

Конспект лекций по курсу «Теория вероятностей»

Лектор: канд. физ.-мат. наук Родионов Игорь Владимирович

Набор: Алексей Шепелев

Содержание

1	Лекция от 10.02.2018	1
2	Лекция от 17.02.2018	4
3	Лекция от 03.03.2018	5
4	Лекция от 10.03.2018	5
5	Лекция от 17.03.2018	8
6	Лекция от 24.03.2018	11
7	Лекция от 31.03.2018	14
8	Лекция от 07.04.2018	18

1 Лекция от 10.02.2018

Будем обозначать вероятностное пространство как $(\Omega, \mathcal{F}, \mathbf{P})$, где

1. Ω — пространство элементарных исходов;
2. \mathcal{F} — σ -алгебра на Ω ;
3. \mathbf{P} — вероятностная, $\mathbf{P} : \mathcal{F} \rightarrow [0, 1]$, причем

а) $\mathbf{P}(\Omega) = 1$;

б) \mathbf{P} — σ -аддитивна, то есть $\forall \{A_n\}_{n=1}^{+\infty} \in \mathcal{F}$, причем $A_n \cap A_m = \emptyset$ при $n \neq m$:

$$\mathbf{P}\left(\bigsqcup_{n=1}^{+\infty} A_n\right) = \sum_{n=1}^{+\infty} \mathbf{P}(A_n).$$

Определение. Последовательность $\{A_n\}$ убывает к A , если $\forall n : A_n \supseteq A_{n+1}$ и $A = \bigcap_{n=1}^{+\infty} A_n$. Последовательность $\{A_n\}$ возрастает к A , если $\forall n : A_n \subseteq A_{n+1}$ и $A = \bigcup_{n=1}^{+\infty} A_n$.

Теорема (о непрерывности вероятностной меры). Пусть (Ω, \mathcal{F}) — измеримое пространство и на нем определена функция $P : \mathcal{F} \rightarrow [0, 1]$, удовлетворяющая следующим свойствам: $P(\Omega) = 1$ и P — конечно аддитивная. Тогда следующие утверждения эквивалентны:

1. P — вероятностная мера;
2. $\forall A \downarrow A : P(A_n) \rightarrow P(A)$ (непрерывность снизу);
3. $\forall A \uparrow A : P(A_n) \rightarrow P(A)$ (непрерывность сверху);
4. $\forall A \downarrow \emptyset : P(A_n) \rightarrow 0$ (непрерывность в нуле).

Теорема (Каратеодори). [б/д] Пусть Ω — некоторое множество, \mathcal{A} — σ -алгебра на Ω и P_σ — вероятностная мера на (Ω, \mathcal{A}) . Тогда существует единственная вероятностная мера на $(\Omega, \sigma(\mathcal{A}))$, являющаяся продолжением P_σ , то есть $\forall A \in \mathcal{A} : P_\sigma(A) = P(A)$.

Рассмотрим измеримое пространство $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ и вероятностную меру P на нем.

Определение. Функция $\mathcal{F}(x), x \in \mathbb{R}$, заданная по правилу $F(x) = P((-\infty, x])$ — функция распределения вероятностной меры P .

Лемма (свойство функции распределения). Пусть $F(x)$ — функция распределения, тогда

1. $F(x)$ не убывает;
2. $\lim_{x \rightarrow +\infty} F(x) = 1; \lim_{x \rightarrow -\infty} F(x) = 0$;
3. $F(x)$ непрерывна справа.

▲ Пусть $y \geq x$, тогда $F(y) - F(x) = P((-\infty, y]) - P((-\infty, x]) = P((x, y]) \geq 0$, следовательно, $F(x)$ неубывает.

Пусть $x_n \rightarrow -\infty$ при $n \rightarrow +\infty$, тогда $(-\infty, x_n] \rightarrow \emptyset$, следовательно, $F(x_n) = P((-\infty, x_n]) \xrightarrow{n \rightarrow +\infty} 0$ по теореме о непрерывности вероятностной меры.

Пусть $x_n \rightarrow +\infty$ при $n \rightarrow +\infty$, тогда $(-\infty, x_n] \rightarrow \mathbb{R}$, следовательно, $F(x_n) = P((-\infty, x_n]) \xrightarrow{n \rightarrow +\infty} P(\mathbb{R}) = 1$.

Пусть $x_n \downarrow x$, тогда $(-\infty, x_n] \downarrow (-\infty, x]$, откуда по теореме о непрерывности вероятностной меры вытекает, что $F(x_n) = P((-\infty, x_n]) \xrightarrow{n \rightarrow +\infty} P((-\infty, x]) = F(x)$. ■

Свойство 1. Функция распределения имеет предел слева $\forall x \in \mathbb{R}$, при этом число точек разрыва не более, чем счетно.

▲ Пусть $x_n \rightarrow x - 0$ — возрастающая последовательность, тогда $F(x_n) = P((-\infty, x_n]) \xrightarrow{n \rightarrow +\infty} P((-\infty, x_n]) = F(x - 0)$. Каждая точка разрыва — скачок функции распределения, каждому скачку сопоставим $[F(x - 0), F(x)]$, а этому отрезку в свою очередь сопоставим некую рациональную точку, которая лежит в $(F(x - 0), F(x))$. Следовательно каждому скачку мы сопоставили точку из \mathbb{Q} , а так как \mathbb{Q} счетно, то число разрывов не более, чем счетно. ■

Определение. Функция $F(x)$, которая удовлетворяет свойствам 1) – 3) из леммы, называется функцией распределения на \mathbb{R} .

Теорема (о взаимно однозначном соответствии между вероятностной мерой и функцией распределения на \mathbb{R}). Пусть $F(X)$ — функция распределения на \mathbb{R} , тогда существует единственная вероятностная мера P на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ такая, что $F(x)$ является ее функцией распределения, то есть $F(x) = P((-\infty, x])$.

▲ Рассмотрим полукольцо $S = \{(a, b]\}$ на \mathbb{R} . Определим σ -аддитивную вероятностную меру $P((a, b]) = F(b) - F(a)$, а по теореме P единственным образом продолжается на всю σ -алгебру $\mathcal{B}(\mathbb{R})$. ■

Классификация вероятностных мер и функций распределения на прямой

① Дискретное распределение

Пусть $\mathcal{X} \subseteq \mathbb{R}$ не более, чем счетно.

Определение. Вероятностная мера P на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, удовлетворяющая свойству $P(\mathbb{R} \setminus \mathcal{X}) = 0$, называется дискретной вероятностной мерой на \mathcal{X} , ее функция распределения также называется дискретной.

Рассмотрим $\mathcal{X} = \{x_k\}$, положим $p_k = P(\{x_k\})$, тогда $P(\mathcal{X}) = 1 = \sum_k P(x_k)$.

Определение. Набор чисел $\{p_k\}$ на \mathcal{X} называется распределением вероятностей на \mathcal{X} .

② Абсолютно непрерывное распределение

Определение. Пусть $F(x)$ — функция распределения вероятностной меры P на \mathbb{R} , причем $\forall x \in \mathbb{R}$ верно $F(x) = \int_{-\infty}^x p(t) dt$, где $p(t) \geq 0$, а $\int_{-\infty}^{+\infty} p(t) dt = 1$. Тогда P абсолютно непрерывна, $F(x)$ также называется абсолютно непрерывной, а $p(t)$ — плотность распределения $F(x)$. Причем $p(t)$ определена однозначно, кроме множества меры нуль.

Примеры:

1. Равномерное распределение $R[a, b]$

$$p(x) = \frac{1}{b-a} \cdot I(x \in [a, b]).$$

2. Нормальное (гауссовское) распределение $N(a, \sigma^2)$

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp \left[-\frac{(x-a)^2}{2\sigma^2} \right].$$

3. Экспоненциальное распределение $\text{Exp}(\alpha)$ ‘

$$p(x) = \alpha e^{-\alpha x} \cdot I(x > 0).$$

4. Распределение Коши $\text{Cauchy}(\theta)$

$$p(x) = \frac{\theta}{\pi (x^2 + \theta^2)}.$$

5. Гамма распределение $\Gamma(\alpha, \gamma)$

$$p(x) = \frac{x^{\alpha-1} \gamma^\alpha}{\Gamma(\alpha)} \cdot e^{-\gamma x} \cdot I(x > 0).$$

Определение. $\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx$, причем $\forall n \in \mathbb{N} : \Gamma(n) = (n-1)!, \forall \lambda \in \mathbb{R} :$
 $\Gamma(\lambda \pm 1) = \lambda \Gamma(\lambda)$, а $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

③ Сингулярные распределения

Определение. Пусть $F(x)$ — функция распределения на \mathbb{R} . Точка $x_0 \in \mathbb{R}$ называется точкой роста $F(x)$, если $\forall \varepsilon > 0 : F(x_0 + \varepsilon) - F(x_0 - \varepsilon) > 0$.

Определение. Функция распределения называется сингулярной, если она непрерывна и множество ее точек роста имеет Лебегову меру нуль. Например, функция Кантора.

