Конспект лекций по курсу «Теория вероятностей»

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

Набор: Алексей Шепелев

Содержание

1 Лекция от 10.02.2018

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

1. Ω — пространство элементарных исходов;
2. $\mathcal{F}$ — σ-алгебра на Ω;
3. P — вероятностная, $\mathsf{P} : \mathcal{F} \to [0, 1]$, причем
   a) $\mathsf{P}(\Omega) = 1$;
   b) P — σ-аддитивна, то есть $\forall \{A_n\}_{n=1}^{+\infty} \in \mathcal{F}$, причем $A_n \cap A_m = \varnothing$ при $n \neq m$: $\mathsf{P}\left(\bigsqcup\limits_{n=1}^{+\infty} A_n\right) = \sum\limits_{n=1}^{+\infty} \mathsf{P}(A_n)$.

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

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

1. P — *вероятностная мера;*
2. $\forall A \downarrow A : \mathsf{P}(A_n) \to \mathsf{P}(A)$ *(непрерывность снизу);*
3. $\forall A \uparrow A : \mathsf{P}(A_n) \to \mathsf{P}(A)$ *(непрерывность сверху);*
4. $\forall A \downarrow \varnothing : \mathsf{P}(A_n) \to 0$ *(непрерывность в нуле).*

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

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

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

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

1. *$F(x)$ не убывает;*
2. $\lim\limits_{x \to +\infty} F(x) = 1$; $\lim\limits_{x \to -\infty} F(x) = 0$;
3. *$F(x)$ непрерывна справа.*

▲ Пусть $y \geqslant x$, тогда $F(y) - F(x) = \mathsf{P}\big((-\infty, y]\big) - \mathsf{P}\big((-\infty, x]\big) = \mathsf{P}\big((x, y]\big) \geqslant 0$, следовательно, $F(x)$ неубывает.

Пусть $x_n \to -\infty$ при $n \to +\infty$, тогда $(-\infty, x_n] \to \varnothing$, следовательно, $F(x_n) = \mathsf{P}\big((-\infty, x_n]\big) \xrightarrow[n \to +\infty]{} 0$ по теореме о непрерувности вероятностной меры.

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

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

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

▲ Пусть $x_n \to x - 0$ — возрастающая последовательность, тогда $F(x_n) = \mathsf{P}\big((-\infty, x_n]\big) \underset{n\to+\infty}{\longrightarrow} \mathsf{P}\big((-\infty, x_n]\big) = F(x-0)$. 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 $F(x)$, котороя 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 $\big(\mathbb{R}, \mathscr{B}(\mathbb{R})\big)$ такая, что $F(x)$ является ее функцией распределения, то есть $F(x) = \mathsf{P}\big(-\infty, x]\big)$.*

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

Классификация вероятностных мер и функций распределения на прямой

① Дискретное распределение

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

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

Рассмотрим $\mathscr{X} = \{x_k\}$, положим $p_k = \mathsf{P}\big(\{x_k\}\big)$, тогда $\mathsf{P}(\mathscr{X}) = 1 = \sum\limits_k \mathsf{P}(x_k)$.

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

② Абсолютно непрерывное распределение

Определение. Пусть $F(x)$ — функция распределения вероятностной меры P на $\mathbb{R}$, причем $\forall x \in \mathbb{R}$ верно $F(x) = \int\limits_{-\infty}^{x} p(t)\,dt$, где $p(t) \geqslant 0$, а $\int\limits_{-\infty}^{+\infty} p(t)\,dt = 1$. 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 $R[a, b]$

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

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 $\mathrm{Exp}(\alpha)$ ‘

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

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

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

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. $\Gamma(\alpha) = \int\limits_0^{+\infty} x^{\alpha-1} e^{-x}\, dx$, причем $\forall n \in \mathbb{N} : \Gamma(n) = (n-1)!$, $\forall \lambda \in \mathbb{R} :$ $\Gamma(\lambda \pm 1) = \lambda\Gamma(\lambda)$, а $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

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, $\alpha_i \geqslant 0, \alpha_1 + \alpha_2 + \alpha_3 = 1$ 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 $\big(\mathbb{R}, \mathscr{B}(\mathbb{R})\big)$

Определение.

