^{n} \mathsf{D}\xi_i$.
Замечание: условие попарной независимости сильное. Достаточным условием является попарная некоррелированность.

7 Неравенства в теории вероятностей

Утверждение 7.1.1. Неравенство Коши-Буняковского
Пусть ξ, η — случайные величины на $(\Omega, \mathscr{F}, \mathsf{P})$, такие, что $\mathsf{E}\xi^2 < +\infty$, $\mathsf{E}\eta^2 < +\infty$. Тогда $\mathsf{E}|\xi\eta| < \infty$ и

$$(\mathsf{E}|\xi\eta|)^2 \leq \mathsf{E}\xi^2 \cdot \mathsf{E}\eta^2$$

Доказательство. Сначала заметим, что если $\mathsf{P}(\xi = 0) = 1$, то $\mathsf{E}\xi = 0$ и $\mathsf{E}|\xi\eta| = 0$. Аналогично с η.

Пусть теперь $\mathsf{P}(\xi = 0) \neq 1; \mathsf{P}(\eta = 0) \neq 1$. Рассмотрим случайные величины $\widetilde{\xi} := \frac{|\xi|}{\sqrt{\mathsf{E}\xi^2}}, \widetilde{\eta} := \frac{|\eta|}{\sqrt{\mathsf{E}\eta^2}}$. Тогда

$$\frac{\xi^2}{\mathsf{E}\xi^2} + \frac{\eta^2}{\mathsf{E}\eta^2} = \widetilde{\xi}^2 + \widetilde{\eta}^2 \geq 2\widetilde{\xi}\widetilde{\eta} \Rightarrow$$
$$\Rightarrow 2 \geq 2\mathsf{E}\widetilde{\xi}\widetilde{\eta} = 2\frac{\mathsf{E}|\xi\eta|}{\sqrt{\mathsf{E}\xi^2\mathsf{E}\eta^2}}$$

□

Утверждение 7.1.2. Неравенство Йенсена
Пусть ξ — случайная величина на $(\Omega, \mathscr{F}, \mathsf{P})$, g — выпуклая вниз функция и $\mathsf{E}g(\xi)$ существует, а $|\mathsf{E}\xi| < \infty$. Тогда

$$g(\mathsf{E}\xi) \leq \mathsf{E}g(\xi)$$

Доказательство. В силу выпуклости g имеем

$$\forall x_0 \in \mathbb{R}\ \exists \lambda(x_0)\ \forall x \in \mathbb{R} : g(x) - g(x_0) \geq \lambda(x_0)(x - x_0)$$
$$g(\xi) \geq g(\mathsf{E}\xi) + \lambda(\mathsf{E}\xi)(\xi - \mathsf{E}\xi)$$
$$\mathsf{E}g(\xi) \geq \mathsf{E}(g(\mathsf{E}\xi) + \lambda(\mathsf{E}\xi)(\xi - \mathsf{E}\xi)) \Rightarrow$$
$$\Rightarrow \mathsf{E}g(\xi) \geq g(\mathsf{E}\xi)$$

□

Утверждение 7.1.3. Неравенство Маркова

Пусть ξ — случайная величина на $(\Omega,\ \mathscr{F},\ \mathsf{P})$, $\mathsf{E}|\xi| < \infty$ и $a > 0$. Тогда

$$\mathsf{P}(|\xi| \geq a) \leq \frac{\mathsf{E}|\xi|}{a}$$

Доказательство. Оценим математическое ожидание $|\xi|$:

$$\mathsf{E}|\xi| = \mathsf{E}|\xi|I(|\xi| \geq a) + \mathsf{E}|\xi|I(|\xi| < a) \geq \mathsf{E}|\xi|I(|\xi| \geq a) \geq \mathsf{E}aI(|\xi| \geq a) = a\mathsf{P}(|\xi| \geq a)$$

□

Утверждение 7.1.4. Неравенство Чебышёва

Пусть ξ — случайная величина на $(\Omega,\ \mathscr{F},\ \mathsf{P})$, $\mathsf{E}\xi^2 < \infty$. Тогда $\forall \varepsilon > 0$

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

Доказательство. Следствие неравенства Маркова для случайной величины $\eta := |\xi - \mathsf{E}\xi|^2$ и $a = \varepsilon^2$ □

8 Виды сходимости случайных величин

8.1 Виды сходимости. Взаимосвязть между ними

Пусть дано вероятностное пространство $(\Omega, \mathscr{F}, \mathsf{P})$ и случайные величины $\xi, \xi_1, \xi_2, \ldots$ на нем.

Определение 8.1. Последовательность $\{\xi_n\}$ сходится *поточечно* к ξ, если для любого $\omega \in \Omega$ числовая последовательность $\{\xi_n(\omega)\}$ сходится к $\xi(\omega)$.

Определение 8.2. Событие $A \in \mathscr{F}$ выполнено *почти наверное*, если $\mathsf{P}(A) = 1$

Определение 8.3. Последовательность $\{\xi_n\}$ сходится к ξ *почти наверное*, если событие $\{\xi_n \to \xi\}$ выполнено почти наверное. (т.е. $\mathsf{P}(\{\xi_n \to \xi\}) = 1$).
Обозначение: $\xi_n \overset{\text{п.н.}}{\to} \xi$.

Определение 8.4. Последовательность $\{\xi_n\}$ сходится к ξ *по вероятнсоти*, если выполнено:

$$\forall \varepsilon > 0 : \ \mathsf{P}(|\xi_n - \xi| > \varepsilon) \to 0$$

Обозначение: $\xi_n \overset{\mathsf{P}}{\to} \xi$

Определение 8.5. Последовательность $\{\xi_n\}$ сходится к ξ в L^p, $(p > 0)$, если $\mathsf{E}|\xi_n - \xi|^p \to 0$ при $n \to \infty$.
Обозначение: $\xi_n \overset{L^p}{\to} \xi$.

Определение 8.6. Последовательность $\{\xi_n\}$ *слабо сходится* к ξ (или сходится *по распределению*), если для любой ограниченной непрерывной функции $f : \mathbb{R} \to \mathbb{R}$ верно:

$$\mathsf{E}f(\xi_n) \to \mathsf{E}f(\xi)$$

Обозначение: $\xi_n \overset{d}{\to} \xi$

Хотя и кажется, что определения сходимостей (быть может, за исключением сходимости по распределнию) эквивалентны, это далеко не так. Для этого рассмотрим следующий пример:

Пример 8.1. $\Omega = [0, 1]$, $\mathscr{F} = \mathscr{B}([0, 1])$, P — равномерное распределение. $\xi_n(\omega) = \omega^n$. Тогда последовательность ξ_n сходится поточечно к $I(\omega = 1)$, но

$$\xi_n \overset{\text{п.н.}}{\to} I(\omega = 1)$$
$$\xi_n \overset{\text{п.н.}}{\to} 0 \text{ т.к. } \mathsf{P}(\xi_n \to 0) = \mathsf{P}([0, 1)) = 1$$
$$\xi_n \overset{\text{п.н.}}{\to} I(\omega \in \mathbb{Q}) \text{ т.к. } \mathsf{P}(\mathbb{Q}) = 0$$

Лемма 8.1. *(Критерий сходимости почти наверное)*

Пусть $\xi, \xi_1, \xi_2, \ldots$ *– случайные величины на* $(\Omega, \mathscr{F}, \mathsf{P})$. *Тогда*

$$\xi_n \overset{п.н.}{\to} \xi \iff \forall \varepsilon > 0 : \ \mathsf{P}(\sup_{k \geq n} |\xi_k - \xi| > \varepsilon) \to 0 \ \textit{при } n \to \infty.$$

Доказательство. Для доказательства леммы рассмотрим следующие события:

$$A_k^\varepsilon = \{|\xi_k - \xi| > \varepsilon\}$$
$$A^\varepsilon = \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k^\varepsilon$$

Заметим тогда, что

$$\{\sup_{k\geq n}|\xi_k-\xi|>\varepsilon\}=\{\exists k\geq n:\ |\xi_k-\xi|>\varepsilon\}=\bigcup_{k\geq n}A_k^{\varepsilon}$$

$$\{\lim\xi_n\neq\xi\}=\{\exists\varepsilon>0:\ \forall n\,\exists k\geq n:\ |\xi_k-\xi|>\varepsilon\}=\bigcup_{m=1}^{\infty}\bigcap_{n=1}^{\infty}\bigcup_{k\geq n}A_k^{\varepsilon=\frac{1}{m}}=\bigcup_{m=1}^{\infty}A^{\frac{1}{m}}$$

Поэтому, имеем

$$\begin{aligned}\xi_n\overset{\text{п.н.}}{\to}\xi\Leftrightarrow \mathsf{P}(\xi_n\not\to\xi)=0 &\Leftrightarrow \mathsf{P}(\bigcup_{m=1}^{\infty}A^{\frac{1}{m}})=0\\ &\Leftrightarrow \forall m\ \mathsf{P}(A^{\frac{1}{m}})=0\\ &\Leftrightarrow \forall\varepsilon\ \mathsf{P}(A^{\varepsilon})=0\\ &\Leftrightarrow \mathsf{P}(\bigcup_{k\geq n}A_k^{\varepsilon})\to 0 \text{ (т.к. } \bigcup_{k\geq n}A_k^{\varepsilon}\downarrow A^{\varepsilon} \text{ и по теореме о непрерывности вероятностой меры)}\\ &\Leftrightarrow \mathsf{P}(\sup_{k\geq n}|\xi_k-\xi|>\varepsilon)\to 0\end{aligned}$$

□

Теорема 8.1. *(взаимосвязь между различными видами сходимости случайных величин)*

Пусть $\xi,\ \xi_1,\ \xi_2,\ldots$ *— случайные величины на* $(\Omega,\ \mathscr{F},\ \mathsf{P})$. *Тогда верны следующие импликации:*

1. $\xi_n\overset{\text{п.н.}}{\to}\xi\Rightarrow\xi_n\overset{\mathsf{P}}{\to}\xi$
2. $\xi_n\overset{L^p}{\to}\xi\Rightarrow\xi_n\overset{\mathsf{P}}{\to}\xi$
3. $\xi_n\overset{\mathsf{P}}{\to}\xi\Rightarrow\xi_n\overset{d}{\to}\xi$

Доказательство. 1. Т.к. $\sup_{k\geq n}|\xi_k-\xi|<\varepsilon\Rightarrow\forall k\geq n:\ |\xi_k-\xi|<\varepsilon$, то по критерию сходимости почти наверное получаем требуемое.

2. $\mathsf{P}(|\xi_n-\xi|>\varepsilon)=\mathsf{P}(|\xi_n-\xi|^p>\varepsilon^p)\leq\dfrac{\mathsf{E}|\xi_n-\xi|^p}{\varepsilon^p}\to 0$ по неравенству Маркова.

3. Пусть f — ограниченная и непрерывная функция, $|f|\leq C$. Зафиксируем $\varepsilon>0$. Т.к. $\mathsf{P}(|\xi|=\infty)=0$, то $\exists N\ \exists\delta$:

   1)$\mathsf{P}(|\xi|>N)\leq\dfrac{\varepsilon}{6C}$ так как $\mathsf{P}(\xi=\infty)=0$

   2)$\mathsf{P}(|\xi_n-\xi|>\delta)\leq\dfrac{\varepsilon}{6C}$ из сходимости по вероятности (при достаточно больших n).