Теорема (Лебега о функции распределения). [б/д] Пусть $F(x)$ — функция распределения на \mathbb{R} . Тогда существуют единственные α_1, α_2 и α_3 , $\alpha_i \geq 0, \alpha_1 + \alpha_2 + \alpha_3 = 1$ и функции распределения $F_1(x), F_2(x)$ и $F_3(x)$ такие, что $F(x) = \alpha_1 F_1(x) + \alpha_2 F_2(x) + \alpha_3 F_3(x)$, где $F_1(x)$ — дискретная функция распределения, $F_2(x)$ — абсолютно непрерывная, а $F_3(x)$ — сингулярная.

2 Лекция от 17.02.2018

Вероятностная мера в $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

Определение.

3 Лекция от 03.03.2018

Свойство 1.

Свойство 2.

Свойство 3.

Свойство 4.

Свойство 5.

Свойство 6.

4 Лекция от 10.03.2018

Свойство 7. Если $\xi = \eta$ почти наверное и $E|\eta| < +\infty$, то $E|\xi| < +\infty$ и $E\xi = E\eta$.

▲ Пусть $A = \{\xi \neq \eta\}$, тогда $I_A = 0$ почти наверное, следовательно $\xi \cdot I_A = 0$ почти наверное и $\eta \cdot I_A = 0$ почти наверное. Так как $\xi = \xi \cdot I_A + \xi \cdot I_{\bar{A}}$, то $\xi = \xi \cdot I_A + \eta \cdot I_{\bar{A}}$, потому что на \bar{A} выполняется $\xi = \eta$. Из свойства 6 имеем $E\xi = E(\xi \cdot I_A) + E(\eta \cdot I_{\bar{A}}) = E(\eta \cdot I_A) + E(\eta \cdot I_{\bar{A}}) = E\eta$. ■

Свойство 8. Пусть $\xi \geq 0$ и $E\xi \geq 0$, тогда $\xi = 0$ почти наверное.

▲ Рассмотрим события $A = \{\xi > 0\}$ и $A_n = \{\xi > \frac{1}{n}\}$, следовательно, $A_n \uparrow A$. Имеем $P(A_n) = EI_{A_n}$, так как $n\xi > 1$ на A_n , то $EI_{A_n} \leq E(n\xi \cdot I_A) \leq nE\xi = 0$, значит, $P(A) = \lim_{n \rightarrow +\infty} P(A_n)$. ■

Свойство 9. Пусть $E\xi$ и $E\eta$ конечны, $\forall A \in \mathcal{F} : E(\xi \cdot I_A) \leq E(\eta \cdot I_A)$. Тогда $\xi \leq \eta$ почти наверное.

▲ Рассмотрим событие $B = \{\xi > \eta\}$. Из условия и построения B получаем, что $E(\eta \cdot I_B) \leq E(\xi \cdot I_B) \leq E(\eta \cdot I_B)$, следовательно, $E(\xi \cdot I_B) = E(\eta \cdot I_B)$, значит $E((\xi - \eta) \cdot I_B) = 0$. Так как $(\xi - \eta) \cdot I_B \geq 0$, то по свойству 8 $(\xi - \eta) \cdot I_B = 0$ почти наверное, следовательно $I_B = 0$ почти наверное, потому что $\xi - \eta > 0$ на B . ■

Теорема (о математическом ожидании произвольной случайной величины). Пусть $\xi \perp \eta$, причем $E\xi$ и $E\eta$ конечны, тогда $E\xi\eta$ конечно и $E\xi\eta = E\xi \cdot E\eta$.

▲ Пусть ξ и η — простые случайные величины, то есть ξ принимает значения $\{x_1, \dots, x_n\}$, η принимает значения $\{y_1, \dots, y_n\}$. Тогда по линейности

$$\begin{aligned} E\xi\eta &= \sum_{k,j=1}^n x_k y_j P(\xi = x_k, \eta = y_j) = \sum_{k,j=1}^n x_k y_j P(\xi = x_k) \cdot P(\eta = y_j) = \\ &= \sum_{k=1}^n x_k P(\xi = x_k) \sum_{j=1}^n y_j P(\eta = y_j) = E\xi \cdot E\eta. \end{aligned}$$

Рассмотрим $\xi_n \uparrow \xi$, $\xi_n = \sum_{k=0}^{n \cdot 2^n - 1} \frac{k}{2^n} I\left(\frac{k}{2^n} \leq \xi \leq \frac{k+1}{2^n}\right) + nI(\xi > n)$, следовательно, $\xi_n = \varphi_n(\xi)$, значит, ξ_n — \mathcal{F}_ξ -измеримая. Пусть $\xi, \eta \geq 0$. Существует последовательность \mathcal{F}_ξ -измеримых (\mathcal{F}_η -измеримых) простых неотрицательных простых функций $\xi_n \uparrow \xi$ ($\eta_n \uparrow \eta$). Так как $\xi \perp \eta$, то $\xi_n = \varphi_n(\xi) \perp \varphi_n(\eta) = \eta_n$. Следовательно, $\xi_n \cdot \eta_n \uparrow \xi \cdot \eta$, а по определению математического ожидания $E\xi\eta = \lim_{n \rightarrow +\infty} E(\xi_n \eta_n) = \lim_{n \rightarrow +\infty} E\xi_n \cdot E\eta_n = E\xi \cdot E\eta$.

Пусть теперь ξ и η — произвольные случайные величины. ξ^+ и ξ^- — функции от ξ , η^+ и η^- — функции от η , следовательно, $\xi^+ \perp \eta^+$ и $\xi^- \perp \eta^-$, отсюда $(\xi\eta)^+ = \xi^+ \eta^+ + \xi^- \eta^-$ значит, $E(\xi\eta)^+ = E\xi^+ \eta^+ + E\xi^- \eta^- = E\xi^+ E\eta^+ + E\xi^- E\eta^-$, аналогично $E(\xi\eta)^- = E\xi^+ \eta^- + E\xi^- \eta^+ = E\xi^+ E\eta^- + E\xi^- E\eta^+$. Осталось заметить, что $E\xi\eta = E(\xi\eta)^+ - E(\xi\eta)^- = E\xi^+ E\eta^+ + E\xi^- E\eta^- - E\xi^+ E\eta^- - E\xi^- E\eta^+ = (E\xi^+ - E\xi^-)(E\eta^+ - E\eta^-) = E\xi \cdot E\eta$. ■

Пусть $\xi = \sum_{i=1}^n x_i \cdot I(\xi = x_i)$ — простая случайная величина. Тогда $Eg(\xi) = \sum_{i=1}^n g(x_i) \cdot P(\xi = x_i) = \sum_{i=1}^n g(x_i) \Delta F_\xi(x_i)$, где $\Delta F_\xi(x_i) = F_\xi(x_i) - F_\xi(x_i - 0)$.

Теорема (о замене переменной в интеграле Лебега). [6/д] Пусть (Ω, \mathcal{F}) и (E, \mathcal{E}) — два измеримых пространства и $X = X(\omega)$ — $\mathcal{F}|\mathcal{E}$ -измеримая функция со значениями в E , то есть $\forall B \in \mathcal{E} : X^{-1}(B) \in \mathcal{F}$. Пусть P — вероятностная мера на (Ω, \mathcal{F}) и P_X — вероятностная мера на (E, \mathcal{E}) , заданная по правилу $P_X(A) = P(\omega : X(\omega) \in A)$ для $A \in \mathcal{E}$. Тогда для любой \mathcal{E} -измеримой функции $g(x), x \in E$, то есть $\forall B \in \mathcal{E} : g^{-1}(B) \in \mathcal{E}$, верно, $\int_A g(x) P_X(dx) = \int_{X^{-1}(A)} g(X(\omega)) P(d\omega)$.

Пусть $\xi : \Omega \rightarrow \mathbb{R}(\mathbb{R}^n)$, в таком случае вероятностная мера P_ξ однозначно восстанавливается по F_ξ , следовательно, по теореме $Eg(\xi) = \int g(\xi) dP = \int g(x) P_\xi(dx) = \int g(x) dF_\xi(x)$.

Пусть ξ — абсолютно непрерывная случайная величина с плотностью $p_\xi(x)$, тогда $dF_\xi(x) = p_\xi(x) dx$, следовательно $Eg(x) = \int_{\mathbb{R}} g(x) p_\xi(x) dx$.

Прямое произведение вероятностных пространств и формула свертки

Определение. Пусть $(\Omega_1, \mathcal{F}_1, P_1)$ и $(\Omega_2, \mathcal{F}_2, P_2)$ — два вероятностных пространства. Тогда (Ω, \mathcal{F}, P) — их прямое произведение, если

1. $\Omega = \Omega_1 \times \Omega_2$;
2. $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$, то есть $\mathcal{F} = \sigma\{\{B_1 \times B_2\} | B_1 \in \mathcal{F}_1, B_2 \in \mathcal{F}_2\}$;
3. $P = P_1 \otimes P_2$, то есть P — продолжение вероятностной меры $P_1 \times P_2$, заданное на прямоугольнике $B_1 \times B_2, B_1 \in \mathcal{F}_1, B_2 \in \mathcal{F}_2$ по правилу $P(B_1 \times B_2) = P_1(B_1) \cdot P_2(B_2)$. Так как $\{B_1 \times B_2\}$ — полукольцо, то P существует и единственна по теореме Каратеодори.