3 Лекция от 03.03.2018

4 Лекция от 10.03.2018

Свойство 2. *Если* $\xi = \eta$ *почти наверное и* $\mathsf{E}|\eta| <= +\infty$*, то* $\mathsf{E}|\xi| < +\infty$ *и* $\mathsf{E}\xi = \mathsf{E}\eta$.

▲ 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 $\xi = \xi \cdot I_A + \xi \cdot I_{\overline{A}}$, то $\xi = \xi \cdot I_A + \eta \cdot I_A$, потому что на $\overline{A}$ выполняется $\xi = \eta$. Из свойства 6 имеем $\mathsf{E}\xi = \mathsf{E}(\xi \cdot I_A) + \mathsf{E}(\eta \cdot I_{\overline{A}}) = \mathsf{E}(\eta \cdot I_A) + E(\eta \cdot I_{\overline{A}}) = \mathsf{E}\eta$. ■

Свойство 3. *Пусть* $\xi \geqslant 0$ *и* $\mathsf{E}\xi \geqslant 0$*, тогда* $\xi = 0$ *почти наверное.*

▲ Рассмотрим события $A = \{\xi > 0\}$ и $A_n = \{\xi > \dfrac{1}{n}\}$, следовательно, $A_n \uparrow A$. Имеем $\mathsf{P}(A_n) = \mathsf{E}I_{A_n}$, так как $n\xi > 1$ на A_n, то $\mathsf{E}I_{A_n} \leqslant \mathsf{E}(n\xi \cdot I_A) \leqslant n\mathsf{E}\xi = 0$, значит, $\mathsf{P}(A) = \lim\limits_{n\to+\infty} \mathsf{P}(A_n)$. ■

Свойство 4. *Пусть* $\mathsf{E}\xi$ *и* $\mathsf{E}\eta$ *конечны,* $\forall A \in \mathcal{F} : \mathsf{E}(\xi \cdot I_A) \leqslant \mathsf{E}(\eta \cdot I_A)$*. Тогда* $\xi \leqslant \eta$ *почти наверное.*

▲ Рассмотрим событие $B = \{\xi > \eta\}$. Из условия и построения B получаем, что $\mathsf{E}(\eta \cdot I_B) \leqslant \mathsf{E}(\xi \cdot I_B) \leqslant \mathsf{E}(\eta \cdot I_B)$, следовательно, $\mathsf{E}(\xi \cdot I_B) = \mathsf{E}(\eta \cdot I_B)$, значит $\mathsf{E}\big((\xi - \eta) \cdot I_B\big) = 0$. Так как $(\xi - \eta) \cdot I_B \geqslant 0$, 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* $\xi \perp\!\!\!\perp \eta$*, причем* $\mathsf{E}\xi$ *и* $\mathsf{E}\eta$ *конечны, тогда* $\mathsf{E}\xi\eta$ *конечно и* $\mathsf{E}\xi\eta = \mathsf{E}\xi \cdot \mathsf{E}\eta$.

▲ 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

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

Рассмотрим $\xi_n \uparrow \xi$, $\xi_n = \sum\limits_{k=0}^{n\cdot 2^n - 1} \dfrac{k}{2^n} I\left(\dfrac{k}{2^n} \leqslant \xi \leqslant \dfrac{k+1}{2^n}\right) + nI(\xi > n)$, следовательно, $\xi_n = \varphi_n(\xi)$, значит, ξ_n — $\mathcal{F}_\xi$-измеримая. Пусть $\xi, \eta \geqslant 0$. Существует последовательность $\mathcal{F}_\xi$-измеримых ($\mathcal{F}_\eta$-измеримых) простых неотрицательных простых функций $\xi_n \uparrow \xi$ ($\eta_n \uparrow \eta$). Так как $\xi \perp\!\!\!\perp \eta$, то $\xi_n = \varphi_n(\xi) \perp\!\!\!\perp \varphi_n(\eta) = \eta_n$. Следовательно, $\xi_n \cdot \eta_n \uparrow \xi \cdot \eta$, а по определению математического ожидания $\mathsf{E}\xi\eta = \lim\limits_{n\to+\infty} \mathsf{E}(\xi_n\eta_n) = \lim\limits_{n\to+\infty} \mathsf{E}\xi_n \cdot \mathsf{E}\eta_n = \mathsf{E}\xi \cdot \mathsf{E}\eta$.