   3)$\forall x\ \forall y\ |x|<N,\ |x-y|<\delta\ |f(x)-f(y)|\leq\dfrac{\varepsilon}{3}$ т.к. f равномерно непрерывна на отрезке $[-N,\ N]$

   Рассмотрим следующие события:

$$A_1=\{|\xi_n-\xi|\leq\delta\}\cap\{|\xi|<N\}$$

$$A_2=\{|\xi_n-\xi|\leq\delta\}\cap\{|\xi|\geq N\}$$

$$A_3=\{|\xi_n-\xi|>\delta\}$$

Очевидно, что эти события образуют разбиение $\Omega = A_1 \sqcup A_2 \sqcup A_3$. Оценим $|\mathsf{E}f(\xi_n) - \mathsf{E}f(\xi)|$:

$$|\mathsf{E}f(\xi_n) - \mathsf{E}f(\xi)| = |\mathsf{E}(f(\xi_n) - f(\xi)| \leq \mathsf{E}|f(\xi_n) - f(\xi)| = \mathsf{E}|f(\xi_n) - f(\xi)|(I_{A_1} + I_{A_2} + I_{A_3}) \leq$$
$$\leq \frac{\varepsilon}{3}\mathsf{P}(A_1) + 2C(\mathsf{P}(A_2) + \mathsf{P}(A_3)) \leq \frac{\varepsilon}{3} + 2C(\frac{\varepsilon}{6C} + \frac{\varepsilon}{6C}) = \varepsilon$$

откуда следует, что $|\mathsf{E}f(\xi_n) - \mathsf{E}f(\xi)| \to 0 \Rightarrow \xi_n \xrightarrow{d} \xi$.

□

Из теоремы следует, что между видами сходимости существует следующая связь: поточечная сходимость $\Rightarrow$ сходимость почти наверное $\Rightarrow$ сходимость по вероятности $\Rightarrow$ сходимость по распределнию. Кроме того, сходимость по вероятности так же следует из сходимости в смысле L^p, однако для всех остальных случаев существуют контр-примеры.

Пример 8.2. $\xi_n \xrightarrow{d} \xi$, но $\xi_n \overset{\mathsf{P}}{\not\to} \xi$.

Пусть $\Omega = \{\omega_1,\ \omega_2\}$, $\mathsf{P}(\{\omega_i\}) = \frac{1}{2}$. Определим для любого n $\xi_n(\omega_1) = 1$, $\xi_n(\omega_2) = -1$. Положим $\xi = -\xi_n$. Тогда:

$$\mathsf{E}f(\xi_n) = \tfrac{f(1)+f(-1)}{2} = \mathsf{E}f(\xi),$$

но $\forall n$ $|\xi_n - \xi| = 2 \Rightarrow \xi_n \overset{\mathsf{P}}{\not\to} \xi$.

Пример 8.3. $\xi_n \xrightarrow{\mathsf{P}} \xi$, но $\xi_n \overset{\text{п.н.}}{\not\to} \xi$.

Положим $\xi_{2^k} = I([0,\ \frac{1}{2^k}])$, $\xi_{2^k+p} = I([\frac{p}{2^k},\ \frac{p+1}{2^k}])$, $1 \leq p < 2^k$. Тогда $\xi_n \xrightarrow{\mathsf{P}} 0$, т.к. $\mathsf{P}(\xi_n > 0) \leq$ длина отрезка в индикаторе $\leq \frac{2}{n} \to 0$, но $\xi_n \overset{\text{п.н.}}{\not\to} 0$ т.к. $\forall\omega\ \exists$ бесконечно много n, таких что $\xi_n(\omega) = 1$.

Пример 8.4. $\xi_n \xrightarrow{L^p} \xi$, но $\xi_n \overset{\text{п.н.}}{\not\to} \xi$. Подходит предыдущий пример.

Пример 8.5. $\xi_n \xrightarrow{\text{п.н.}} \xi$, но $\xi_n \overset{L^p}{\not\to} \xi$ и $\xi_n \xrightarrow{\mathsf{P}} \xi$ но $\xi_n \overset{L^p}{\not\to} \xi$.

$\Omega = [0, 1]$, $\mathscr{F} = \mathscr{B}([0, 1])$, P — равномерное распределение. Определим для $k \geq 1$ $\xi_k = 2^{k-1}I([0,\ \frac{1}{2^{k-1}}])$. Тогда $\forall k$ $\mathsf{E}\xi_k = 1$, но $\xi = I(\omega = 0)$

Пример 8.6. $\xi_n \xrightarrow{\mathsf{P}} \xi$, но $\xi_n \overset{\text{п.н.}}{\not\to} \xi$ и $\xi_n \overset{L^p}{\not\to} \xi$.

Положим $\xi_{2^k} = e^{2^k} \cdot I([0,\ \frac{1}{2^k}])$, $\xi_{2^k+p} = e^{2^k+p} \cdot I([\frac{p}{2^k},\ \frac{p+1}{2^k}])$, $1 \leq p < 2^k$ (т.е. случайные величины из примера 8.3, домноженные на e^n). Аналогично примеру 8.3, имеется сходимость по вероятности к $\xi = 0$, но нет сходимости почти наверное.

Заметим теперь, что $\mathsf{E}|\xi_n - \xi|^p = \mathsf{E}\xi_n^p \not\to 0$ при $n \to \infty$.

Определение 8.7. Последовательность $\{F_n\}_{n=1}^{\infty}$ функций распределения на $\mathbb{R}$ *сходится в основном* к функции распределения F на $\mathbb{R}$, если $\forall x \in C(F)\ :\ F_n(x) \to F(x)$, где $C(F)$ — множество точек непрерывности функции F.

Обозначение: $F_n \Rightarrow F$.

Определение 8.8. Последовательность $\{\mathsf{P}_n\}_{n=1}^{\infty}$ вероятностных мер на $(\mathbb{R}^k,\ \mathscr{B}(\mathbb{R}^k))$ *сходится в основном* к вероятностной мере P на $(\mathbb{R}^k,\ \mathscr{B}(\mathbb{R}^k))$, если $\forall A \in \mathscr{B}(\mathbb{R}^k)$, таких что $\mathsf{P}(\partial A) = 0$ выполнено, что

$\mathsf{P}_n(A) \to \mathsf{P}(A)$.

Обозначение: $\mathsf{P}_n \Rightarrow \mathsf{P}$.

Теорема 8.2. *(Александрова, б/д)*

Следующие условия эквивалентны:

1. $\xi_n \xrightarrow{d} \xi$
2. $F_{\xi_n} \Rightarrow F_\xi$
3. $\mathsf{P}_{\xi_n} \Rightarrow \mathsf{P}_\xi$

8.2 Лемма Бореля-Кантелли. Критерии Коши сходимости случайных величин

Определение 8.9. Пусть $\{A_n\}_{n=1}^{\infty}$ – последовательность событий на $(\Omega,\ \mathscr{F},\ \mathsf{P})$. Тогда событием "$A_n$ бесконечно часто" называется событие $B = \{\omega \in \Omega :\ |k \in \mathbb{N} :\ \omega \in A_k| = \infty\}$.

Обозначение: $\{A_n$ б.ч. $\}$.

$\{A_n \text{ б.ч. }\} = \{\omega |\ \forall n\ \exists k \geq n\ \omega \in A_k\} = \bigcap\limits_{n=1}^{\infty} \bigcup\limits_{k \geq n} A_k$

Теорема 8.3. *(лемма Бореля-Кантелли)*

1. *Если* $\sum\limits_{n=1}^{\infty} \mathsf{P}(A_n) < \infty$, *то* $\mathsf{P}(A_n \text{ б.ч. }) = 0$.
2. *Если* $\sum\limits_{n=1}^{\infty} \mathsf{P}(A_n) = \infty$ *и* $\{A_n\}$ *независимы в совокупности, то* $P(A_n \text{ б.ч. }) = 1$.

Доказательство. 1. $\mathsf{P}(A_n \text{ б.ч. }) = \mathsf{P}(\bigcap\limits_{n=1}^{\infty} \bigcup\limits_{k \geq n} A_k) = \lim\limits_{n \to \infty} \mathsf{P}(\bigcup\limits_{k \geq n} A_k) \leq \lim\limits_{n \to \infty} \sum\limits_{k=n}^{\infty} \mathsf{P}(A_k) = 0$ (как остаточный член сходящегося ряда).

2.

$$\begin{aligned}
\mathsf{P}(A_n \text{ б.ч. }) &= \mathsf{P}(\bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k) \\
&= \lim_{n\to\infty} \mathsf{P}(\bigcup_{k\geq n} A_k) \\
&= 1 - \lim_{n\to\infty} \mathsf{P}(\bigcap_{k\geq n} \overline{A}_k) \\
&= 1 - \lim_{n\to\infty} (\lim_{N\to\infty} \mathsf{P}(\bigcap_{k=n}^{N} \overline{A}_k)) \\
&= 1 - \lim_{n\to\infty} \lim_{N\to\infty} \prod_{k=n}^{N} (1 - \mathsf{P}(A_k)) \\
&\geq 1 - \lim_{n\to\infty} \lim_{N\to\infty} e^{-\sum\limits_{k=n}^{N} \mathsf{P}(A_k)} \\
&= 1 - \lim_{n\to\infty} e^{-\sum\limits_{k=n}^{\infty} \mathsf{P}(A_k)} \\
&= 1 - \lim_{n\to\infty} e^{-\infty} = 1
\end{aligned}$$

$\square$

Определение 8.10. Последовательность случайных величин $\{\xi_n\}$ на $(\Omega, \mathscr{F}, \mathsf{P})$ называется *фундаментальной по вероятности*, если $\forall \varepsilon > 0 : \ \mathsf{P}(|\xi_n - \xi_m| \geq \varepsilon) \to 0$ при $n, m \to \infty$.

Теорема 8.4. *(критерий Коши сходимости случайных величин по вероятности)*

$\{\xi_n\}$ фундаментальна по вероятности $\iff$ $\exists \xi$ — случайная величина на $(\Omega, \mathscr{F}, \mathsf{P})$, такая что $\xi_n \xrightarrow{\mathsf{P}} \xi$.

Доказательство. $\Leftarrow$:

$$\{|\xi_n - \xi| \leq \frac{\varepsilon}{2}\} \cap \{|\xi_m - \xi| \leq \frac{\varepsilon}{2}\} \subseteq \{|\xi_n - \xi_m| \leq \varepsilon\} \Rightarrow \{|\xi_n - \xi_m| > \varepsilon\} \subseteq \{|\xi_n - \xi| > \frac{\varepsilon}{2}\} \cup \{|\xi_m - \xi| > \frac{\varepsilon}{2}\} \downarrow \varnothing$$

По условию вероятность каждого события стремится к 0, а значит $\mathsf{P}(|\xi_n - \xi_m| > \varepsilon) \to 0$ при $n, m \to \infty$.