Теорема (Фубини). [б/д] Пусть (Ω, \mathcal{F}, P) — прямое произведение вероятностных пространств $(\Omega_1, \mathcal{F}_1, P_1)$ и $(\Omega_2, \mathcal{F}_2, P_2)$. Пусть $\xi : \Omega \rightarrow \mathbb{R}$ такая, что $\int_{\Omega} |\xi(\omega_1, \omega_2)| dP < +\infty$. Тогда интегралы $\int_{\Omega_1} \xi(\omega_1, \omega_2) P_1(d\omega_1)$ и $\int_{\Omega_2} \xi(\omega_1, \omega_2) P_2(d\omega_2)$ определены почти наверное относительно P_2 и P_1 соответственно, являются измеримыми случайными величинами относительно \mathcal{F}_2 и \mathcal{F}_1 , следовательно,

$$\int_{\Omega} \xi(\omega_1, \omega_2) dP = \int_{\Omega_2} \int_{\Omega_1} \xi(\omega_1, \omega_2) P_1(d\omega_1) P_2(d\omega_2) + \int_{\Omega_1} \int_{\Omega_2} \xi(\omega_1, \omega_2) P_2(d\omega_2) P_1(d\omega_1).$$

Из всего этого следует, что двойной интеграл равен повторному.

Утверждение. Пусть $\xi \perp \eta$ — случайные величины, тогда $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), P_{(\xi, \eta)}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_{\xi}) \otimes (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_{\eta})$.

▲ Достаточно проверить свойство прямого произведения:

1. $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$;
2. $\mathcal{B}(\mathbb{R}^2) = \sigma(\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}))$ по определению борелевской σ -алгебры в \mathbb{R}^2 ;
3. $P_{(\xi, \eta)}(B_1 \times B_2) = P(\xi \in B_1, \eta \in B_2) = P(\xi \in B_1) \cdot P(\eta \in B_2) = P_{\xi}(B_1) \cdot P_{\eta}(B_2)$. ■

Лемма (о свертке). Пусть случайные величины ξ и η независимы с функциями распределения F_{ξ} и F_{η} . Тогда

$$F_{\xi+\eta}(z) = \int_{\mathbb{R}} F_{\xi}(z-x) dF_{\eta}(x) = \int_{\mathbb{R}} F_{\eta}(z-x) dF_{\xi}(x).$$

Если ξ и η имеют плотности распределения f_{ξ} и f_{η} соответственно, то $\xi + \eta$ имеет плотность распределения

$$f_{\xi+\eta}(z) = \int_{\mathbb{R}} f_{\xi}(z-x) f_{\eta}(x) dx = \int_{\mathbb{R}} f_{\eta}(z-x) f_{\xi}(x) dx$$

▲ Заметим, $F_{\xi+\eta}(z) = P(\xi + \eta \leq z)$, а по теореме о замене переменных в интеграле Лебега это равно $\int_{\mathbb{R}^2} I(x+y \leq z) P_{\xi}(dx) P_{\eta}(dy)$, полученный двойной интеграл по Фубини можно записать как повторный:

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} I(x+y \leq z) P_{\xi}(dx) \right) P_{\eta}(dy) = \int_{\mathbb{R}} \left(\int_{-\infty}^{z-y} P_{\xi}(dx) \right) P_{\eta}(dy) = \int_{\mathbb{R}} F_{\xi}(z-y) dF_{\eta}(y).$$

Перейдем ко второму пункту доказательства:

$$\begin{aligned} F_{\xi+\eta}(z) &= \int_{\mathbb{R}^2} I(x+y \leq z) P_{\xi}(dx) P_{\eta}(dy) = \int_{\mathbb{R}^2} I(x+y \leq z) f_{\xi}(x) f_{\eta}(y) dx dy \stackrel{t=x+y}{=} \\ &\stackrel{t=x+y}{=} \int_{\mathbb{R}^2} I(t \leq z) f_{\xi}(x) f_{\eta}(t-x) dx dt = \int_{-\infty}^z \left(\int_{\mathbb{R}} f_{\xi}(x) f_{\eta}(t-x) dx \right) dt. \end{aligned}$$

Следовательно, по определению плотности, $f_{\xi+\eta} = \int_{\mathbb{R}} f_{\xi}(x)f_{\eta}(t-x) dx$. ■

5 Лекция от 17.03.2018

Дисперсия и ковариация

Определение. Дисперсией случайной величины ξ называется $D\xi = E(\xi - E\xi)^2$, если $E\xi < +\infty$. Очевидно, $D\xi \geq 0$.

Определение. Ковариация двух случайных величин называется $\text{cov}(\xi, \eta) = E((\xi - E\xi)(\eta - E\eta))$. Легко заметить, что $\text{cov}(\xi, \xi) = D\xi$. Если $\text{cov}(\xi, \eta) = 0$, то случайные величины ξ и η называются некоррелированными.

Определение. Величина $\rho(\xi, \eta) = \frac{\text{cov}(\xi, \eta)}{\sqrt{D\xi \cdot D\eta}}$ называется коэффициентом корреляции случайных величин ξ и η при условии, что $D\xi$ и $D\eta$ не равны нулю и конечны.

Свойства ковариации и дисперсии

Свойство 1 (Билинейность ковариации). $\text{cov}(a\xi + b\zeta, \eta) = a \text{cov}(\xi, \eta) + b \text{cov}(\zeta, \eta)$

Свойство 2. $\text{cov}(\xi, \eta) = E\xi\eta - E\xi \cdot E\eta \Rightarrow D\xi = E\xi^2 - (E\xi)^2$

▲ $\text{cov}(\xi, \eta) = E(\xi - E\xi)(\eta - E\eta) = E\xi\eta - E((E\xi) \cdot \eta) - E((E\eta) \cdot \xi) + E\xi \cdot E\eta = E\xi\eta - E\xi \cdot E\eta$ ■

Свойство 3. Пусть $c \in \mathbb{R}$, тогда $D(c\xi) = c^2 D\xi$, $D(\xi + c) = D\xi$, $Dc = 0$.

▲ $D(c\xi) = E c^2 \xi^2 - (E c \xi)^2 = c^2 E \xi^2 - c^2 (E \xi)^2 = c^2 D\xi$;
 $D(c + \xi) = E(c + \xi - E(c + \xi))^2 = E(c + \xi - c - E\xi)^2 = D\xi$;
 $Dc = E(c - Ec)^2 = E(c - c)^2 = 0$. ■

Свойство 4 (Неравенство Коши-Буняковского). $|E\xi\eta|^2 \leq E\xi^2 \cdot E\eta^2$

▲ Рассмотрим для $\lambda \in \mathbb{R}$ функцию $f(\lambda) = E(\xi - \lambda\eta)^2 \geq 0$. Имеем $f(\lambda) = E\xi^2 + 2\lambda E\xi\eta + \lambda^2 E\eta^2 \geq 0$. Для выполнения неравенства дискриминант полученного многочлена должен быть меньше нуля: $D = 4E\xi\eta - 4E\xi^2 E\eta^2 \leq 0$, откуда следует неравенство. ■

Свойство 5. $|\rho(\xi, \eta)| \leq 1$, причем $\rho(\xi, \eta) = \pm 1 \Leftrightarrow \xi = a\eta + b$ почти наверное.

▲ Рассмотрим случайные величины $\xi_1 = \xi - E\xi$ и $\eta_1 = \eta - E\eta$, следовательно $\rho(\xi, \eta) = \frac{E\xi_1\eta_1}{\sqrt{E\xi_1^2 \cdot E\eta_1^2}} \leq 1$ по неравенству Коши-Буняковского. Пусть $|\rho(\xi, \eta)| = 1$, тогда дискриминант $D = 0$, следовательно, $\exists! \lambda_0 : f(\lambda_0) = 0$, то есть $E(\xi_1 + \lambda_0\eta_1)^2 = 0$, отсюда $(\xi_1 + \lambda_0\eta_1)^2 = 0$ почти наверное, а, значит, и $\xi_1 + \lambda_0\eta_1 = 0$ почти наверное. Теперь можно заключить, что $\xi = E\xi - \lambda_0(\eta - E\eta)$. ■

Свойство 6. Если $\xi \perp \eta$, то $\text{cov}(\xi, \eta) = 0$, обратное неверное.

▲ $\text{cov}(\xi, \eta) = E\xi\eta - E\xi \cdot E\eta$, но так как $\xi \perp \eta$, то $E\xi\eta = E\xi \cdot E\eta$, следовательно, $\text{cov}(\xi, \eta) = 0$. ■

Лемма. Пусть ξ_1, \dots, ξ_n — попарно некоррелированные случайные величины (например, независимые в совокупности), $D\xi_1, \dots, D\xi_n < +\infty$, тогда $D(\xi_1, \dots, \xi_n) = D\xi_1 + \dots + D\xi_n$.