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

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

Теорема (о замене переменной в интеграле Лебега). *[б/д] Пусть $(\Omega, \mathcal{F})$ 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 $(\Omega, \mathcal{F})$ и P_X — вероятностная мера на $(E, \mathcal{E})$, заданная по правилу $\mathsf{P}_X(A) = \mathsf{P}(\omega : X(\omega) = A)$ для $A \in \mathcal{E}$. Тогда для любой $\mathcal{E}$-измеримой функции $g(x), x \in E$, то есть $\forall B \in \mathcal{E} : g^{-1}(B) \in E$, верно, $\int\limits_A g(x)\mathsf{P}_X(dx) = \int\limits_{X^{-1}(A)} g\big(X(\omega)\big)\mathsf{P}(d\omega)$.*

Пусть $\xi : \Omega \to \mathbb{R}(\mathbb{R}^n)$, в таком случае вероятностная мера P_ξ однозначно восстанавливается по F_ξ, следовательно, по теореме $\mathsf{E}g(\xi) = \int g(\xi)\,d\mathsf{P} = \int g(x)\mathsf{P}_\xi(dx) = \int g(x)\,d\mathcal{F}_\xi(x)$.

Пусть ξ — абсолютно непрерывная случайная величина с плотностью $p_\xi(x)$, тогда $d\mathcal{F}_\xi(x) = p_\xi(x)$, следовательно $\mathsf{E}g(x) = \int\limits_{\mathbb{R}} g(x)p_\xi(x)\,dx$.

Прямое произведение вероятностных пространств и формула свертки

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

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

Теорема (Фубини). *[б/д] Пусть $(\Omega, \mathcal{F}, \mathsf{P})$ — прямое произведение вероятностных пространств $(\Omega_1, \mathcal{F}_1, \mathsf{P}_1)$ и $(\Omega_2, \mathcal{F}_2, \mathsf{P}_2)$. Пусть $\xi : \Omega \to \mathbb{R}$ такая, что $\int\limits_\Omega \big|\xi(\omega_1, \omega_2)\big|\,d\mathsf{P} < +\infty$. Тогда интегралы $\int\limits_{\Omega_1} \xi(\omega_1, \omega_2)\mathsf{P}_1(d\omega_1)$ и $\int\limits_{\Omega_2} \xi(\omega_1, \omega_2)\mathsf{P}_2(d\omega_2)$ определены почти наверное относительно P_2 и P_1 соответственно, являются измеримыми случайными величинами относительно $\mathcal{F}_2$ и $\mathcal{F}_1$ соответственно* *и* $\int\limits_{\Omega} \xi(\omega_1, \omega_2)\, d\mathsf{P} = \int\limits_{\Omega_2} \int\limits_{\Omega_1} \xi(\omega_1, \omega_2) \mathsf{P}_1(d\omega_1) \mathsf{P}_2(d\omega_2) + \int\limits_{\Omega_1} \int\limits_{\Omega_2} \xi(\omega_1, \omega_2) \mathsf{P}_2(d\omega_2) \mathsf{P}_1(d\omega_1)$. *Из всего этого следует, что двойной интеграл равен повторному.*

Утверждение. *Пусть* $\xi \perp\!\!\!\perp \eta$ *— случайные величины, тогда* $(\mathbb{R}^2, \mathscr{B}(\mathbb{R}^2), \mathsf{P}_{(\xi,\eta)}) = (\mathbb{R}, \mathscr{B}(\mathbb{R}), \mathsf{P}_\xi) \otimes (\mathbb{R}, \mathscr{B}(\mathbb{R}), \mathsf{P}_\eta)$.