$\Rightarrow$:

Сначала докажем следующую лемму:

Лемма 8.2. *Если ξ_n фундаментальна по вероятности, то существует такая случайная величина ξ и подпоследовательность ξ_{n_k}, что $\xi_{n_k} \xrightarrow{п.н.} \xi$ при $k \to \infty$.*

Доказательство. Пусть $n_1 = 1$, $n_k = \min\{j > n_{k-1} \, \forall r, \, s \geq j \; \mathsf{P}(|\xi_r - \xi_s| \geq 2^{-k}) < 2^{-k}\}$ (это можно сделать, т.к. последовательность фундаментальна по вероятности). Далее, рассмотрим событие $\{|\xi_{n_k} - \xi_{n_{k-1}}| \geq 2^{-(k-1)}$ б.ч. $\}$. По лемме Бореля-Кантелли $\sum\limits_{k=2}^{\infty} \mathsf{P}(|\xi_{n_k} - \xi_{n_{k-1}}| \geq 2^{-(k-1)}) < \sum\limits_{k=2}^{\infty} 2^{-(k-1)} < \infty$ а значит

$$\begin{aligned}
&\mathsf{P}(|\xi_{n_k} - \xi_{n_{k-1}}| \geq 2^{-(k-1)} \text{ б.ч. }) = 0 \\
&\Rightarrow \mathsf{P}(\sum_{k=2}^{\infty} |\xi_{n_k} - \xi_{n_{k-1}}| < \infty) = 1 \\
&\Rightarrow \mathsf{P}(\sum_{k=2}^{\infty} \xi_{n_k} - \xi_{n_{k-1}} < \infty) = 1
\end{aligned}$$

Положим $A = \{\omega : \sum\limits_{k=2}^{\infty} |\xi_{n_k} - \xi_{n_{k-1}}| < \infty\}$. Как было показано выше, $\mathsf{P}(A) = 1$. Определим случайную величину ξ для любого $\omega \in A$ как $\lim\limits_{k\to\infty} \xi_{n_k}$ (опять же, выше показано, что на множестве A этот предел существует). Для $\omega \in \overline{A}$ положим, например $\xi(\omega) = 0$. На сходимость почти наверное это никак не повлияет. Лемма доказана. $\square$

Выделим сходяющуюся почти наверное подпоследовательность $\xi_{n_k} \overset{\text{п.н.}}{\to} \xi$. Рассмотрим событие

$$\{|\xi_n - \xi| \leq \varepsilon\} \supseteq \{|\xi_{n_k} - \xi| \leq \tfrac{\varepsilon}{2}\} \cap \{|\xi_{n_k} - \xi_n| \leq \tfrac{\varepsilon}{2}\}$$

Вероятность каждого из событий справа стремится к 1. Значит, и $\mathsf{P}(|\xi_n - \xi| \leq \varepsilon) \to 1$. $\square$

Определение 8.11. Последовательность случайных величин ξ_n фундаментальна почти наверное, если $\mathsf{P}(\xi_n \text{ фундаментальна }) = 1$.

Теорема 8.5. *(критерий Коши сходимости почти наверное)*

Пусть $\xi_1, \xi_2, \ldots$ — случайные величины на $(\Omega, \mathscr{F}, \mathsf{P})$. Тогда следующие утверждения эквивалентны:

1. *$\{\xi_n\}$ почти наверное фундаментальна*

2. *Существует случайная величина ξ на $(\Omega, \mathscr{F}, \mathsf{P})$, такая что $\xi_n \overset{п.н.}{\to} \xi$*

3. *$\forall \varepsilon > 0 : \mathsf{P}(\sup\limits_{k\geq 1} |\xi_n - \xi_{n+k}| \geq \varepsilon) \to 0$ при $n \to \infty$*

Доказательство. 1) $\Leftrightarrow$ 2):

$$\begin{aligned}
\xi_n \text{ п.н. фундаментальна } &\Leftrightarrow \mathsf{P}(\xi_n \text{ фундаментальна }) = 1 \\
&\Leftrightarrow \mathsf{P}(\xi_n \text{ сходится }) = 1.
\end{aligned}$$

1) $\Leftrightarrow$ 3):

$$\begin{aligned}
\xi_n \text{ п.н. фундаментальна } &\Leftrightarrow \mathsf{P}(\forall \varepsilon > 0\, \exists N\, \forall n,\, m \geq N |\xi_n - \xi_m| \leq \varepsilon) = 1 \\
&\Leftrightarrow \mathsf{P}(\forall k \in \mathbb{N}\, \exists N\, \forall n,\, m \geq N |\xi_n - \xi_m| \leq \frac{1}{k}) = 1 \\
&\Leftrightarrow \forall k \in \mathbb{N} : \mathsf{P}(\exists N\, \forall n,\, m \geq N |\xi_n - \xi_m| \leq \frac{1}{k}) = 1 \\
&\Leftrightarrow \forall \varepsilon > 0 : \mathsf{P}(\exists N\, \forall n,\, m \geq N |\xi_n - \xi_m| \leq \varepsilon) = 1 \\
&\Leftrightarrow \forall \varepsilon > 0 : \mathsf{P}(\forall N\, \exists n,\, m \geq N |\xi_n - \xi_m| \geq \varepsilon) = 0 \\
&\Leftrightarrow \forall \varepsilon > 0 : \mathsf{P}(\forall N\, \exists n \geq N\, \exists k \in \mathbb{N}\, |\xi_n - \xi_{n+k}| \geq \varepsilon) = 0 \text{ — } (1)
\end{aligned}$$

С другой стороны,

$$\forall\varepsilon > 0 : \ \mathsf{P}(\sup_{k\geq 1}|\xi_n - \xi_{n+k}| \geq \varepsilon) \to 0 \iff \forall\varepsilon > 0 : \ \mathsf{P}(\exists k \geq 1 : \ |\xi_{n+k} - \xi_n| \geq \varepsilon) \to 0 - (2).$$

(1) $\Leftrightarrow \forall\varepsilon > 0 : \ \lim\limits_{N\to\infty} \mathsf{P}(\exists n \geq N \, \forall k \in \mathbb{N} \ |\xi_n - \xi_{n+k}| \geq \varepsilon) = 0$

Отсюда, очевидно, следует (2). Наоборот, если (2) выполнено, то

$$\forall\varepsilon > 0 \, \forall\delta > 0 \, \exists n_0 \, \forall N > n_0 \ \mathsf{P}(\exists k \geq 1 : |\xi_{N+k} - \xi_N| \geq \varepsilon/2) < \delta$$

тогда

$$\mathsf{P}(\exists n \geq N \, \exists k \geq 1 \ |\xi_{n+k} - \xi_n| \geq \varepsilon) \leq \mathsf{P}(\exists n \geq N \, \exists k \geq 1 \ |\xi_{N+k} - \xi_n| \geq \varepsilon/2)$$

или

$$\mathsf{P}(|\xi_N - \xi_n| \geq \varepsilon/2) \leq \mathsf{P}(\exists k \geq 1 \ |\xi_{n+k} - \xi_n| \geq \varepsilon/2) < \delta.$$

$\square$

9 Случайное блуждание

9.1 Определение. Закон повторного логарифма

Определение 9.1. Пусть $\{\xi_n\}_{n=1}^{\infty}$ — незавсимо одинаково распределенные случайные величины на $(\Omega, \mathscr{F}, \mathsf{P})$. Обозначим за $S_0 = 0$, $S_n = \sum\limits_{k=1}^{n} \xi_k$. Тогда последовательность $\{S_n\}_{n=1}^{\infty}$ называется *случайным блужданием*. Если $\mathsf{P}(\xi_n = 1) = p$, $\mathsf{P}(\xi_n = -1) = 1 - p$, то последовательность S_n называется *простейшим* случайным блужданием. Если $p = \frac{1}{2}$, то случайное блуждание называется *симметричным*. Последовательность $\{S_n(\omega)\}$ для фиксированного $\omega \in \Omega$ называется *траекторией*.

Теорема 9.1. *(закон повторного логарифма, б/д)*
Пусть $\xi_1, \xi_2, \ldots$ — независимые одинаково распределенные случайные величины, $\mathsf{E}\xi_i = 0$, $\mathsf{D}\xi_i = \sigma$, $0 < \sigma < +\infty$. *Тогда:*

$$\mathsf{P}(\varlimsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\sigma \ln\ln n}} = 1) = 1$$

Утверждение 9.1.1. Пусть $\{\xi_n\}$ — последовательность независимых экспоненциально распределенных случайных величин с параметром 1. Тогда

$$\mathsf{P}(\varlimsup_{n\to\infty} \frac{\xi_n}{\ln n} = 1) = 1.$$

Доказательство. 1. Докажем, что $\mathsf{P}(\varlimsup\limits_{n\to\infty} \frac{\xi_n}{\ln n} \geq 1) = 1$. Для этого рассмотрим подробнее это событие:

$$\begin{aligned}\{\varlimsup_{n\to\infty} \frac{\xi_n}{\ln n} \geq 1\} &= \{\forall \varepsilon > 0\ \frac{\xi_n}{\ln n} \geq 1 - \varepsilon \text{ б. ч. }\} \\ &= \{\forall k \in \mathbb{N}\ \frac{\xi_n}{\ln n} \geq 1 - \frac{1}{k} \text{ б. ч. }\} \\ &= \bigcap_{k\in\mathbb{N}} \{\frac{\xi_n}{\ln n} \geq 1 - \frac{1}{k} \text{ б. ч. }\}\end{aligned}$$

Заметим, что достаточно доказать, что вероятность каждого события под знаком пересечения равна одному. Очевидно тогда, что и вероятность всего пересечения будет равна одному. По лемме Бореля-Кантелли имеем:

$$\begin{aligned}\sum_{n=1}^{\infty} \mathsf{P}(\frac{\xi_n}{\ln n} \geq 1 - \frac{1}{k}) &= \sum_{n=1}^{\infty} e^{-(1-\frac{1}{k})\ln n} \\ &= \sum_{n=1}^{\infty} n^{-(1-\frac{1}{k})} = \infty\end{aligned}$$

а значит $\mathsf{P}(\{\varlimsup\limits_{n\to\infty} \frac{\xi_n}{\ln n} \geq 1\}) = 1$.

2. Докажем, что $\mathsf{P}(\overline{\lim\limits_{n\to\infty}} \frac{\xi_n}{\ln n} \le 1) = 1$ аналогично:

$$\{\overline{\lim} \frac{\xi_n}{\ln n} \le 1\} = \overline{\{\exists \varepsilon > 0 : \frac{\xi_n}{\ln n} \ge 1 + \varepsilon \text{ б.ч. }\}}$$
$$= \overline{\{\exists k \in \mathbb{N}\ \frac{\xi_n}{\ln n} \ge 1 + \frac{1}{k} \text{ б. ч. }\}}$$
$$= \bigcap_{k\in\mathbb{N}} \overline{\{\frac{\xi_n}{\ln n} \ge 1 + \frac{1}{k} \text{ б. ч. }\}}$$

Покажем, что вероятность каждого события из пересечения равна 1 $\iff$ $\mathsf{P}(\{\frac{\xi_n}{\ln n} \ge 1+\frac{1}{k} \text{ б. ч. }\}) = 0$

$$\sum_{n=1}^{\infty} \mathsf{P}(\{\frac{\xi_n}{\ln n} \ge 1 + \frac{1}{k}\}) = \sum_{n=1}^{\infty} e^{-(1+\frac{1}{k})\ln n} = \sum_{n=1}^{\infty} n^{-(1+\frac{1}{k})} < \infty$$

откуда по лемме Бореля-Кантелли следует требуемое.