▲

$$D\left(\sum_{i=1}^n \xi_i\right) = \text{cov}\left(\sum_{i=1}^n \xi_i, \sum_{j=1}^n \xi_j\right) = \sum_{i,j=1}^n \text{cov}(\xi_i, \xi_j).$$

По условию, если $i \neq j$, то $\text{cov}(\xi_i, \xi_j) = 0$, следовательно

$$D\left(\sum_{i=1}^n \xi_i\right) = \sum_{i=1}^n \text{cov}(\xi_i, \xi_i) = \sum_{i=1}^n D\xi_i.$$

■

Многомерный случай

Определение. Пусть $\vec{\xi} = (\xi_1, \dots, \xi_n)$ — случайный вектор, тогда его математическим ожиданием называется вектор из математических ожиданий его компонент, то есть $E\vec{\xi} = (E\xi_1, \dots, E\xi_n)$.

Определение. Матрицей ковариаций случайного вектора $\vec{\xi}$ называется

$$\text{Var } \vec{\xi} = \begin{pmatrix} \text{cov}(\xi_1, \xi_1) & \cdots & \text{cov}(\xi_1, \xi_n) \\ \vdots & \ddots & \vdots \\ \text{cov}(\xi_n, \xi_1) & \cdots & \text{cov}(\xi_n, \xi_n) \end{pmatrix} = \|\text{cov}(\xi_i, \xi_j)\|_{i,j=1}^n.$$

Лемма. Матрица ковариаций случайного вектора — симметрическая и неотрицательно определенная¹.

▲ Матрица $\text{Var } \vec{\xi} = \|\text{cov}(\xi_i, \xi_j)\|_{i,j=1}^n$ — симметрическая, так как $r_{ij} \equiv \text{cov}(\xi_i, \xi_j) = \text{cov}(\xi_j, \xi_i) \equiv r_{ji}$. Пусть $\vec{x} \in \mathbb{R}^n$, тогда

$$\begin{aligned} \vec{x}^T \text{Var } \vec{\xi} \vec{x} &= (\vec{x}, \text{Var } \vec{\xi} \vec{x}) = \sum_{i,j=1}^n \text{cov}(\xi_i, \xi_j) x_i x_j = \sum_{i,j=1}^n \text{cov}(x_i \xi_i, x_j \xi_j) = \\ &= \text{cov}\left(\sum_{i=1}^n x_i \xi_i, \sum_{j=1}^n x_j \xi_j\right) = \text{cov}\left(\sum_{i=1}^n x_i \xi_i, \sum_{i=1}^n x_i \xi_i\right) = D\left(\sum_{i=1}^n x_i \xi_i\right) \geq 0. \end{aligned}$$

■

¹Матрица A неотрицательно определена, если $\forall \vec{x} \in \mathbb{R}^n : \vec{x}^T A \vec{x} \geq 0$

Неравенства

Лемма (Неравенство Маркова). Пусть $\xi \geq 0$ — случайная величина, $E\xi < +\infty$ (существует). Тогда $\forall \varepsilon > 0 : P(\xi \geq \varepsilon) \leq \frac{E\xi}{\varepsilon}$.

▲ $P(\xi \geq \varepsilon) = EI(\xi \geq \varepsilon)$. На множестве $\xi \geq \varepsilon$ случайная величина $\frac{\xi}{\varepsilon} \geq 1$, следовательно $EI(\xi \geq \varepsilon) \leq E\left(\frac{\xi}{\varepsilon} \cdot I(\xi \geq \varepsilon)\right) \leq \frac{1}{\varepsilon} \cdot E\xi$. ■

Лемма (Неравенство Чебышёва). Пусть ξ — случайная величина такая, что $D\xi < +\infty$, тогда $\forall \varepsilon > 0 : P(|\xi - E\xi| \geq \varepsilon) \leq \frac{D\xi}{\varepsilon^2}$.

▲ $P(|\xi - E\xi| \geq \varepsilon) = P(|\xi - E\xi|^2 \geq \varepsilon^2)$. Из неравенства Маркова имеем, что $P(|\xi - E\xi|^2 \geq \varepsilon^2) \leq \frac{E(\xi - E\xi)^2}{\varepsilon^2} = \frac{D\xi}{\varepsilon^2}$. ■

Лемма (Неравенство Йенсена). Пусть $g(x)$ — борелевская выпуклая вниз (вверх) функция и $E\xi < +\infty$. Тогда $Eg(\xi) \geq g(E\xi)$ ($Eg(\xi) \leq g(E\xi)$).

▲ Так как $g(x)$ выпукла вниз, то $\forall x_0 \in \mathbb{R} : g(x) \geq g(x_0) + \lambda(x_0)(x - x_0)$. Положим $x = \xi$ и $x_0 = E\xi$, тогда $g(\xi) \geq g(E\xi) + \lambda(E\xi)(\xi - E\xi)$, считая математическое ожидание от обеих частей неравенства, получаем $Eg(\xi) \geq g(E\xi) + 0$. ■

Определение. Пусть ξ и $\{\xi_i\}_{i=1}^{+\infty}$ — случайные величины, тогда $\xi_n \xrightarrow{P} \xi$ сходится по вероятности, если $\forall \varepsilon > 0 : P(\omega : |\xi_n(\omega) - \xi(\omega)| > \varepsilon) \rightarrow 0$ при $n \rightarrow +\infty$.

Теорема (Закон больших чисел в форме Чебышёва). Пусть $\{\xi_1, \dots, \xi_n\}_{i=1}^{+\infty}$ — последовательность попарно некоррелированных случайных величин таких, что $\forall n \in \mathbb{N} : D\xi_n \leq C$. Обозначим $S_n = \sum_{i=1}^n \xi_i$, тогда $\frac{S_n - ES_n}{n} \xrightarrow{P} 0$ при $n \rightarrow +\infty$.

▲ По неравенству Чебышёва $P\left|\frac{S_n - ES_n}{n}\right| > \varepsilon \geq \frac{D(S_n - ES_n)}{n^2\varepsilon^2}$, по свойству дисперсии с сдвиге это равно $\frac{DS_n}{n^2\varepsilon^2}$. Применяя лемму о дисперсии суммы, получаем $\frac{\sum_{i=1}^n D\xi_i}{n^2\varepsilon^2} \rightarrow 0$. ■

Следствие. Пусть $\{\xi_n\}_{i=1}^{+\infty}$ — независимые случайные величины такие, что $\forall n \in \mathbb{N} : D\xi_n \leq C \wedge E\xi_n = a$. Тогда $\frac{S_n}{n} \xrightarrow{P} a$ при $n \rightarrow +\infty$.

Условные математические ожидания (УМО)

Пусть (Ω, \mathcal{F}, P) — вероятностное пространство; $\xi : \Omega \rightarrow \mathbb{R}$ — случайная величина; $\mathcal{F}_\xi = \{\xi^{-1}(B), B \in \mathcal{B}(\mathbb{R})\}$ — σ -алгебра, порожденная ξ . Если \mathcal{G} — под σ -алгебра σ -алгебры \mathcal{F} , то ξ называется \mathcal{G} -измеримой, если $\mathcal{F}_\xi \subset \mathcal{G}$.

Определение. Пусть ξ — случайная величина на (Ω, \mathcal{F}, P) , \mathcal{G} — под σ -алгебра \mathcal{F} . Условным математическим ожиданием случайной величины ξ относительно \mathcal{G} называется случайная величина $E(\xi|\mathcal{G})$, обладающая следующими свойствами:

1. $E(\xi|\mathcal{G})$ является σ -измеримой случайной величиной;
2. $\forall A \in \mathcal{G} : E(\xi \cdot I_A) = E(E(\xi|\mathcal{G}) \cdot I_A)$ или, что тоже самое, $\int_A \xi dP = \int_A E(\xi|\mathcal{G}) dP$.

Обозначаем $E(\xi|\eta) \equiv E(\xi|\mathcal{F}_\eta)$, если такая η существует.

Определение. Пусть (Ω, \mathcal{F}, P) — вероятностное пространство. Функция множеств $\nu : \mathcal{F} \rightarrow \mathbb{R}$ — заряд (мера со знаком), если ν — σ -аддитивна на \mathcal{F} , то есть $\nu\left(\bigcup_{i=1}^{+\infty} A_i\right) = \sum_{i=1}^{+\infty} \nu(A_i)$ для $\{A_i\}_{i=1}^{+\infty} \in \mathcal{F}$, ряд в правой части сходится абсолютно и $\sup_{A \in \mathcal{F}} |\nu(A)| < +\infty$.

Определение. Заряд ν называется абсолютно непрерывным относительно меры P , если $\forall A \in \mathcal{F} : (P(A) = 0 \Rightarrow \nu(A) = 0)$.

Теорема (Радона-Никодима). [б/д] Пусть (Ω, \mathcal{F}, P) — вероятностное пространство, ν — заряд на \mathcal{F} , абсолютно непрерывный относительно меры P . Тогда существует и единственна случайная величина η на (Ω, \mathcal{F}, P) такая, что $E\eta < +\infty$ и $\nu(A) = \int_A \eta dP = E\eta \cdot I_A$.