▲ Достаточно проверить свойство прямого произведения:

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

5 Лекция от 17.03.2018

Дисперсия и ковариация

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

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

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

Свойства ковариации и дисперсии

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

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

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

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

▲
$$\begin{aligned} \mathsf{D}(c\xi) &= \mathsf{E}c^2\xi^2 - (\mathsf{E}c\xi)^2 = c^2 \mathsf{E}\xi^2 - c^2 (\mathsf{E}\xi)^2 = c^2 \mathsf{D}\xi; \\ \mathsf{D}(c + \xi) &= \mathsf{E}\big(c + \xi - \mathsf{E}(c + \xi)\big)^2 = \mathsf{E}\,(c + \xi - c - \mathsf{E}\xi)^2 = \mathsf{D}\xi; \\ \mathsf{D}c &= \mathsf{E}(c - \mathsf{E}c)^2 = \mathsf{E}(c - c)^2 = 0. \end{aligned}$$
■

Свойство 4 (Неравенство Коши-Буняковского)**.** $|\mathsf{E}\xi\eta|^2 \leqslant \mathsf{E}\xi^2 \cdot \mathsf{E}\eta^2$

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

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

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

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

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

Лемма. *Пусть* $\xi_1, \ldots, \xi_n$ *— попарно некоррелированные случайные величины (например, независимые в совокупности),* $\mathsf{D}x_1, \ldots, \mathsf{D}\xi_n < +\infty$*, тогда* $\mathsf{D}(\xi_1, \ldots, \xi_n) = \mathsf{D}\xi_1 + \ldots + \mathsf{D}\xi_n$.

▲

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

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

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

■

Многомерный случай

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

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

$$\operatorname{Var}\vec{\xi} = \begin{pmatrix} \operatorname{cov}(\xi_1, \eta_1) & \cdots & \operatorname{cov}(\xi_1, \eta_n) \\ \vdots & \ddots & \vdots \\ \operatorname{cov}(\xi_n, \eta_1) & \cdots & \operatorname{cov}(\xi_n, \eta_n) \end{pmatrix} = \left\| \operatorname{cov}(\xi_i, \eta_j) \right\|_{i,j=1}^{n}.$$

Лемма. *Матрица ковариаций случайного вектора — симметрическая и неотрицательно определенная*[1].

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

$$\vec{x}^T \operatorname{Var}\vec{\xi}\vec{x} = (\vec{x}, \operatorname{Var}\vec{\xi}\vec{x}) = \sum_{i,j=1}^n \operatorname{cov}(\xi_i, \xi_j) x_i x_j = \sum_{i,j=1}^n \operatorname{cov}(x_i\xi_i, x_j\xi_j) =$$

$$= \operatorname{cov}\left(\sum_{i=1}^n x_i\xi_i, \sum_{j=1}^n x_j\xi_j\right) = \operatorname{cov}\left(\sum_{i=1}^n x_i\xi_i, \sum_{i=1}^n x_i\xi_i\right) = \mathsf{D}\left(\sum_{i=1}^n x_i\xi_i\right) \geqslant 0.$$

■

Неравенства

Лемма (Неравенство Маркова)**.** *Пусть* $\xi \geqslant 0$ *— случайная величина,* $\mathsf{E}\xi < +\infty$ *(существует). Тогда* $\forall \varepsilon > 0 : \mathsf{P}(\xi \geqslant \varepsilon) \leqslant \dfrac{\mathsf{E}\xi}{\varepsilon}$.

▲ $\mathsf{P}(\xi \geqslant \varepsilon) = \mathsf{E}I(\xi \geqslant \varepsilon)$. На множестве $\xi \geqslant \varepsilon$ случайная величина $\dfrac{\xi}{\varepsilon} \geqslant 1$, следовательно $\mathsf{E}I(\xi \geqslant \varepsilon) \leqslant \mathsf{E}\left(\dfrac{\xi}{\varepsilon} \cdot I(\xi \geqslant \varepsilon)\right) \leqslant \dfrac{1}{\varepsilon} \cdot \mathsf{E}\xi$. ■

Лемма (Неравенство Чебышёва)**.** *Пусть* ξ *— случайная величина такая, что* $\mathsf{D}\xi < +\infty$*, тогда* $\forall \varepsilon > 0 : \mathsf{P}\big(|\xi - \mathsf{E}\xi| \geqslant \varepsilon\big) \leqslant \dfrac{\mathsf{D}\xi}{\varepsilon^2}$.