Объединяя результаты 1) и 2) получаем утверждение. □

9.2 Некоторые факты о случайном блуждании

Утверждение 9.2.1. Рассмотрим простейшее случайное блуждание, $\mathsf{P}(\xi_i = 1) = p$, $\mathsf{P}(\xi_i = 0) = 1 - p$. Тогда

$$\mathsf{P}(S_n = k) = \begin{cases} C_n^{\frac{n+k}{2}} p^{\frac{n+k}{2}} (1-p)^{\frac{n-k}{2}}, & |k| \le n,\ 2 \mid n + k \\ 0, & \text{иначе} \end{cases}$$

Доказательство. Представим, что мы находимся на прямой, и должны прийти в точку k. Для начала заметим, что четность числа шагов и возможных положений за такое число шагов всегда совпадает. Теперь, чтобы прийти в точку k, мы должны сделать k шагов вправо, а потом x шагов влево и x шагов вправо. Тогда, $k + 2x = n$ и количество шагов вправо: $k + x$, количество шагов влево: x. □

Задача. Найти $\mathsf{P}(S_1 > 0,\ S_2 > 0,\ \ldots,\ S_{n-1} > 0,\ S_n = k)$ — вероятность того, что случайное блуждание никогда не возвратиться в 0.

Решение: Найдем количество траекторий, пересекающих 0. Если траектория пересекает 0 первый раз в точке $x = x_0$, то отразим дальнейшую ее часть относительно оси x. Тогда, если раньше траектория приходила в k, то теперь она приходит в $-k$. Заметим, что между траекториями, где $S_1 > 0$ до точки k за еще $n-1$ шаг, касающимися 0, и траекториями где $S_1 > 0$ до точки $-k$ тогда существует биекция. Обозначим за $N(n,\ k)$ количество путей из 0 в k за n шагов. По предыдущей задаче, это $C_n^{\frac{n+k}{2}}$. Тогда искомая вероятность равна

$$(N(n-1,\ k-1) - N(n-1, -k-1))p^{\frac{n+k}{2}} q^{\frac{n-k}{2}}$$

□

10 Закон больших чисел

10.1 Закон больших чисел в форме Чебышева

Теорема 10.1. *(Закон больших чисел в форме Чебышева)*

Пусть $\xi_1, \xi_2, \ldots$ *— независимые одинаково распределенные случайные величины на* $(\Omega, \mathscr{F}, \mathsf{P})$*, такие что* $\mathsf{E}\xi_i^2 < \infty$*. Тогда для любого* $\delta > 0$

$$\frac{S_n - \mathsf{E}S_n}{n^{1/2+\delta}} \xrightarrow{\mathsf{P}} 0$$

Доказательство.

$$P(|S_n - \mathsf{E}S_n| \geq \varepsilon n^{\frac{1}{2}+\delta}) \leq \frac{\mathsf{D}S_n}{\varepsilon^2 n{1+2\delta}} = \frac{n\mathsf{D}\xi_1}{\varepsilon^2 n^{1+2\delta}} = \frac{\mathsf{D}\xi_1}{\varepsilon^2 n^{2\delta}}$$

□

Замечание: это утверждение верно для попарно некоррелированых величин.

10.2 Усиленные законы больших чисел

Теорема 10.2. *(неравенство Колмогорова)*

Пусть $\xi_1,\ \xi_2,\ \ldots$ *— независимые случайные величины, такие что* $\mathsf{E}\xi_i = 0,\ \mathsf{E}\xi_i^2 < \infty$*. Тогда:*

1. $\mathsf{P}(\sup\limits_{1\leq k\leq n} |S_k| \geq \varepsilon) \leq \frac{\mathsf{E}S_n^2}{\varepsilon^2}$

2. *Если дополнительно* $\forall i:\ \mathsf{P}(|\xi_i| \leq C) = 1$*, то* $\mathsf{P}(\sup\limits_{1\leq k\leq n} |S_k| \geq \varepsilon) \geq 1 - \frac{(C+\varepsilon)^2}{\mathsf{E}S_n^2}$

Доказательство. Определим события

$$A = \{\sup_{1\leq k\leq n} |S_k| \geq \varepsilon\}$$
$$A_k = \{|S_1| < \varepsilon,\ |S_2| < \varepsilon,\ \ldots,\ |S_{k-1}| < \varepsilon,\ |S_k| \geq \varepsilon\}.$$

Тогда $A = \bigsqcup\limits_{k=1}^{n} A_k$ и

$$\mathsf{E}S_n^2 \geq \mathsf{E}S_n^2 I_A = \sum_{k=1}^{n} \mathsf{E}S_n^2 I_{A_k}$$

Но

$$\begin{aligned}\mathsf{E}S_n^2 I_{A_k} &= \mathsf{E}(S_k + (\xi_{k+1} + \ldots + \xi_n))^2 I_{A_k} = \\ &= \mathsf{E}S_k^2 I_{A_k} + 2\mathsf{E}S_k(\xi_{k+1} + \ldots + \xi_n) I_{A_k} + \mathsf{E}(\xi_{k+1} + \ldots + \xi_n)^2 I_{A_k} \geq \mathsf{E}S_k^2 I_{A_k}\end{aligned}$$

посклькy $\mathsf{E}S_k(\xi_{k+1} + \ldots + \xi_n) I_{A_k} = \mathsf{E}S_k I_{A_k} \mathsf{E}(\xi_{k+1} + \ldots + \xi_n) = 0$ в силу предположений независимости и $\mathsf{E}\xi_i = 0$. Поэтому

$$\mathsf{E}S_n^2 \geq \sum_{k=1}^{n} \mathsf{E}S_n^2 I_{A_k} \geq \varepsilon^2 \sum_{k=1}^{n} \mathsf{P}(A_k) = \varepsilon^2 \mathsf{P}(A)$$

что и доказывает неравенство 1). Для доказательства неравенства 2) заметим, что

$$\mathsf{E}S_n^2 I_A = \mathsf{E}S_n^2 - \mathsf{E}S_n^2 I_{\overline{A}} \geq \mathsf{E}S_n^2 - \varepsilon^2 \mathsf{P}(\overline{A}) = \mathsf{E}S_n^2 - \varepsilon^2 + \varepsilon^2 \mathsf{P}(A)$$

С другой стороны, на множестве A_k выполнено $|S_{k-1}| \leq \varepsilon$, $|S_k| \leq |S_{k-1}| + |\xi_k| \leq \varepsilon + C$. И значит

$$\begin{aligned}\mathsf{E}S_n^2 I_A = \sum_{k=1}^n \mathsf{E}S_k^2 I_{A_k} + \sum_{k=1}^n \mathsf{E}(I_{A_k}(S_n - S_k)^2) \leq \\ \leq (\varepsilon + C)^2 \sum_{k=1}^n \mathsf{P}(A_k) + \sum_{k=1}^n \mathsf{P}(A_k) \sum_{j=k+1}^n \mathsf{E}\xi_j^2 \\ \leq \mathsf{P}(A)\Big[(\varepsilon + C)^2 + \sum_{j=1}^n \mathsf{E}\xi_j^2\Big] = \mathsf{P}(A)\big[(\varepsilon + C)^2 + \mathsf{E}S_n^2\big]\end{aligned}$$

Из этого находим, что

$$\mathsf{P}(A) \geq \frac{\mathsf{E}S_n^2 - \varepsilon^2}{(\varepsilon + C)^2 + \mathsf{E}S_n^2 - \varepsilon^2} = 1 - \frac{(\varepsilon + C)^2}{(\varepsilon + C)^2 + \mathsf{E}S_n^2 - \varepsilon^2} \geq 1 - \frac{(\varepsilon + C)^2}{\mathsf{E}S_n^2}$$

□

Теорема 10.3. *(теорема Колмогорова-Хинчина)*

Пусть $\{\xi_n\}$ *— последовательность независимых случайных величин, такая что* $\mathsf{E}\xi_i = 0$, $\mathsf{E}\xi_i^2 < \infty$. *Тогда*

1. *Из сходимости ряда* $\sum_{i=1}^{\infty} \mathsf{E}\xi_i^2$ *следует сходимость почти наверное ряда* $\sum_{i=1}^{\infty} \xi_i$
2. *Если дополнительно для любого* i *и некоторого* $c < \infty$ *выполнено* $\mathsf{P}(|\xi_i| \leq c) = 1$ *и ряд* $\sum_{i=1}^{\infty} \xi_i$ *сходится почти наверное, то сходится и ряд* $\sum_{i=1}^{\infty} \mathsf{E}\xi_i^2$

Доказательство. Заметим, что последовательность $\{\eta_n\}$ сходится почти наверное тогда и только тогда, когда $\forall \varepsilon > 0 :\ \mathsf{P}(\sup_{k\geq 1} |\eta_{n+k} - \eta_n| > \varepsilon) \to 0$ при $n \to \infty$. Возьмем $\eta_n := \sum_{i=1}^n \xi_i$. Тогда:

$$\eta_n \text{ сходится почти наверное } \iff \forall \varepsilon > 0 :\ \mathsf{P}(\sup_{k\geq 1} |\xi_{n+1} + \ldots + \xi_{n+k}| > \varepsilon) \to 0$$

Рассмотрим $S_k = \xi_{n+1} + \ldots + \xi_{n+k}$. По неравенству Колмогорова имеем

$$\begin{aligned}\mathsf{P}(\sup_{1\leq k\leq N} |S_k| > \varepsilon) < \frac{\mathsf{E}S_N^2}{\varepsilon^2} = \frac{\mathsf{E}(\xi_{n+1} + \ldots + \xi_{n+N})^2}{\varepsilon^2} = \\ = \frac{\mathsf{E}\xi_{n+1}^2 + \ldots + \mathsf{E}\xi_{n+N}^2}{\varepsilon^2} \leq \frac{\mathsf{E}\xi_{n+1}^2 + \mathsf{E}\xi_{n+2}^2 + \ldots}{\varepsilon^2} \to 0\end{aligned}$$

как остаточный член сходящегося ряда. Но тогда сходится и ряд $\sum_{i=1}^{\infty} \xi_i$ поскольку его остаточный член стремиться к 0, что и доказывает первую часть теоремы.