6 Лекция от 24.03.2018

Лемма (о существовании УМО). Пусть ξ — случайная величина с $E|\xi| < +\infty$. Тогда $\forall \mathcal{G} \subset \mathcal{F}$ (под σ -алгебра) : $E(\xi|\mathcal{G})$ существует и единственна почти наверное.

▲ Рассмотрим вероятностное пространство (Ω, \mathcal{G}, P) . Положим, что $\forall A \in \mathcal{G} : Q(A) = \int_A \xi dP = E(\xi \cdot I_A)$, следовательно, $Q(A)$ — заряд на (Ω, \mathcal{G}, P) , абсолютно непрерывный относительно меры P . Тогда по теореме Радона-Никодима существует и единственна почти наверное случайная величина η на (Ω, \mathcal{G}, P) с $E\eta < +\infty$ такая, что $Q(A) = \int_A \eta dP$. Значит, η — УМО. Действительно, η \mathcal{G} -измерима и $\forall A \in \mathcal{G} : \int_A \eta dP = \int_A \xi dP$. ■

Теорема. Пусть σ -алгебра \mathcal{G} порождена разбиением $\Omega \setminus \{D_n\}_{n=1}^{+\infty}$, причем, $P(D_n) > 0$. Тогда, если $E\xi < +\infty$, то $E(\xi|\mathcal{G}) = \sum_{n=1}^{+\infty} \frac{E(\xi \cdot I(D_n))}{P(D_n)} \cdot I(D_n)$.

▲ Пусть η \mathcal{G} -измерима. Покажем, что $\eta = \sum_{n=1}^{+\infty} c_n I_{D_n}(\omega)$. Пусть $\eta \neq \text{const}$ на D_n , тогда $\exists a \neq b : \{\omega : \eta(\omega) = a\} \cap D_n \neq \emptyset$ и $\{\omega : \eta(\omega) = b\} \cap D_n \neq \emptyset$, следовательно, $\{\omega : \eta(\omega) = a\} \cap D_n = D_n$ и $\{\omega : \eta(\omega) = b\} \cap D_n \neq D_n$, иначе $\{\omega : \eta(\omega) = a\} \notin \mathcal{G}$, то есть η не \mathcal{G} -измерима. Получили противоречие.

Найдем $c_n : E(\xi|\mathcal{G}) = \sum_{n=1}^{+\infty} c_n I_{D_n}$, так как $E(\xi|\mathcal{G})$ \mathcal{G} -измерима по определению.

$$E(\xi \cdot I_{D_n}) = E(E(\xi|\mathcal{G}) \cdot I_{D_n}) = E\left(\sum_{m=1}^{+\infty} c_m I_{D_m} I_{D_n}\right) = E(c_n I_{D_n}) = c_n P(D_n).$$

Следовательно, $c_n = \frac{E(\xi \cdot I_{D_n})}{P(D_n)}$. ■

Свойства УМО

Свойство УМО: если $\forall A \in \mathcal{F} : E(\xi \cdot I_A) = E(\eta \cdot I_A)$, то $\xi = \eta$ гочти наверное на (Ω, \mathcal{F}, P) .

Свойство 1. Если ξ \mathcal{G} -измерима, то $E(\xi|\mathcal{G}) = \xi$ почти наверное.

▲ ξ удовлетворяет свойствам УМО: первому по условию, а второму, поскольку $\int_A \xi dP = \int_A \xi dP$. Следовательно, $E(\xi|\mathcal{G}) = \xi$ почти наверное. ■

Свойство 2 (формула полной вероятности). $E(E(\xi|\mathcal{G})) = E\xi$.

▲ Так как $\Omega \in \mathcal{G}$, то по интегральному свойству $E(E(\xi|\mathcal{G})) = E(E(\xi|\mathcal{G}) \cdot I_\Omega) = E(\xi \cdot I_\Omega) = E\xi$. ■

Свойство 3 (линейность). $E(\alpha\xi + \beta\eta|\mathcal{G}) = \alpha E(\xi|\mathcal{G}) + \beta E(\eta|\mathcal{G})$.

▲ $\alpha E(\xi|\mathcal{G}) + \beta E(\eta|\mathcal{G})$ \mathcal{G} -измерима. Осталось проверить интегральное свойство:

$$\begin{aligned} \forall A \in \mathcal{G} : \int_A (\alpha E(\xi|\mathcal{G}) + \beta E(\eta|\mathcal{G})) dP &= \alpha \int_A E(\xi|\mathcal{G}) dP + \beta \int_A E(\eta|\mathcal{G}) dP = \\ &= \alpha \int_A \xi dP + \beta \int_A \eta dP = \int_A (\alpha\xi + \beta\eta) dP = \int_A E(\alpha\xi + \beta\eta) dP \end{aligned}$$

■

Свойство 4. Пусть ξ не зависит от \mathcal{G} , то есть $\mathcal{F}_\xi \perp \mathcal{G}$. Тогда $E(\xi|\mathcal{G}) = E\xi$ почти наверное.

▲ Пусть $\xi \perp \mathcal{G}$, что равносильно $\forall A \in \mathcal{G} : \xi \perp I_A$. $E\xi$ — константа, следовательно, она измерима относительно \mathcal{G} , так как $\mathcal{F}_{E\xi} = \{\Omega, \emptyset\}$. Интегральное свойство УМО: $E(\xi \cdot I_A) = \boxed{E(E(\xi|\mathcal{G}) \cdot I_A)} = E\xi \cdot P(A) = \boxed{E(E(\xi) \cdot I_A)}$, следовательно, $E\xi = E(\xi|\mathcal{G})$. ■

Свойство 5. Пусть $\xi \leq \eta$ почти наверное, тогда $E(\xi|\mathcal{G}) \leq E(\eta|\mathcal{G})$ почти наверное.

▲ $\xi \leq \eta$ почти наверное, следовательно, $\forall A \in \mathcal{G} \int_A \xi dP \leq \int_A \eta dP$, что равносильно $\int E(\xi|\mathcal{G}) dP \leq \int E(\eta|\mathcal{G}) dP$, а из свойств математического ожидания вытекает, что $E(\xi|\mathcal{G}) \leq E(\eta|\mathcal{G})$ почти наверное. ■

Свойство 6. $|E(\xi|\mathcal{G})| \leq E(|\xi||\mathcal{G})$.

▲ $-|\xi| \leq \xi \leq |\xi|$. ■

Свойство 7 (телескопическое свойство). Пусть $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{F}$, тогда

1. $E(E(\xi|\mathcal{G}_1)|\mathcal{G}_2) = E(\xi|\mathcal{G}_1)$ почти наверное,
2. $E(E(\xi|\mathcal{G}_2)|\mathcal{G}_1) = E(\xi|\mathcal{G}_1)$ почти наверное.

▲ (1) $E(\xi|\mathcal{G}_1)$ \mathcal{G}_2 -измерима, следовательно, по первому свойству $E(E(\xi|\mathcal{G}_1)|\mathcal{G}_2) = E(\xi|\mathcal{G}_1)$. (2) Пусть $A \in \mathcal{G}_1$, следовательно, $A \in \mathcal{G}_2$.

$$E(E(\xi|\mathcal{G}_1) \cdot I_A) = E(\xi \cdot I_A) = E(E(\xi|\mathcal{G}_2) \cdot I_A) = E(E(E(\xi|\mathcal{G}_2)|\mathcal{G}_1) \cdot I_A).$$

По свойству математического ожидания $E(\xi|\mathcal{G}_1) = E(E(\xi|\mathcal{G}_2)|\mathcal{G}_1)$. ■

Свойство 8. [б/д] Пусть $\forall n > 1 : |\xi_n| \leq \eta$, $E\eta < +\infty$ и $\xi_n \xrightarrow{n.n.} \xi$. Тогда $\forall \mathcal{G} \subset \mathcal{F} : E(\xi_n|\mathcal{G}) \xrightarrow{n.n.} E(\xi|\mathcal{G})$.

Свойство 9. Пусть η \mathcal{G} -измерима, $E|\xi\eta| < +\infty$, $E|\xi| < +\infty$, $E|\eta| < +\infty$. Тогда $E(\xi\eta|\mathcal{G}) = \eta E(\xi|\mathcal{G})$ почти наверное.

▲ Пусть $\eta = I_B$, где $B \in \mathcal{G}$. Тогда

$$\begin{aligned} \forall A \in \mathcal{G} : E(E(\xi\eta|\mathcal{G}) \cdot I_A) &= E(\xi\eta \cdot I_A) = E(\xi I_B I_A) = \\ &= E(\xi I_{A \cap B}) = E(E(\xi|\mathcal{G}) \cdot I_{A \cap B}) = E(\eta E(\xi|\mathcal{G}) \cdot I_A). \end{aligned}$$

Следовательно, $E(\xi\eta|\mathcal{G}) = \eta E(\xi|\mathcal{G})$ почти наверное по свойству математического ожидания.