▲ $\mathsf{P}\big(|\xi - \mathsf{E}\xi| \geqslant \varepsilon\big) = \mathsf{P}\big(|\xi - \mathsf{E}\xi|^2 \geqslant \varepsilon^2\big)$. Из неравенства Маркова имеем, что $\mathsf{P}\big(|\xi - \mathsf{E}\xi|^2 \geqslant \varepsilon^2\big) \leqslant \dfrac{\mathsf{E}(\xi - \mathsf{E}\xi)^2}{\varepsilon^2} = \dfrac{\mathsf{D}\xi}{\varepsilon^2}$. ■

Лемма (Неравенство Йенсена)**.** *Пусть* $g(x)$ *— борелевская выпуклая вниз (вверх) функция и* $\mathsf{E}\xi < +\infty$*. Тогда* $\mathsf{E}g(\xi) \geqslant g(\mathsf{E}\xi)$ *(* $\mathsf{E}g(\xi) \leqslant g(\mathsf{E}\xi)$ *).*

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

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

Теорема (Закон больших чисел в форме Чебышёва)**.** *Пусть* $\{\xi_1, \ldots, \xi_n\}_{i=1}^{+\infty}$ *— последовательность попарно некоррелированных случайных величин таких, что* $\forall n \in \mathbb{N} : \mathsf{D}\xi_n \leqslant C$*. Обозначим* $S_n = \sum\limits_{i=1}^n \xi_i$*, тогда* $\dfrac{S_n - \mathsf{E}S_n}{n} \xrightarrow{\mathsf{P}} 0$ *при* $n \to +\infty$*.*

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

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

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

Условные математические ожидания (УМО)

Пусть $(\Omega, \mathcal{F}, \mathsf{P})$ — вероятностное пространство; $\xi : \Omega \to \mathbb{R}$ — случайная величина; $\mathcal{F}_\xi = \{\xi^{-1}(B), B \in \mathscr{B}(\mathbb{R})\}$ — 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 случайная величина на $(\Omega, \mathcal{F}, \mathsf{P})$, $\mathcal{G}$ — подσ-алгебра $\mathcal{F}$. Условным математическим ожиданием случайной величины ξ относительно $\mathcal{G}$ называется случайная величина $\mathsf{E}(\xi|\mathcal{G}$, обладающая следующими свойствами:

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

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

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

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

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

6 Лекция от 24.03.2018

Лемма (о существовании УМО). *Пусть* ξ *— случайная величина с* $\mathsf{E}|\xi| < +\infty$. *Тгда* $\forall \mathcal{G} \subset \mathcal{F}$ *(подσ-алгебра)* : $\mathsf{E}(\xi|\mathcal{G})$ *существует и единственно почти наверное.*

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

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

▲ Пусть ■

7 Лекция от 31.03.2018

Условные распределения

Определение. Величиной $\mathsf{E}(\xi|\eta = y)$ называется такая борелевска функция $\varphi(y)$, что $\forall B \in \mathscr{B}(\mathbb{R}) : \mathsf{E}(\xi \cdot I(\eta \in B) = \int\limits_B \varphi(y)\mathsf{P}_\eta(dy)$.