Для доказательства второй части заметим, что, по тому же неравенству Колмогорова, в случае когда $\mathsf{P}(|\xi_i| \leq c) = 1, c < \infty$ верно следующее:

$$\mathsf{P}(\sup_{1\leq k\leq N} |S_k| > \varepsilon) \geq 1 - \frac{(\varepsilon + c)^2}{\mathsf{E}S_N^2} = 1 - \frac{(\varepsilon + c)^2}{\mathsf{E}\xi_{n+1}^2 + \ldots + \mathsf{E}\xi_{n+N}^2}$$

Предположим, что ряд $\sum_{i=1}^{\infty} \mathsf{E}\xi_i^2$ расходится. Тогда сумму, стоящую в знаменателе, можно сделать сколь угодно большой, а значит $\mathsf{P}(\sup_{1\leq k\leq N} |S_k| > \varepsilon) \to 1$.

С другой стороны

$$\sum_{n=1}^{\infty} \xi_n < \infty \Leftrightarrow \forall \varepsilon > 0 : \mathsf{P}(\sup_{k \geq 1} |S_k| > \varepsilon) \to 0$$

$$\Leftrightarrow \forall \varepsilon > 0 : \mathsf{P}(\lim_{N\to\infty} \sup_{1 \leq k \leq N} |S_k| > \varepsilon) \to 0 \text{ при } n \to \infty.$$

Возьмем $\varepsilon = \frac{1}{2}$. Тогда, из выше сказанного, следует, что $\exists n : \mathsf{P}(\lim\limits_{N\to\infty} \sup\limits_{1 \leq k \leq N} |S_k| > \frac{1}{2}) < \frac{1}{2}$. Так как события $\{\sup\limits_{1 \leq k \leq N} |S_k| > \frac{1}{2}\}$ вложены друг в друга при росте N, имеем

$$\lim_{N\to\infty} \mathsf{P}(\sup_{1 \leq k \leq N} |S_k| > \tfrac{1}{2}) < \tfrac{1}{2}$$

что противоречит $\mathsf{P}(\sup\limits_{1 \leq k \leq N} |S_k| > \varepsilon) \to 1$, а значит предположение о расходимости ряда $\sum\limits_{i=1}^{\infty} \mathsf{E}\xi_i^2$ неверно. □

Лемма 10.1. *(лемма Теплица)*

Пусть $\{x_n\}$ — сходящаяся к x последовательность чисел, $\{a_n\}$ такая последовательность чисел, что $a_i \geq 0$, $\forall k \geq 1 :$ $\sum\limits_{n=1}^{k} a_n > 0$ и $\sum\limits_{n=1}^{k} a_n \to \infty$ при $k \to \infty$. Тогда

$$\frac{\sum\limits_{n=1}^{k} x_n a_n}{\sum\limits_{n=1}^{k} a_n} \to x,\ k \to \infty$$

Доказательство. Из сходимости последовательности x_n получаем $\forall \varepsilon > 0\ \exists n_0\ \forall n \geq n_0 :\ |x_n - x| < \frac{\varepsilon}{2}$. Кроме этого, $\exists n_1 \geq n_0 \forall n \geq n_1 :$ $\left| \frac{\sum_{i=1}^{n_0} (x_i - x) a_i}{\sum_{i=1}^{n} a_i} \right| < \frac{\varepsilon}{2}$, т.к. числитель это фиксированное число, а знаменатель стремиться к бесконечности.

Тогда для $n \geq n_1$ имеем

$$\left| \frac{\sum\limits_{i=1}^{n} x_i a_i}{\sum\limits_{i=1}^{n} a_i} - x \right| = \left| \frac{\sum\limits_{i=1}^{n} (x_i - x) a_i}{\sum\limits_{i=1}^{n} a_i} \right| \leq \left| \frac{\sum\limits_{i=1}^{n_0} (x_i - x) a_i}{\sum\limits_{i=1}^{n} a_i} \right| + \left| \frac{\sum\limits_{i=n_0+1}^{n} (x_i - x) a_i}{\sum\limits_{i=1}^{n} a_i} \right|$$

$$\leq \frac{\varepsilon}{2} + \frac{\sum\limits_{i=n_0+1}^{n} |x_i - x| a_i}{\sum\limits_{i=1}^{n} a_i} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \cdot \frac{\sum\limits_{i=n_0+1}^{n} a_i}{\sum\limits_{i=1}^{n} a_i} \leq \varepsilon$$

□

Лемма 10.2. *(лемма Кронекера)*

Пусть даны две числовые последовательности: неубывающая $\{b_n\} \to +\infty$, такая что $\forall n\, b_n > 0$, и $\{x_n\}$, такая что $\sum\limits_{n=1}^{\infty} x_n = x < \infty$. Тогда

$$\frac{1}{b_n} \sum_{k=1}^{n} b_k x_k \to 0 \text{ при } n \to \infty.$$

Доказательство. Положим $a_k := b_k - b_{k-1}$ и $b_0 := 0$. Тогда $b_n = \sum\limits_{k=1}^{n} a_k$. При этом

$$\frac{1}{b_n}\sum_{k=1}^{n} b_k x_k = \frac{1}{\sum\limits_{k=1}^{n} a_k}\sum_{k=1}^{n} x_k \sum_{i=1}^{k} a_i = \frac{1}{\sum\limits_{k=1}^{n} a_k}\sum_{i=1}^{n} a_i \sum_{j=i}^{n} x_j = \frac{1}{\sum\limits_{k=1}^{n} a_k}\left(\sum_{i=1}^{n} a_i\left(\sum_{j=i}^{\infty} x_j - \sum_{j=n+1}^{\infty} x_j\right)\right)$$

Положим $y_{n+1} := \sum\limits_{k=n+1}^{\infty} x_k$. Тогда, поскольку ряд $\sum\limits_{n=1}^{\infty} x_n$ сходится, имеем $\lim\limits_{n\to\infty} y_n = 0$, а значит $\forall \varepsilon > 0:$ $\exists n_0\ \forall n \geq n_0 |y_{n+1}| < \varepsilon$. По лемме 10.1 $\exists n_1 \geq n_0\ \forall n \geq n_1:$ $\left|\frac{\sum\limits_{i=1}^{n} a_i y_i}{\sum\limits_{i=1}^{n} a_i}\right| < \frac{\varepsilon}{2}$.

Тогда $\forall n \geq n_1$ имеем

$$\left|\frac{1}{b_n}\sum_{k=1}^{n} b_k x_k\right| \leq \left|\frac{\sum\limits_{i=1}^{n} a_i y_i}{\sum\limits_{i=1}^{n} a_i}\right| + \left|\frac{\sum\limits_{i=1}^{n} a_i y_{n+1}}{\sum\limits_{i=1}^{n} a_i}\right| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

□

Теорема 10.4. *(Усиленный закон больших чисел 1)*

Пусть $\xi_1, \xi_2, \ldots$ — независимые случайные величины с $\mathsf{E}\xi_i^2 < \infty$. Если существует последовательность чисел $\{b_n\}$, такая что $0 < b_1 \leq b_2 \leq \ldots,\ b_n \leq \ldots \to \infty$ и $\sum\limits_{n=1}^{\infty} \frac{\mathsf{D}\xi_n}{b_n^2} < \infty$, то

$$\frac{S_n - \mathsf{E}S_n}{b_n} \stackrel{\text{п.н.}}{\to} 0$$

Доказательство. Заметим, что если последовательность ξ_n такая, что $\forall n:\ \mathsf{D}\xi_n \leq C$, то можно положить $b_n := n^{1/2+\delta}$. Тогда $\sum\limits_{n=1}^{\infty} \frac{\mathsf{D}\xi_n}{b_n^2} < C \sum\limits_{n=1}^{\infty} \frac{1}{b_n^2} < \infty$ (т.е. для некоторых последовательностей случайных величин такая последовательность $\{b_n\}$ существует. Сравните это с ЗБЧ в форме Чебышева.).

Приступим к доказательству теоремы. Без ограничения общности предположим, что $\mathsf{E}\xi_i = 0$, то есть $\mathsf{D}\xi_n = \mathsf{E}\xi_n^2$ (рассмотреть случайные величины $\eta = \xi - \mathsf{E}\xi$). Тогда

$$\sum_{n=1}^{\infty} \mathsf{E}\left(\frac{\xi_n}{b_n}\right)^2$$

сходится, а значит, по теореме Колмогорова-Хинчина ряд $\sum\limits_{n=1}^{\infty} \frac{\xi_n}{b_n}$ сходится. Отсюда и из леммы Кронекера следует, что

$$\frac{1}{b_n}\sum_{k=1}^{n} b_k \cdot \frac{\xi_k}{b_k} = \frac{\sum\limits_{k=1}^{n} \xi_k}{b_n} \stackrel{\text{п.н.}}{\to} 0$$

□

Теорема 10.5. *(Усиленный закон больших чисел 2)*

Пусть $\xi_1, \xi_2, \ldots$ — независимые одинаково распределенные случайные величины, $\mathsf{E}|\xi_1| < \infty$. Тогда

$$\frac{S_n - \mathsf{E}S_n}{n} \stackrel{\text{п.н.}}{\to} 0.$$

Доказательство. Опять же, положим $\mathsf{E}\xi_n = 0$. Рассмотрим ряд

$$\sum_{n=1}^{\infty} \mathsf{P}(|\xi_1| \geq n) = \sum_{n=1}^{\infty} n\mathsf{P}(n \leq |\xi_1| \leq n+1) = \sum_{n=1}^{\infty} \mathsf{E}nI(n \leq |\xi_1| \leq n+1) =$$
$$= \mathsf{E}\sum_{n=1}^{\infty} nI(n \leq |\xi_1| \leq n+1) = \mathsf{E}\lfloor|\xi_1|\rfloor \in \Big[\mathsf{E}|\xi_1| - 1;\ \mathsf{E}|\xi_1|\Big]$$

а значит сходимость ряда $\sum\limits_{n=1}^{\infty} \mathsf{P}(|\xi_1| \geq n)$ эквивалентна конечности мат.ожидания $\mathsf{E}|\xi_1| < \infty$. Однако случайные величины одинаково распределенные, а значит

$$\sum_{n=1}^{\infty} \mathsf{P}(|\xi_1| \geq n) < \infty \Leftrightarrow \sum_{n=1}^{\infty} \mathsf{P}(|\xi_n| \geq n) < \infty$$
$$\Rightarrow \mathsf{P}(|\xi_n| \geq n \text{ б. ч.}) = 0 \text{ по лемме Бореля-Кантелли}$$

Рассмотрим $\tilde{\xi}_n := \xi_n I(|\xi_n| \leq n)$. Из того, что $\mathsf{P}(|\xi_n| \geq n \text{ б. ч.}) = 0$ следует, что лишь конечное число членов последовательности $\{\xi_k\}$ больше n по модулю, а значит, для достаточо больших n, выполнено $\tilde{\xi}_n = \xi_n$ почти наверное, и поэтому

$$\frac{\xi_1 + \ldots + \xi_n}{n} \overset{\text{п.н.}}{\to} 0 \iff \frac{\tilde{\xi}_1 + \ldots + \tilde{\xi}_n}{n} \overset{\text{п.н.}}{\to} 0$$

Покажем, что $\frac{\mathsf{E}\tilde{\xi}_1 + \ldots + \mathsf{E}\tilde{\xi}_n}{n} \to 0$. Если это так, то

$$\frac{\xi_1 + \ldots + \xi_n}{n} \overset{\text{п.н.}}{\to} 0 \Leftrightarrow \frac{\tilde{\xi}_1 + \ldots + \tilde{\xi}_n - \mathsf{E}\tilde{\xi}_1 - \ldots - \mathsf{E}\tilde{\xi}_n}{n} \overset{\text{п.н.}}{\to} 0$$

Действительно, имеем

$$\mathsf{E}\tilde{\xi}_n = \mathsf{E}\xi_n I(|\xi_n| \leq n) = \mathsf{E}\xi_1 I(|\xi_1| \leq n)$$

Заметим, что $\xi_1 I(|\xi_1| \leq n) \to \xi_1$ при $n \to \infty$, а $|\xi_1 I(|\xi_1| \leq n)| \leq |\xi_1|$, $\mathsf{E}|\xi_1| < \infty$, а значит, по теореме Лебега о мажорируемой сходимости

$$\mathsf{E}\tilde{\xi}_n \to \mathsf{E}\xi_1 = 0$$

Применяя лемму Теплица для $a_n = 1$ получаем, что $\frac{1}{n}(\mathsf{E}\tilde{\xi}_1 + \ldots + \mathsf{E}\tilde{\xi}_n) \to 0$.