Так как доказали для индикаторов, то доказали и для любой простой функции. Теперь пусть η — произвольная случайная величина. Возьмем последовательно простых \mathcal{F}_η -измеримых случайных величин $\eta_n : |\eta_n| \leq |\eta|$ и $\eta_n \xrightarrow{n.n.} \eta$. По свойству 8 $E(\xi\eta|\mathcal{G}) = \lim_{n \rightarrow \infty} E(\xi\eta_n|\mathcal{G}) \xrightarrow{n.n.} E(\xi\eta|\mathcal{G})$, то есть $E(\xi\eta|\mathcal{G}) = \eta E(\xi|\mathcal{G})$ почти наверное. ■

Теорема (о наилучшем квадратичном прогнозе). Пусть ξ — случайная величина, \mathcal{G} — под-алгебра \mathcal{F} . Обозначим $\mathcal{A}_{\mathcal{G}} = \{\eta | \eta - \mathcal{G}\text{-измеримая сл. вел.}\}$. Тогда $\inf_{\eta \in \mathcal{A}_{\mathcal{G}}} E(\xi - \eta)^2 = E(\xi - E(\xi|\mathcal{G}))^2$.

▲ Пусть $\eta \in \mathcal{A}_{\mathcal{G}}$, тогда

$$\begin{aligned} E(\xi - \eta)^2 &= E(\xi - E(\xi|\mathcal{G}) + E(\xi|\mathcal{G}) - \eta)^2 = \\ &= E(\xi - E(\xi - E(\xi|\mathcal{G})))^2 + E(E(\xi|\mathcal{G}) - \eta)^2 + 2E((\xi - E(\xi|\mathcal{G}))(E(\xi|\mathcal{G}) - \eta)). \end{aligned}$$

Пусть $\varkappa \equiv \xi - E(\xi|\mathcal{G})$, $\psi \equiv E(\xi|\mathcal{G}) - \eta$. Рассмотрим $E(\varkappa\psi)$, по свойству 2 это равно $E(E(\varkappa\psi|\mathcal{G}))$, а по свойству 3, это можно переписать, как $E(\psi E(\varkappa|\mathcal{G}))$. Но $E(\varkappa|\mathcal{G}) = E((\xi - E(\xi|\mathcal{G}))|\mathcal{G}) = 0$, следовательно, $E(\varkappa\psi) = 0$. Значит $E(\xi - \eta)^2 = E(\xi - E(\xi - E(\xi|\mathcal{G})))^2 + E(E(\xi|\mathcal{G}) - \eta)^2 \geq E(\xi - E(\xi|\mathcal{G}))^2$. ■

7 Лекция от 31.03.2018

Условные распределения

Определение. Пусть $A \in \mathcal{F}$, тогда по определению $P(A|\mathcal{G}) = E(I_A|\mathcal{G})$, $\mathcal{G} \subset \mathcal{F}$. Если ξ, η — случайные величины на (Ω, \mathcal{F}, P) , то $E(\xi|\eta) = E(\xi|\mathcal{F}_{\eta})$.

Определение. Величиной $E(\xi|\eta = y)$ называется такая борелевская функция $\varphi(y)$, что $\forall B \in \mathcal{B}(\mathbb{R}) : E(\xi \cdot I(\eta \in B)) = \int_B \varphi(y) P_{\eta}(dy)$.

Лемма. Если $E\xi$ существует, то $E(\xi|\eta = y)$ существует и единственно почти наверное относительно P_{η} .

▲ Рассмотрим $\psi(B) = E(\xi \cdot I(\eta \in B))$ — заряд на $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_{\eta})$, потому что $\psi(B)$ σ -аддитивна по свойству интеграла Лебега и конечна, так как $E(\xi) < +\infty$. ψ абсолютно непрерывна относительно P_{η} , так как если $P_{\eta}(B) = 0$, то $I(\eta \in B) = 0$ почти наверное, следовательно, $E(\xi \cdot I(\eta \in B)) = 0$, а, значит, выполнены условия теоремы Радона-Никодима, то есть существует и единственна почти наверное случайная величина φ на $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_{\eta})$ (борелевская функция) такая, что $\psi(B) = \int_B \varphi(y) P_{\eta}(dy)$. ■

Лемма. $E(\xi|\eta = y) = \varphi(y)$ тогда и только тогда, когда $E(\xi|\eta) = \varphi(\eta)$ почти наверное.

▲ Пусть $B \in \mathcal{B}(\mathbb{R})$, тогда $E(E(\xi|\eta) \cdot I(\eta \in B)) = E(\xi \cdot I(\eta \in B)) = \int_B \varphi(y) P_{\eta}(dy)$.

По теореме о замене переменных в интеграле Лебега это можно переписать, как $\int_{\{\eta \in B\}} \varphi(\eta) dP = E(\varphi(\eta) \cdot I(\eta \in B))$, что равносильно условию $E(\xi|\eta) = \varphi(\eta)$ почти наверное по Свойству. Обратно аналогично, по тем же равенствам. ■

Следствие. Пусть ξ — \mathcal{F}_η -измеримая случайная величина, тогда существует борелевская функция $\psi(x)$ такая, что $\xi = \psi(x)$ почти наверное.

▲ Так как ξ — \mathcal{F}_η -измеримая, то по свойству 1 $\xi = E(\xi|\eta)$ почти наверное. С другой стороны, так как существует единственная $\psi(x) : \psi(x) = E(\xi|\eta = x)$, то $\xi = E(\xi|\eta) = \psi(\eta)$. ■

Определение. Условным распределением случайной величины ξ при условии $\eta = y$ называется вероятностная мера $P(\xi \in B|\eta = y) = E(I(\xi \in B)|\eta = y)$. Является мерой на $\mathcal{B}(R)$.

Определение. Условной плотностью случайной величины ξ относительно η называется плотность условного распределения $P(\xi \in B|\eta = y)$, то есть функция $f_{\xi|\eta}(x|y)$ такая, что $P(\xi \in B|\eta = y) = \int_B f_{\xi|\eta}(x|y) dx$.

Теорема (о свойстве условной плотности). Пусть существует условная плотность случайной величины ξ относительно случайной величины η $f_{\xi|\eta}(x|y)$. Тогда для любой борелевской функции $g(x)$ такой, что $E|g(x)|$ существует, выполнено $E(g(\xi)|\eta = y) = \int_{\mathbb{R}} g(x) f_{\xi|\eta}(x|y) dx$ относительно P_η почти наверное.

▲ Пусть $B \in \mathcal{B}(\mathbb{R})$, пусть также $g(x) = I_A(x)$, $A \in \mathcal{B}(\mathbb{R})$. Тогда

$$\begin{aligned} \int_{\mathbb{R}} g(x) \cdot f_{\xi|\eta}(x|y) dx &= \int_{\mathbb{R}} I_A(x) \cdot f_{\xi|\eta}(x|y) dx = \int_A f_{\xi|\eta}(x|y) dx = \\ &= P(\xi \in A|\eta = y) = E(I(\xi \in A)|\eta = y) = E(g(\xi)|\eta = y). \end{aligned}$$

Так как доказали для индикаторов, то доказали и для всех простых функций $g(x)$. Далее с помощью теоремы Лебега для условных математических ожиданий доказываем для всех $g(x)$. $(E(\xi_n|\eta) \xrightarrow{п.н.} E(\xi|\eta))$, где $\xi_n \xrightarrow{п.н.} \xi$, ξ_n — простые) ■

Теорема (о виде условной плотности). Пусть ξ и η — случайные величины такие, что существует их совместная плотность $f_{(\xi,\eta)}(x,y)$. Пусть $f_\eta(y)$ — плотность случайной величины η , тогда функция

$$\varphi(x,y) = \frac{f_{(\xi,\eta)}(x,y)}{f_\eta(y)} \cdot I(f_\eta(y) > 0)$$

есть условная плотность $f_{\xi|\eta}(x|y)$.

▲ Для любых $A \in \mathcal{B}(\mathbb{R})$, $B \in \mathcal{B}(\mathbb{R})$ выполнено

$$P(\xi \in B, \eta \in A) = \int_{B \times A} f_{(\xi,\eta)}(x,y) dx dy = \int_A \left(\int_B \frac{f_{(\xi,\eta)}(x,y)}{f_\eta(y)} dx \right) f_\eta(y) dy,$$

с другой стороны

$$P(\xi \in B, \eta \in A) = E(I(\xi \in B, \eta \in A)) = \int_{\{\eta \in A\}} I(\xi \in B) dP.$$

Далее по интегральному свойству получаем, что

$$P(\xi \in B, \eta \in A) = \int_{\{\eta \in A\}} E(I(\xi \in B) | \eta) dP,$$

заменяя переменные, окончательно имеем следующее:

$$\begin{aligned} P(\xi \in B, \eta \in A) &= \int_A E(I(\xi \in B) | \eta = y) P_\eta(dy) = \\ &= \int_A P(\xi \in B | \eta = y) P_\eta dy = \int_A P(\xi \in B | \eta = y) f_\eta(y) dy. \end{aligned}$$

■

Алгоритм подсчета УМО

1. Найти совместную плотность $f_{(\xi, \eta)}(x, y)$, затем $f_\eta(y) = \int_{\mathbb{R}} f_{(\xi, \eta)}(x, y) dx$, тогда условная плотность $f_{\xi|\eta}(x|y) = \frac{f_{(\xi, \eta)}(x, y)}{f_\eta(y)}$.
2. Вычислить $\varphi(y) = E(g(\xi) | \eta = y) = \int_{\mathbb{R}} g(x) f_{\xi|\eta}(x|y) dx$.
3. Тогда $E(g(x) | \eta) = \varphi(\eta)$.