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

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

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

▲ Пусть $B \in \mathscr{B}(\mathbb{R})$, тогда $\mathsf{E}\big(\mathsf{E}(\xi|\eta) \cdot I(\eta \in B)\big) = \mathsf{E}\big(\xi \cdot I(\eta \in B)\big) = \int\limits_B \varphi(y)\mathsf{P}_\eta(dy)$. По теореме о замене переменных в интеграле Лебега это можно переписать, как $\int\limits_{\{\eta \in B\}} \varphi(\eta)\, d\mathsf{P} = \mathsf{E}\big(\varphi(\eta) \cdot I(\eta \in B)\big)$, что равносильно условию $\mathsf{E}(\xi|\eta) = \varphi(\eta)$ почти наверное по Свойству. Обратно аналогично, по тем же равенствам. ■

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

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

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

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

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 $\mathsf{E}\big|g(x)$ существует, выполнено $\mathsf{E}\big(g(\xi)|\eta = y\big) = \int\limits_{\mathbb{R}} g(x) f_{\xi|\eta}(x|y)\,dx$ относительно P_η почти наверное.*

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

$$\int\limits_{\mathbb{R}} g(x) \cdot f_{\xi|\eta}(x|y)\,dx = \int\limits_{\mathbb{R}} I_A(x) \cdot f_{\xi|\eta}(x|y)\,dx = \int\limits_{A} \cdot f_{\xi|\eta}(x|y)\,dx =$$
$$= \mathsf{P}(\xi \in A|\eta = y) = \mathsf{E}\big(I(\xi \in A)|\eta \in y\big) = \mathsf{E}\big(g(\xi)|\eta = y\big).$$

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 $g(x)$. ($\mathsf{E}(\xi_n|\eta) \xrightarrow{\text{п.н.}} \mathsf{E}(\xi|\eta)$, где $\xi_n \xrightarrow{\text{п.н.}} \to \xi$, 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*

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

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

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

$$\mathsf{P}(\xi \in B, \eta \in A) = \int\limits_{B \rtimes A} f_{(\xi,\eta)}(x, y)\,dx\,dy = \int\limits_{A} \left(\int\limits_{B} \frac{f_{(\xi,\eta)}(x, y)}{f_\eta(y)}\,dx \right) f_\eta(y)\,dy,$$

с другой стороны

$$\mathsf{P}(\xi \in B, \eta \in A) = \mathsf{E}\big(I(\xi \in B, \eta \in A)\big) = \int\limits_{\{\eta \in A\}} I(\xi \in B)\,d\mathsf{P}.$$

Далее по интегральному свойству получаем, что

$$\mathsf{P}(\xi \in B, \eta \in A) = \int\limits_{\{\eta \in A\}} \mathsf{E}\big(I(\xi \in B)|\eta\big)\, d\mathsf{P},$$

заменяя переменные, окончательно имеем следующее:

$$\mathsf{P}(\xi \in B, \eta \in A) = \int\limits_A \mathsf{E}\big(I(\xi \in B|\eta = y)\mathsf{P}_\eta(dy) =$$

$$= \int\limits_A \mathsf{P}(\xi \in B|\eta = y)\mathsf{P}_\eta\, dy = \int\limits_A \mathsf{P}(\xi \in B|\eta = y) f_\eta(y)\, dy.$$

■

Алгоритм подсчета УМО

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

Виды сходимости случайных величин

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

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

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

Лемма (критерий сходимости почти наверное). $\xi_n \xrightarrow{\text{п.н.}} \xi$ *тогда и только тогда, когда* $\forall \varepsilon > 0 : \mathsf{P}\big(\omega : \sup\limits_{k\geqslant n} |\xi_k(\omega) - \xi(\omega)| \geqslant \varepsilon\big) \xrightarrow[n\to+\infty]{} 0$.

▲ Пусть $A_k^\varepsilon = \{\omega : |\xi_k - \xi| \geqslant \varepsilon\}$, $A^\varepsilon = \bigcap\limits_{n=1}^{+\infty} \bigcup\limits_{k\geqslant n} A_k^\varepsilon = \{\omega : \forall n\ \exists f \geqslant n : |\xi_n - \xi| \geqslant \varepsilon\}$. Тогда $\{\omega : \xi_n(\omega) \not\to \xi(\omega)\} = \bigcup\limits_{m=1}^{+\infty} A^{\frac{1}{m}} = \{\omega : \exists m\ \forall n\ \exists k \geqslant n : |\xi_k(\omega) - \xi)\omega)| > \varepsilon\}$. Следовательно,

$$\mathsf{P}\big(\omega : \xi_n(\omega) \not\to \xi(\omega)\big) = 0 \Leftrightarrow \mathsf{P}\left(\bigcup_{m=1}^{+\infty} A^{\frac{1}{m}}\right) = 0 \Leftrightarrow$$
$$\Leftrightarrow \forall m \in \mathbb{N} : \mathsf{P}\left(A^{\frac{1}{m}}\right) = 0 \Leftrightarrow \forall \varepsilon > 0 : \mathsf{P}\left(A^\varepsilon\right) = 0,$$