Рассмотрим теперь $\sum\limits_{n=1}^{\infty} \frac{\mathsf{D}\tilde{\xi}_n}{n^2}$:

$$\sum_{n=1}^{\infty} \frac{\mathsf{D}\tilde{\xi}_n}{n^2} \leq \sum_{n=1}^{\infty} \frac{\mathsf{E}\tilde{\xi}_n^2}{n^2} = \sum_{n=1}^{\infty} \frac{\mathsf{E}\xi_n^2 I(|\xi_n| \leq n)}{n^2}$$
$$= \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{k=1}^{n} \mathsf{E}\xi_n^2 I(k-1 < |\xi_n| \leq k)$$
$$\leq \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{k=1}^{n} k\mathsf{E}|\xi_n| I(k-1 < |\xi_n| \leq k)$$
$$= \sum_{k=1}^{\infty} k \sum_{n=k}^{\infty} \frac{1}{n^2} \mathsf{E}|\xi_n| I(k-1 < |\xi_n| \leq k)$$
$$= \sum_{k=1}^{\infty} k \sum_{n=k}^{\infty} \frac{1}{n^2} \mathsf{E}|\xi_1| I(k-1 < |\xi_1| \leq k)$$
$$= \sum_{k=1}^{\infty} k\mathsf{E}|\xi_1| I(k-1 < |\xi_1| \leq k) \sum_{n=k}^{\infty} \frac{1}{n^2}$$

Для того, чтобы оценить получившуюся сумму, воспользуемся неравенством $\sum\limits_{n=k}^{\infty}\frac{1}{n^2} \le \int\limits_{k-1}^{+\infty}\frac{dt}{t^2} = \frac{2}{k-1}$. С учетом этого, имеем:

$$\sum_{k=1}^{\infty} k\mathsf{E}|\xi_1| I(k-1<|\xi_1|\le k)\sum_{n=k}^{\infty}\frac{1}{n^2} \le 4\mathsf{E}|\xi_1|I(0<|\xi_1|\le 1)+\sum_{k=2}^{\infty}\frac{2k}{k-1}\mathsf{E}|\xi_1|I(k-1<|\xi_1|\le k)$$
$$\le 4\mathsf{E}|\xi_1|I(0<|\xi_1|\le 1)+4\sum_{k=2}^{\infty}\mathsf{E}|\xi_1|I(k-1<|\xi_1|\le k) = 4\mathsf{E}|\xi_1| < \infty$$

Мы показали, что $\sum\limits_{n=1}^{\infty}\frac{\mathsf{D}\tilde{\xi_n}}{n^2}$ сходится. Теперь, по теореме Колмогорова-Хинчина (как и в УЗБЧ-1), получаем

$$\sum_{n=1}^{\infty}\frac{\tilde{\xi_n}-\mathsf{E}\tilde{\xi_n}}{n} < \infty$$

откуда по лемме Кронекера

$$\frac{1}{n}\sum_{k=1}^{n}\tilde{\xi}_k - \mathsf{E}\tilde{\xi}_k \to 0$$

что и завершает доказательство теоремы. □

10.3 Неравенство больших уклонений

Теорема 10.6. *(Неравенство больших уклонений)*

Пусть $\xi_1, \xi_2, \dots$ — последовательность независимых одинаково распределенных случайных величин. Пусть $\mathsf{E}\xi_i = \mathsf{E}\xi_1 = a$. Тогда, если у ξ_1 существуют все моменты, то $\exists c_1,\ c_2 > 0$, такие что

$$\mathsf{P}\left(\left|\frac{S_n - \mathsf{E}S_n}{n}\right| \ge \varepsilon\right) \le e^{-c_1 n} + e^{-c_2 n}$$

Доказательство. Пусть случайная величина ξ имеет такое же распределение, как и ξ_i. Рассмотрим функцию $\varphi_\xi(\lambda) = \ln \mathsf{E}e^{\lambda\xi}$. Пусть U это некоторая окрестность 0, в которой функция φ_ξ бесконечно диффференцируема (такая окрестность найдется поскольку у ξ существуют все моменты). Тогда

$$\varphi'_\xi(\lambda)\Big|_{\lambda=0} = \frac{\mathsf{E}\xi e^{\lambda\xi}}{\mathsf{E}e^{\lambda\xi}}\Big|_{\lambda=0} = a$$

$$\varphi''_\xi(\lambda)\Big|_{\lambda=0} = \frac{\mathsf{E}\xi^2 e^{\lambda\xi}\cdot\mathsf{E}e^{\lambda\xi} - \left(\mathsf{E}\xi e^{\lambda\xi}\right)^2}{(\mathsf{E}e^{\lambda\xi})^2}\Big|_{\lambda=0} = \mathsf{D}\xi > 0$$

Рассмотрим новую функцию $\psi(b) = \sup\limits_{\lambda\in U}(b\lambda - \varphi_\xi(\lambda))$. Поскольку φ_ξ выпукла вниз, то при $b > a$ супремум достигается при $\lambda > 0$, а при $b < a$ соответственно при $\lambda < 0$:

$$b > a \Rightarrow \psi(b) = \sup_{\substack{\lambda\in U\\ \lambda>0}}(b\lambda - \varphi_\xi(\lambda))$$

$$b < a \Rightarrow \psi(b) = \sup_{\substack{\lambda\in U\\ \lambda<0}}(b\lambda - \varphi_\xi(\lambda))$$

$$\psi(a) = 0$$

Распишем по неравенству Маркова ($\lambda > 0$):

$$\begin{aligned}
\mathsf{P}\left(\frac{S_n}{n} - a > \varepsilon\right) &= \mathsf{P}\left(\frac{S_n}{n} > a + \varepsilon\right) \\
&= \mathsf{P}\left(exp\left(\lambda \cdot \frac{S_n}{n}\right) > exp\left(\lambda(a+\varepsilon)\right)\right) \\
&\leq \frac{\mathsf{E}\left(exp\left(\lambda \cdot \frac{S_n}{n}\right)\right)}{exp(\lambda(a+\varepsilon))} \\
&= exp\left(-\lambda(a+\varepsilon) + \ln \mathsf{E}\left(exp(\lambda \cdot \frac{S_n}{n})\right)\right) \\
&= exp\left(-\left(\lambda(a+\varepsilon) - \sum_{i=1}^{n} \ln \mathsf{E}\left(exp\left(\frac{\lambda}{n}\xi_i\right)\right)\right)\right) \\
&= exp\left(-n\left(\frac{\lambda}{n}(a+\varepsilon) - \ln \mathsf{E}\left(exp\left(\frac{\lambda}{n}\xi\right)\right)\right)\right) = exp\left(-n\left(\frac{\lambda}{n}(a+\varepsilon) - \varphi(\frac{\lambda}{n})\right)\right)
\end{aligned}$$

Так как это верно для любого положительного λ, то верно и для того, где достигается супремум (поскольку $a + \varepsilon > a$). А значит

$$\mathsf{P}\left(\frac{S_n}{n} - a > \varepsilon\right) \leq e^{-n\psi(a+\varepsilon)}$$

Абсолютно аналогично

$$\mathsf{P}\left(\frac{S_n}{n} - a < -\varepsilon\right) \leq e^{-n\psi(a-\varepsilon)}$$

Тогда

$$\mathsf{P}\left(\left|\frac{S_n}{n} - a\right| > \varepsilon\right) \leq e^{-n\psi(a-\varepsilon)} + e^{-n\psi(a+\varepsilon)}$$

$\square$

11 Характеристические функции. Центральная предельная теорема

11.1 Характеристические функции

Определение 11.1. Пусть ξ — случайная величина на $(\Omega, \mathscr{F}, \mathsf{P})$. Тогда ее характеристической функцией называется функция

$$\varphi_\xi(t) = \mathsf{E}e^{it\xi} \text{ (прямое преобразование Фурье)}$$

Если ξ — случайный вектор, то

$$\varphi_\xi(t) = \mathsf{E}e^{i\langle t,\, \xi\rangle}, \ \ t \in \mathbb{R}^n$$

где $\langle t,\, \xi\rangle$ — скалярное произведение.

Пример 11.1. $\xi \sim Bern(p)$. Тогда $\varphi_\xi(t) = e^{it} \cdot p + (1-p)$.

Пример 11.2. $\xi \sim U([a,\, b])$. Тогда

$$\varphi_\xi(t) = \int\limits_a^b e^{itx}\frac{1}{b-a}dx = \left.\frac{e^{itx}}{(b-a)it}\right|_a^b = \frac{e^{itb}-e^{ita}}{it(b-a)}$$

Пример 11.3. $\xi \sim N(0,\, 1)$.

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

Пример 11.4. $\xi \sim N(a,\, \sigma^2)$. Тогда $\frac{\xi-a}{\sigma} \sim N(0,\, 1)$.

$$\varphi_{\frac{\xi-a}{\sigma}}(t) = e^{-\frac{t^2}{2}} = \mathsf{E}e^{it\frac{\xi-a}{\sigma}}$$

сделаем замену $t' = t\sigma$ т.к. t — произвольное

$$e^{-\frac{t^2\sigma^2}{2}} = \mathsf{E}e^{-it(\xi-a)}$$

$$\varphi_\xi(t) = \mathsf{E}e^{it\xi} = e^{ita-\frac{1}{2}\sigma^2t^2}$$

Теорема 11.1. *(теорема о единственности, б/д)*

Пусть $\varphi_{\xi_1}(t) = \varphi_{\xi_2}(t)\ \forall t \in \mathbb{R}$. Тогда у ξ_1 и ξ_2 одинаковые распределения (аналогичная теорема верна и для случайных векторов).