Виды сходимости случайных величин

Определение. Последовательность $\{\xi_n\}_{n \geq 1}$ сходится к случайной величине ξ

1. по вероятности ($\xi_n \xrightarrow{P} \xi$), если $\forall \varepsilon > 0 : P(\omega : |\xi_n(\omega) - \xi(\omega)| \geq \varepsilon) \xrightarrow{n \rightarrow +\infty} 0$,
2. почти наверное ($\xi_n \xrightarrow{\text{п.н.}} \xi$), если $P(\omega : \xi_n \rightarrow \xi) = 1$,
3. в L_p ($\xi_n \xrightarrow{L_p} \xi$), если $E|\xi_n|^p < +\infty$, $E|\xi|^p < +\infty$ и $E|\xi_n - \xi|^p \xrightarrow{n \rightarrow +\infty} 0$ ($p > 0$),
4. по распределению ($\xi_n \xrightarrow{d} \xi$), если для любой непрерывной ограниченной функции $f(x)$ выполнено $E f(\xi_n) \xrightarrow{n \rightarrow +\infty} E f(\xi)$.

Теорема (Александрова). $[b/\partial] \xi_n \xrightarrow{d} \xi$ тогда только тогда, когда $F_{\xi_n}(x) \xrightarrow{\text{в основном}} F_\xi(x)$, то есть $F_{\xi_n}(x) \rightarrow F_\xi(x)$ во всех точках непрерывности функции распределения $F_\xi(x)$.

Лемма (критерий сходимости почти наверное). $\xi_n \xrightarrow{n.н.} \xi$ тогда и только тогда, когда $\forall \varepsilon > 0 : P(\omega : \sup_{k \geq n} |\xi_k(\omega) - \xi(\omega)| \geq \varepsilon) \xrightarrow{n \rightarrow +\infty} 0$.

▲ Пусть $A_k^\varepsilon = \{\omega : |\xi_k - \xi| \geq \varepsilon\}$, $A^\varepsilon = \bigcap_{n=1}^{+\infty} \bigcup_{k \geq n} A_k^\varepsilon = \{\omega : \forall n \exists k \geq n : |\xi_n - \xi| \geq \varepsilon\}$.

Тогда $\{\omega : \xi_n(\omega) \not\xrightarrow{n.н.} \xi(\omega)\} = \bigcup_{m=1}^{+\infty} A_m^{\frac{1}{m}} = \{\omega : \exists m \forall n \exists k \geq n : |\xi_k(\omega) - \xi(\omega)| > \varepsilon\}$.

Следовательно,

$$\begin{aligned} P(\omega : \xi_n(\omega) \not\xrightarrow{n.н.} \xi(\omega)) = 0 &\Leftrightarrow P\left(\bigcup_{m=1}^{+\infty} A_m^{\frac{1}{m}}\right) = 0 \Leftrightarrow \\ &\Leftrightarrow \forall m \in \mathbb{N} : P\left(A_m^{\frac{1}{m}}\right) = 0 \Leftrightarrow \forall \varepsilon > 0 : P(A^\varepsilon) = 0, \end{aligned}$$

так как всегда существует m , что $\frac{1}{m} \geq \varepsilon \geq \frac{1}{m+1}$, то есть $A_{m+1}^{\frac{1}{m+1}} \supseteq A^\varepsilon \supseteq A_m^{\frac{1}{m}}$. Но $\bigcup_{k \geq n} A_k^\varepsilon \downarrow A^\varepsilon$, следовательно,

$$\begin{aligned} 0 = P(A^\varepsilon) &= \lim_{n \rightarrow +\infty} P\left(\bigcup_{k \geq n} A_k^\varepsilon\right) \Leftrightarrow \forall \varepsilon > 0 : P\left(\bigcup_{k \geq n} A_k^\varepsilon\right) \xrightarrow{n \rightarrow +\infty} 0 \Leftrightarrow \\ &\Leftrightarrow \forall \varepsilon > 0 : P(\omega : \sup_{k \geq n} |\xi_k(\omega) - \xi(\omega)| \geq \varepsilon) \xrightarrow{n \rightarrow +\infty} 0. \end{aligned}$$

■

Теорема (взаимоотношения различных видов сходимости).

$$\begin{array}{c} n.н. \\ \searrow \\ L_p \nearrow P \rightarrow d \end{array}$$

▲ (п.н. \Rightarrow P) $\xi_n \xrightarrow{n.н.} \xi \Leftrightarrow \forall \varepsilon > 0 : P(\omega : \sup_{k \geq n} |\xi_k(\omega) - \xi(\omega)| > \varepsilon) \rightarrow 0$, но

$$\{\omega : |\xi_n(\omega) - \xi(\omega)| > \varepsilon\} \subset \{\omega : \sup_{k \geq n} |\xi_k(\omega) - \xi(\omega)| \geq \varepsilon\},$$

следовательно, $P(|\xi_n - \xi| \geq \varepsilon) \rightarrow 0$.

($L_p \Rightarrow$ P) $P(|\xi_n - \xi| \geq \varepsilon) = P(\omega : |\xi_n(\omega) - \xi(\omega)|^p > \varepsilon^p)$, а по неравенству Маркова это меньше или равно $\frac{E|\xi_n(\omega) - \xi(\omega)|^p}{\varepsilon^p} \xrightarrow{n \rightarrow +\infty} 0$.

(P \Rightarrow d) Пусть $f(x)$ — ограниченная непрерывная функция, тогда $\exists C \in \mathbb{R} \forall x \in \mathbb{R} : |f(x)| \leq C$. Зафиксируем $\varepsilon > 0$, возьмем $N \in \mathbb{R} : P(|\xi| > N) \leq \frac{\varepsilon}{4C}$. На отрезке $[-N, N]$ $f(x)$ равномерно непрерывна, следовательно,

$$\exists \delta > 0 \forall x, y \in \mathbb{R} : (|x - y| < \delta \Rightarrow |f(x) - f(y)| < \frac{\varepsilon}{2}).$$

Рассмотрим разбиение Ω :

$$\begin{aligned} A_1 &= \{\omega : |\xi(\omega)| < N, |\xi_n(\omega) - \xi(\omega)| \leq \delta\}, \\ A_2 &= \{\omega : |\xi(\omega)| > N, |\xi_n(\omega) - \xi(\omega)| \leq \delta\}, \\ A_3 &= \{\omega : |\xi_n(\omega) - \xi(\omega)| > \delta\}. \end{aligned}$$

Оценим

$$|Ef(\xi_n) - Ef(\xi)| \leq E|f(\xi_n) - f(\xi)| = E[|f(\xi) - f(\xi_n)| \cdot (I_{A_1} + I_{A_2} + I_{A_3})] \leq \boxed{\quad}.$$

Пусть $\omega \in A_1$, тогда $|f(\xi_n) - f(\xi)| \leq \frac{\varepsilon}{2}$, следовательно, $E[|f(\xi_n) - f(\xi)| \cdot I_{A_1}] \leq \frac{\varepsilon}{2} \cdot EI_{A_1} = \frac{\varepsilon}{2} \cdot P(A_1) \leq \frac{\varepsilon}{2}$. Если же $\omega \in A_2, A_3$, то $|f(\xi_n) - f(\xi)| \leq 2C$.

$$\text{Значит, } \boxed{\leq} \frac{\varepsilon}{2} + 2C \cdot P(A_2) + 2C \cdot P(A_3) \leq \frac{\varepsilon}{2} + 2C \cdot P(|\xi| > N) + 2C \cdot P(|\xi_n - \xi| > \delta) \leq C_1 \varepsilon. \quad \xrightarrow{0, \text{ т.к. } \xi_n \xrightarrow{P} \xi}$$

Следовательно, $Ef(\xi_n) \rightarrow Ef(\xi)$, то есть $\xi_n \xrightarrow{d} \xi$. ■

8 Лекция от 07.04.2018

Контрпримеры

Пример (п.н. $\not\Rightarrow L_p$, а значит, $P \not\Rightarrow L_p$ и $d \not\Rightarrow L_p$). Рассмотрим $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}([0, 1])$, $P = \lambda$. Пусть $\xi_n = e^n \cdot I_{[0, \frac{1}{n}]}$, $\xi = 0$, тогда $\xi_n \xrightarrow{\text{п.н.}} \xi$, но $E|\xi_n - \xi|^p = e^{np} \cdot \frac{1}{n} \rightarrow +\infty$.