так как всегда существует m, что $\frac{1}{m} \geqslant \varepsilon \geqslant \frac{1}{m+1}$, то есть $A^{\frac{1}{m+1}} \supseteq A^\varepsilon \supseteq A^{\frac{1}{m}}$. Но $\bigcup\limits_{k\geqslant n} A_k^\varepsilon \downarrow A^\varepsilon$, следовательно,

$$0 = \mathsf{P}\left(A^\varepsilon\right) = \lim_{n\to+\infty} \mathsf{P}\left(\bigcup_{k\geqslant n} A_k^\varepsilon\right) \Leftrightarrow \forall \varepsilon > 0 : \mathsf{P}\left(\bigcup_{k\geqslant n} A_k^\varepsilon\right) \xrightarrow[n\to+\infty]{} 0 \Leftrightarrow$$
$$\Leftrightarrow \forall \varepsilon > 0 : \mathsf{P}\Big(\omega : \sup_{k\geqslant n} |\xi_k(\omega) - \xi(\omega)| \geqslant \varepsilon\Big) \xrightarrow[n\to+\infty]{} 0.$$

■

Теорема (взаимоотношения различных видов сходимости).

$$\begin{matrix} \textit{п.н.} & \searrow & & & \\ & & \mathsf{P} & \longrightarrow & d \\ L_p & \nearrow & & & \end{matrix}$$

▲ (п.н. $\Rightarrow$ P) $\quad \xi_n \xrightarrow{\text{п.н.}} \xi \Leftrightarrow \forall \varepsilon > 0 : \mathsf{P}\big(\omega : \sup\limits_{k\geqslant n} |\xi_k(\omega) - \xi(\omega)| > \varepsilon\big) \to 0$, но

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

следовательно, $\mathsf{P}\big(|\xi_n - \xi| \geqslant \varepsilon\big) \to 0$.

($L_p \Rightarrow \mathsf{P}$) $\quad \mathsf{P}\big(|\xi_n - \xi| \geqslant \varepsilon\big) = \mathsf{P}\big(\omega : |\xi_n(\omega) - \xi(\omega)|^p > \varepsilon^p$, а по неравенству Маркова это меньше или равно $\dfrac{\mathsf{E}|\xi_n(\omega) - \xi(\omega)|^p}{\varepsilon p} \xrightarrow[n\to+\infty]{} 0$.

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

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

Рассмотрим разбиение Ω:

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

Оценим

$$\big|\mathsf{E}f(\xi_n) - \mathsf{E}f(\xi)\big| \leqslant \mathsf{E}\big|f(\xi_n) - f(\xi)\big| = \mathsf{E}\big[|f(\xi) - f(\xi_n)| \cdot (I_{A_1} + I_{A_2} + I_{A_3})\big] \boxed{\leqslant}.$$

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

$0, \text{ т.к. } \xi_n \xrightarrow{\mathsf{P}} \xi$

Значит, $\boxed{\leqslant} \dfrac{\varepsilon}{2} + 2C \cdot \mathsf{P}(A_2) + 2C \cdot \mathsf{P}(A_3) \leqslant \dfrac{\varepsilon}{2} + 2C \cdot \mathsf{P}\big(|\xi| > N\big) + 2C \cdot \mathsf{P}\big(|\xi_n - \xi| > \delta\big) \leqslant C_1\varepsilon.$
Следовательно, $\mathsf{E}f(\xi_n) \to \mathsf{E}f(\xi)$, то есть $\xi_n \xrightarrow{d} \xi$. ■

8 Лекция от 07.04.2018