Теорема 11.2. *(О независимости)*

Пусть $\xi = (\xi_1,\, \ldots,\, \xi_n)$ — случайный вектор. Тогда $\xi_1,\, \ldots,\, \xi_n$ — независимы в совокупности тогда и только тогда, когда

$$\varphi_\xi(t_1,\, \ldots,\, t_n) = \prod_{i=1}^n \varphi_{\xi_i}(t_i)\ \ \forall t_1,\, \ldots,\, t_n \in \mathbb{R}$$

Доказательство. $\Rightarrow$:

Т.к. ξ_i независимы в совокупности, то, по утверждению 4.3.2, независимы в совокупности случайные вектора $(\cos \xi_i,\ \sin \xi_i)$. Имеем

$$\begin{aligned}\varphi_\xi(t_1,\ \ldots,\ t_n) = \mathsf{E}e^{i\sum \xi_i t_i} &= \mathsf{E}(\cos \xi_1 t_1 + i \sin \xi_1 t_1) \ldots (\cos \xi_n t_n + i \sin \xi_n t_n) \\ &= \mathsf{E}(\cos \xi_1 t_1 + i \sin \xi_1 t_1) \ldots \mathsf{E}(\cos \xi_n t_n + i \sin \xi_n t_n) = \varphi_{\xi_1}(t_1) \ldots \varphi_{\xi_n}(t_n)\end{aligned}$$

$\Rightarrow$:

Рассмотрим независимые с.в. $\eta_1,\ \ldots,\ \eta_n$, такие что $\forall i:\ \ \eta_i \overset{d}{=} \xi_i$ (*существование б/д*). Тогда

$$\varphi_{(\eta_1,\ \ldots,\ \eta_n)}(t_1,\ \ldots,\ t_n) = \prod_{i=1}^{n} \varphi_{\eta_i}(t_i) = \prod_{i=1}^{n} \varphi_{\xi_i}(t_i) = \varphi_\xi(t_1,\ \ldots,\ t_n)$$

и по теореме о единственности $(\eta_1,\ \ldots,\ \eta_n) \overset{d}{=} \xi$. $\square$

Пример 11.5 (Формула свертки). Пусть $\xi \perp\!\!\!\perp \eta$. Тогда

$$F_{\xi+\eta}(x) = \mathsf{P}(\xi + \eta \le x) = \int\limits_{u+v \le x} \mathsf{P}_\xi(du)\mathsf{P}_\eta(dv) = \int\limits_{-\infty}^{+\infty} \mathsf{P}_\xi(du) \int\limits_{-\infty}^{x-u} \mathsf{P}_\eta(dv) = \int\limits_{-\infty}^{+\infty} F_\eta(x-u)\mathsf{P}_\xi(du)$$

С другой стороны, $\varphi_{\eta+\xi}(x) = \mathsf{E}e^{i(\eta+\xi)x} = \varphi_\xi(x)\varphi_\eta(x)$. Пусть, например, $\xi \sim N(a_1,\ \sigma_1^2),\ \ \eta \sim N(a_2,\ \sigma_2^2)$. Для хар.функции имеем:

$$\varphi_{\xi+\eta}(t) = e^{ita_1 - \frac{1}{2}\sigma_1^2 t^2} e^{ita_2 - \frac{1}{2}\sigma_2^2 t^2} = e^{it(a_1+a_2) - \frac{1}{2}(\sigma_1^2 + \sigma_2^2)t^2}$$

а значит, по теореме о единственности, $\xi + \eta \sim N(a_1 + a_2,\ \sigma_1^2 + \sigma_2^2)$. Это *удобнее*, чем формула свертки.

Теорема 11.3. *(формула обращения, б/д)*

Пусть $\varphi_\xi(t)$ — характеристическая функция ξ. Тогда для любых точек $a < b$, в которых F_ξ непрерывна, выполнено

$$F_\xi(b) - F_\xi(a) = \frac{1}{2\pi} \lim_{c \to \infty} \int\limits_{-c}^{c} \frac{e^{-ita} - e^{-itb}}{it} \cdot \varphi_\xi(t)dt$$

Если дополнительно $\varphi_\xi(t)$ абсолютно интегрируема, то

$$p_\xi(x) = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} e^{-itx} \varphi_\xi(t)dt \text{ (обратное преобразование Фурье)}$$

Утверждение 11.1.1. $|\varphi(t)| \le 1$ и $\varphi(0) = 1$.

Доказательство. $|\mathsf{E}e^{it\xi}| \le \mathsf{E}|e^{it\xi}| = 1$ $\square$

Утверждение 11.1.2. Характеристическая функция равномерно непрерывна на $\mathbb{R}$.

Доказательство.

$$\begin{aligned}\forall\varepsilon>0: \ \left|\mathsf{E}e^{it\xi}-\mathsf{E}e^{is\xi}\right| &= \left|\mathsf{E}(e^{it\xi}-e^{is\xi})\right| \\ &= \left|\mathsf{E}(e^{is\xi})\cdot(e^{i(t-s)\xi}-1))\right| \\ &\le \mathsf{E}|e^{i(t-s)\xi}-1| \\ &= \mathsf{E}\sqrt{(\cos(t-s)\xi-1)^2+\sin^2(t-s)\xi} \\ &= \mathsf{E}\sqrt{2-2\cos(t-s)\xi} = 2\cdot\mathsf{E}\left|\sin\frac{(t-s)\xi}{2}\right|\end{aligned}$$

Положим $\delta := t-s$. Тогда

$$\left.\begin{aligned}&\left|\sin\tfrac{\delta\xi}{2}\right| \xrightarrow[\delta\to 0]{} 0\\ &\left|\sin\tfrac{\delta\xi}{2}\right| \le 1\end{aligned}\right\} \Rightarrow \text{По теореме Лебега: } \mathsf{E}\left|\sin\frac{\delta\xi}{2}\right| \to 0$$

а значит, что $\forall\varepsilon>0\ \exists\delta_0\ \forall\delta<\delta_0: \ \mathsf{E}\sin^2\frac{\delta\xi}{2}<\varepsilon$. Возьмем $|t-s|<\delta_0$ и все. □

Утверждение 11.1.3. $\varphi_\xi(x)$ принимает только действительные значения $\iff \xi \stackrel{d}{=} -\xi$ (т.е. распределение ξ симметрично относительно 0).

Доказательство.

$\Rightarrow$:

$\varphi_\xi(t)=\mathsf{E}\cos t\xi,\ \varphi_{-\xi}=\mathsf{E}\cos(-t\xi)=\mathsf{E}\cos t\xi$. Из теоремы о единственности следует, что $\xi\stackrel{d}{=}-\xi$.

$\Leftarrow$:

$\varphi_\xi(t)=\mathsf{E}\left(\cos t\xi+i\sin t\xi\right)=\varphi_{-\xi}(t)=\mathsf{E}\left(\cos t\xi-i\sin t\xi\right)\Rightarrow\mathsf{E}\sin t\xi=0\Rightarrow\varphi_\xi=\mathsf{E}\cos t\xi$ □

Утверждение 11.1.4. $\varphi_\xi(-t)=\overline{\varphi_\xi(t)}$ — очевидно по определению.

Утверждение 11.1.5. (б/д)

Пусть $\mathsf{E}|\xi|^n<+\infty$. Тогда $\varphi_\xi(t)$ n раз дифференцируема и

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

$$\varphi_\xi(t)=\sum_{k=0}^{n}\frac{(it)^k}{k!}\cdot\mathsf{E}\xi^k+\varepsilon_n(t)$$

где $|\varepsilon_n(t)|\le 3\mathsf{E}|\xi|^n,\ \varepsilon_n(t)\to 0$ при $t\to\infty$.

Утверждение 11.1.6. (б/д)

Если существует производная четного порядка $\varphi^{(2n)}$ в точке $x=0$, то существует $\mathsf{E}\xi^{2n}<+\infty$.

Утверждение 11.1.7. (б/д)

Пусть $\forall n\in\mathbb{N}\ \mathsf{E}|\xi|^n<+\infty$ и $\varlimsup\limits_{n\to\infty}\frac{\sqrt[n]{\mathsf{E}|\xi|^n}}{n}=T<\infty$. Тогда для любого $x\in(-\frac{1}{T};\ \frac{1}{T})$

$$\varphi_\xi(x)=\sum_{k=0}^{+\infty}\frac{(ix)^k}{k!}\cdot\mathsf{E}\xi^k$$

Теорема 11.4. *(о непрерывности, б/д)*

1. *Пусть* $\xi_n \xrightarrow{d} \xi$. *Тогда* $\forall t \in \mathbb{R}: \ \varphi_{\xi_n}(t) \to \varphi_\xi(t)$ *при* $n \to \infty$.

2. *Пусть* $\xi_1,\ \xi_2,\ \ldots$ *— случайные величины и* $\forall t \in \mathbb{R}: \ \varphi_{\xi_n}(t) \to \varphi(t)$ *при* $n \to \infty$, *причем* $\varphi(t)$ *непрерывна в* 0. *Тогда* $\varphi(t)$ *— характеристическая функция некоторой случайной величины* ξ, *причем* $\xi_n \xrightarrow{d} \xi$.

Теорема 11.5. *(Бохнера-Хинчина)*

Пусть φ *— непрерывна и* $\varphi(0) = 1$. *Тогда она является характеристической тогда и только тогда, когда она неотрицательно определена, то есть*

$$\forall n \in \mathbb{N}\ \forall z_1,\ z_2,\ \ldots,\ z_n \in \mathbb{C}\ \forall t_1,\ \ldots,\ t_n \in \mathbb{R}: \sum_{1 \le i,\ j \le n} \varphi(t_i - t_j) z_i \overline{z_j} \ge 0$$

(Это условие означает, что для любых $t_1, \ldots, t_n \in \mathbb{R}$ *матрица* $\begin{pmatrix} \varphi(t_1 - t_1) & \varphi(t_1 - t_2) & \ldots & \varphi(t_1 - t_n) \\ \varphi(t_2 - t_1) & \varphi(t_2 - t_2) & \ldots & \varphi(t_2 - t_n) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi(t_n - t_1) & \varphi(t_n - t_2) & \ldots & \varphi(t_n - t_n) \end{pmatrix}$ *неотрицательно определена.)*

Доказательство.

$\Rightarrow$:

Пусть φ — характеристическая. Тогда

$$\sum_{i,\ j} \mathsf{E} e^{i(t_i - t_j)\xi} z_i \overline{z_j} = \sum_{i,\ j} \mathsf{E} \left(z_i e^{i t_i \xi} \right) \left(\overline{z_j e^{i t_j \xi}} \right) = \mathsf{E} |\sum_j z_j e^{i t_j \xi}|^2 \ge 0$$

$\Leftarrow$:

Без доказательства. □

11.2 Гауссовские векторы

Определение 11.2. $\xi = (\xi_1,\ \ldots,\ \xi_n)^T$ — гауссовский случайный вектор, если

$$\varphi_\xi(x) = e^{i\langle a,\ x \rangle - \frac{1}{2}\langle \Sigma x,\ x \rangle}$$

где $\langle x,\ y \rangle$ — скалярное произведение, Σ — неотрицательно определенная симметричная матрица $n \times n$, $a \in \mathbb{R}^n$.