Пример ($L_p \not\Rightarrow$ п.н., $P \not\Rightarrow$ п.н., $d \not\Rightarrow$ п.н.). Рассмотрим $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}([0, 1])$, $P = \lambda$. Возьмем $\xi_{2^n+i} = I\left(\omega \in \left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)\right)$, $i = 0, \dots, 2^n - 1$; $n \in \mathbb{Z}_+$. Тогда $\xi_k \xrightarrow{L_p} 0$ при $k \rightarrow +\infty$, так как $E|\xi_k|^p = \frac{1}{2^n}$, где $n = [\log_2 k]$. Но для любой точки из $[0, 1]$ существует бесконечно много ξ_i таких, что $\xi_i(\omega) = 1$ и $\xi_i(\omega) = 0$, следовательно, $\forall \omega : \xi_i(\omega) \xrightarrow{i \rightarrow +\infty} 0$.

Пример ($d \not\Rightarrow P$). Пусть $\Omega = \{\omega_1, \omega_2\}$, $P(\omega_i) = \frac{1}{2}$, $\forall n \in \mathbb{Z}_+ : \xi_n(\omega_1) = 1, \xi_n(\omega_2) = 0$. Тогда $\xi_n \sim \text{Bern}\left(\frac{1}{2}\right)$. $\xi(\omega_1) = 1, \xi(\omega_2) = 0$, значит, $\xi \sim \text{Bern}\left(\frac{1}{2}\right)$, следовательно, по теореме Александрова $\xi_n \xrightarrow{d} \xi$, но $P(|\xi_n - \xi| > 0.5) = 1$, значит, $\xi_n \not\xrightarrow{P} \xi$.

Определение. Последовательность чисел $\{x_n\}$ называется фундаментальной, если $|x_n - x_m| \rightarrow 0$ при $n, m \rightarrow +\infty$.

Теорема (критерий Коши сходимости числовой последовательности). *[б/д] Последовательность чисел $\{x_n\}$ сходится тогда и только тогда, когда $\{x_n\}$ фундаментальна.*

Теорема (критерий Коши сходимости почти наверное). *Последовательно случайных величин $\{\xi_n\}$ сходится почти наверное тогда и только тогда, когда $\{\xi_n\}$ фундаментальна почти наверное, то есть $P(\omega : |\xi_n(\omega) - \xi(\omega)| \rightarrow 0) = 1$ при $n, m \rightarrow +\infty$.*

▲ (\Rightarrow) Пусть $\xi_n \xrightarrow{п.н.} \xi$, тогда, если $\omega \in \{\omega : \xi_n(\omega) - \xi(\omega)\}$, то $\omega \in \{\omega : \{\xi_n\} - \text{фундаментальная}\}$, следовательно, $P(\omega : \{\xi_n(\omega)\} - \text{фундаментальная}) \geq P(\omega : \xi_n(\omega) \rightarrow \xi(\omega)) = 1$.

(\Leftarrow) Обозначим $A = \{\omega : \{\xi_n\} - \text{фундаментальная}\}$. Построим такую случайную величину ξ , что $\xi_n \xrightarrow{п.н.} \xi$. По критерию Коши для любого $\omega \in A$ у последовательности $\{\xi_n(\omega)\}$ существует предел $\xi(\omega)$. Положим по определению $\xi(\omega) = \lim_{n \rightarrow +\infty} \xi_n(\omega) \cdot I_A(\omega)$. Тогда $\xi_n \cdot I_A \rightarrow \xi$, то есть ξ — случайная величина, как предел случайных величин, и $P(\omega : \xi_n(\omega) \rightarrow \xi(\omega)) = P(A) = 1$. ■

Лемма (критерий фундаментальности почти наверное). *[б/д] Последовательность случайных величин $\{\xi_n\}$ фундаментальна почти наверное тогда и только тогда, когда $\forall \varepsilon > 0 : P(\omega : \sup_{k \geq n} |\xi_k(\omega) - \xi_n(\omega)| > \varepsilon) \xrightarrow{n \rightarrow +\infty} 0$.*

Определение. Пусть $\{A_n\}_{n \in \mathbb{N}}$ — последовательность событий, тогда событием $\{A_n \text{ бесконечно часто (б.ч.)}\}$ называется событие $\{\omega : \forall n \exists k \geq n : \omega \in A_k\}$, то есть все такие ω , что ω принадлежит бесконечному числу элементов из $\{A_n\}_{n \in \mathbb{N}}$.
 $\{A_n \text{ б.ч.}\} = \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k$.

Лемма (Бореля-Кантелли). 1. Если $\sum_{k=1}^{\infty} P(A_k) < +\infty$, то $P(A_n \text{ б.ч.}) = 0$.

2. Если $\sum_{k=1}^{\infty} P(A_k) = +\infty$ и $\{A_k\}$ независимы в совокупности, то $P(A_n \text{ б.ч.}) = 1$.

▲ $P(A_n \text{ б.ч.}) = P\left(\bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k\right) \stackrel{[1]}{=} \square$. Известно, что $\bigcup_{k \geq n} A_k \downarrow \{A_n \text{ б.ч.}\}$, следо-

вательно, по непрерывности вероятностной меры имеем $\square \stackrel{[2]}{=} \lim_{n \rightarrow \infty} P\left(\bigcup_{k \geq n} A_k\right) \leq$

$\lim_{n \rightarrow \infty} \sum_{k \geq n} P(A_k) = 0$.

Заметим, что $P(A_n \text{ б.ч.}) = \lim_{n \rightarrow \infty} P\left(\bigcup_{k \geq n} A_k\right) = \lim_{n \rightarrow \infty} \left(1 - P\left(\bigcup_{k \geq n} \overline{A_k}\right)\right)$, но

$$\begin{aligned} P\left(\bigcup_{k \geq n} \overline{A_k}\right) &= \lim_{N \rightarrow \infty} P\left(\bigcap_{k=n}^N \overline{A_k}\right) = \lim_{N \rightarrow \infty} \prod_{k=n}^N P(\overline{A_k}) = \lim_{N \rightarrow \infty} \prod_{k=n}^N [1 - P(A_k)] \leq \\ &\leq \lim_{N \rightarrow \infty} \prod_{k=n}^N P(\overline{A_k}) = \lim_{N \rightarrow \infty} \exp\left(-\sum_{k=n}^N P(A_k)\right) = \exp\left(-\sum_{k=n}^{\infty} P(A_k)\right). \end{aligned}$$

■

Теорема (Рисса). *Есть последовательность случайных величин $\{\xi_n\}$ фундаментальна (или сходится) по вероятности, то из нее можно выделить подпоследовательность $\{\xi_{n_k}\}$ фундаментальную (сходящуюся) почти наверное.*

▲ Пусть $\{\xi_n\}$ фундаментальна по вероятности, то есть $\forall \varepsilon : \mathbf{P}(|\xi_k - \xi_n| > \varepsilon) \xrightarrow{n, k \rightarrow \infty} 0$. Докажем, что можно выделить подпоследовательность $\{\xi_{n_k}\}$, сходящуюся почти наверное. Пусть $n_1 = 1$. По индукции определим n_k , как наименьшее $n > n_{k-1}$ такое, что $\forall s \geq n, t \geq n : \mathbf{P}(|\xi_t - \xi_s| > 2^{-k}) < 2^{-k}$. Тогда $\sum_{k=1}^{\infty} \mathbf{P}(|\xi_{n_{k+1}} - \xi_{n_k}| > 2^{-k}) < \sum_{k=1}^{\infty} 2^{-k} < +\infty$, следовательно, по лемме Бореля-Кантелли $\mathbf{P}(|\xi_{n_{k+1}} - \xi_{n_k}| > 2^{-k} \text{ б.ч.}) = 0$, значит, почти наверное $\sum_{k=1}^{+\infty} |\xi_{n_{k+1}} - \xi_{n_k}| < +\infty$. Пусть $N = \left\{ \omega : \sum_{n=1}^{+\infty} |\xi_{n_{k+1}}(\omega) - \xi_{n_k}(\omega)| = +\infty \right\}$, тогда $\mathbf{P}(N) = 0$. Положим $\xi(\omega) = \left(\xi_{n_1}(\omega) + \sum_{k=1}^{+\infty} (\xi_{n_{k+1}}(\omega) - \xi_{n_k}(\omega)) \right) \cdot I_{\bar{N}}(\omega)$. Получаем, $\sum_{j=1}^k (\xi_{n_{j+1}} - \xi_{n_j}) + \xi_{n_1} = \xi_{n_{k+1}} \xrightarrow{\text{п.н.}} \xi$.

Пусть теперь $\xi_n \xrightarrow{\mathbf{P}} \xi$, тогда

$$\mathbf{P}(|\xi_m - \xi_n| \geq \varepsilon) \leq \mathbf{P}\left(|\xi_n - \xi| \geq \frac{\varepsilon}{2}\right) + \mathbf{P}\left(|\xi_m - \xi| \geq \frac{\varepsilon}{2}\right) \xrightarrow{n, m \rightarrow \infty} 0.$$

Следовательно, из сходимости по вероятности следует фундаментальность по вероятности, а дальше все тоже самое. ■