Теорема 11.6. *(эквивалентные определения гауссовского вектора)*

Следующие определения гауссовского вектора эквиваленты:

1. *Определение 11.2*

2. $\xi = A\eta + b$, *где* $\eta = (\eta_1,\, \ldots,\, \eta_m)^T$ *и* $\eta_i \sim N(0,\, 1)$, *независимые, а* $A \in Mat_{n\times m},\;\; b \in \mathbb{R}^n$

3. $\forall \lambda_1,\, \ldots,\, \lambda_n \in \mathbb{R}:\;\; \lambda_1\xi_1 + \ldots + \lambda_n\xi_n$ *— нормальная случайная величина.*

Доказательство.

1) $\Rightarrow$ 2):

$\varphi_\xi(x) = e^{i\langle a,\, x\rangle - \frac{1}{2}\langle \Sigma x,\, x\rangle}$. Тогда $\varphi_{\xi-a}(x) = \mathsf{E}e^{i\langle \xi - a,\, x\rangle} = e^{-i\langle a,\, x\rangle}\varphi_\xi(x) = e^{-1/2\langle \Sigma x,\, x\rangle}$. Так как Σ — симметричная, то $\Sigma = R^T D R$, где R — отрогональная, а D — диагональная с неотрицательными числами на главной диагонали (спектральное разложение). Тогда Σ можно представить в виде $\Sigma = R^T D R = R^T\sqrt{D}\sqrt{D}R = (\sqrt{D}R)^T(\sqrt{D}R)$, а значит

$$\langle \Sigma x,\, x\rangle = (\Sigma x)^T x = x^T \Sigma^T x = x^T(\sqrt{D}R)^T(\sqrt{D}R)x = (\sqrt{D}Rx)^T(\sqrt{D}Rx)$$

Сделаем замену $x = (\sqrt{D}R)^{-1}y$. Тогда $\varphi_{\xi-a}((\sqrt{D}R)^{-1}y) = e^{-1/2\langle y,\, y\rangle} = \prod\limits_{i=1}^{n} e^{-\frac{y_i^2}{2}}$. С другой стороны

$$\begin{aligned}
\varphi_{\xi-a}((\sqrt{D}R)^{-1}y) &= \mathsf{E}e^{i\langle \xi - a,\, (\sqrt{D}R)^{-1}y\rangle} \\
&= \mathsf{E}\Big[exp\Big(i(\xi-a)^T(\sqrt{D}R)^{-1}y\Big)\Big] \\
&= \mathsf{E}\Big[exp\Big(i(\xi-a)^T R^T(\sqrt{D})^{-1}y\Big)\Big] \\
&= \mathsf{E}\Big[exp\Big(i((\sqrt{D})^{-1}R(\xi-a))^T y\Big)\Big] = \varphi_{(\sqrt{D})^{-1}R(\xi-a)}(y)
\end{aligned}$$

то есть характеристическая функция вектора $\eta := (\sqrt{D})^{-1}R(\xi - a)$ это произведение характеристических функций стандартных нормальных распределений, а значит его компоненты $\sim N(0,\, 1)$.

2) $\Rightarrow$ 3):

Пусть $\xi = A\eta + b$. Тогда

$$\langle \lambda,\, \xi\rangle = \lambda^T(A\eta + b) = \lambda^T A\eta + \lambda^T b \sim N(\lambda^T b,\, \langle \lambda^T A,\, \lambda^T A\rangle)$$

3) $\Rightarrow$ 1):

Перепишем функцию распределения:

$$\varphi_\xi(x) = \mathsf{E}e^{i\langle \xi,\, x\rangle} = \mathsf{E}e^{i(x_1\xi_1 + \ldots + x_n\xi_n)} = \varphi_{x_1\xi_1 + \ldots + x_n\xi_n}(1) = e^{ia\cdot 1 - \frac{1}{2}\sigma^2\cdot 1^2}$$

где $a = \mathsf{E}(x_1\xi_1 + \ldots + x_n\xi_n)$, а $\sigma^2 = \mathsf{D}(x_1\xi_1 + \ldots + x_n\xi_n)$ Заметим, что $a = \langle (\mathsf{E}\xi_1,\, \ldots,\, \mathsf{E}\xi_n)^T,\, x\rangle$ и

$$0 \le \sigma^2 = \sum_{1\le j,\, k\le n} \mathrm{cov}(\xi_j x_j,\, \xi_k x_k) = \sum_{1\le k,\, j\le n} x_j x_k \mathrm{cov}(\xi_j,\, \xi_k) = \langle \Sigma x,\, x\rangle$$

где Σ — матрица ковариаций вектора ξ. $\square$

Следствие 11.6.1. *В условиях определения 11.2* a *— вектор мат. ожиданий, а* Σ *— матрица ковариаций.*

Следствие 11.6.2. *Пусть ξ — гауссовский вектор и* $\Sigma = \begin{pmatrix} \Sigma_1 & 0 & \dots & 0 \\ 0 & \Sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Sigma_k \end{pmatrix}$ *имеет блочно-диагональный вид. Обозначим за x_i, ξ^i, a_i соответственно те координаты векторов x, ξ, a, которые отвечают Σ_i в Σ. Тогда $\xi = (\xi^1, \dots, \xi^k)$, где ξ^i случайный вектор и $\xi^1, \dots, \xi^k$ независимы в совокупности.*

Доказательство.

$$\begin{aligned}
\varphi_\xi(x) &= e^{i\langle a,\, x\rangle - \frac{1}{2}\langle \Sigma x,\, x\rangle} \\
&= exp\left(i\langle a_1,\, x_1\rangle + \ldots + i\langle a_k,\, x_k\rangle - \frac{1}{2}\langle \Sigma_1 x_1,\, x_1\rangle - \ldots - \frac{1}{2}\langle \Sigma_k x_k,\, x_k\rangle\right) \\
&= \prod_{j=1}^{k} e^{i\langle a_j,\, x_j\rangle - \frac{1}{2}\langle \Sigma_j x_j,\, x_j\rangle} \\
&= \prod_{j=1}^{k} \varphi_{\xi_j}(x_j)
\end{aligned}$$

Применяя критерий независимости получаем требуемое. □

11.3 Центральная предельная теорема

Теорема 11.7. *(центральная предельная теорема)*

Пусть $\xi_1, \xi_2, \ldots$ — независимые одинаково распределенные случайные величины с конечным вторым моментом. Тогда

$$\frac{S_n - \mathsf{E}S_n}{\sqrt{\mathsf{D}S_n}} \xrightarrow{d} \eta \sim N(0,\, 1)$$

Доказательство. Рассмотрим $\tilde{S_n} = \frac{S_n - \mathsf{E}S_n}{\sqrt{\mathsf{D}S_n}}$, $\tilde{\xi_k} = \frac{\xi_k - \mathsf{E}\xi_k}{\sqrt{\mathsf{D}S_n}}$. Покажем, что характеристические функции $\tilde{S_n}$ сходятся к характеристической функции случайной величины $\sim N(0,\, 1)$

$$\begin{aligned}
\varphi_{\tilde{S_n}}(x) &= \varphi_{\sum\limits_{j=1}^{n} \tilde{\xi_j}}(x) \\
&= \prod_{j=1}^{n} \varphi_{\tilde{\xi_j}}(x) \qquad \text{(по независимости)} \\
&= \left(\mathsf{E}e^{ix\tilde{\xi_1}}\right)^n \qquad \text{(т.к. одинаково распределенные)} \\
&= \left(\mathsf{E}\left[exp\left(\frac{ix}{\sqrt{\mathsf{D}S_n}}(\xi_1 - \mathsf{E}\xi_1)\right)\right]\right)^n \\
&= \left(\varphi_{\xi_1 - \mathsf{E}\xi_1}\left(\frac{x}{\sqrt{\mathsf{D}S_n}}\right)\right)^n \\
&= \left(1 - \frac{1}{2}x^2\frac{\mathsf{D}\xi_1}{\mathsf{D}S_n} + o\left(\frac{1}{\mathsf{D}S_n}\right)\right)^n \\
&= \left(1 - \frac{1}{2n}x^2 + o\left(\frac{1}{n}\right)\right)^n \to e^{-\frac{x^2}{2}}
\end{aligned}$$

□

Теорема 11.8. *(многомерная ЦПТ, б/д)*

Пусть $\xi_1,\, \xi_2, \ldots$ *— независимые одинаково распределенные векторы, второй момент каждых компонент которых конечен.* Σ *— матрица ковариаций* ξ_1 *(существует т.к.* $\mathsf{E}\xi_i\xi_j \leq \sqrt{\mathsf{E}\xi_i^2 \cdot \mathsf{E}\xi_j^2}$*). Тогда*

$$\frac{S_n - \mathsf{E}S_n}{\sqrt{n}} \xrightarrow{d} \eta \sim N(\overline{0},\, \Sigma)$$

где сходимость по распределению случайных векторов означает сходимость в основном, определенную выше.

Теорема 11.9. *(локальная предельная теорема, б/д)*

Пусть $\xi_1,\, \xi_2,\, \ldots \sim Bern(p)$ *— независимые одинаково распределенные случайные величины. Пусть* $\varphi(n) = o(n^{2/3})$*. Тогда*

$$\sup_{k\colon |k-np|<\varphi(n)} \left| \frac{\mathsf{P}(S_n = k)}{\frac{1}{\sqrt{2\pi npq}} exp\left(-\frac{(k-np)^2}{2npq}\right)} - 1 \right| \to 0$$

(Теорема утверждает, что число успехов в таком случае имеет распределение, почти равное $N(np, npq)$*)*

Теорема 11.10. *(интегральная предельная теорема Муавра-Лапласа, б/д)*

Пусть $\xi_1,\, \xi_2,\, \ldots \sim Bern(p)$ *— независимые одинаково распределенные случайные величины. Тогда*

$$\sup_{-\infty<a<b<+\infty} \left| \mathsf{P}\left(a < \frac{S_n - np}{\sqrt{npq}} < b\right) - \int\limits_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \right| \to 0$$

(Теорема утверждает, что выполнена ЦПТ с равномерной сходимостью)

Пример 11.6. Игральный кубик бросается 6 млн раз. Тогда вероятность того, что количество выпавших троек отличается от 1 миллиона не более чем на 3000 примерно равна $0,999$.

Теорема 11.11. *(предельная теорема Пуассона)*

Пусть $\xi_1,\, \xi_2,\, \ldots,\, \xi_n \sim Bern(p_n)$ *— независимые случайные величины и* $pn \to \lambda \in \mathbb{R}$ *при* $n \to \infty$*. Тогда*

$$S_n \xrightarrow{d} \eta \sim Pois(\lambda)$$

Доказательство. Из условия $p_n = \frac{\lambda}{n} + o\left(\frac{\lambda}{n}\right)$. Тогда имеем

$$\begin{aligned} \mathsf{P}(S_n = k) = C_n^k p^k (1-p)^{n-k} &= \frac{n(n-1)\ldots(n-k+1)}{n^k}(np)^k \frac{1}{k!}(1-p)^n(1-p)^{-k} \\ &\to 1 \cdot \lambda^k \cdot \frac{1}{k!} \cdot e^{-\lambda} \cdot 1 \\ &= \frac{e^{-\lambda}\lambda^k}{k!} \end{aligned}$$

□