

Теория вероятностей «Числовые характеристики случайных величин»

Математическое ожидание

Про так, как вводится мат.ожидание в 3 шага: сначала для простых случайных величин, потом для неотрицательных, потом для всех, можно посмотреть ниже в приложении.

Определение 1. Математическим ожиданием случайной величины ξ на вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$ называется число

$$\mathbb{E}\xi = \int_{\mathbb{R}} x d\mathbb{P}_{\xi}(x) = \int_{\mathbb{R}} x dF_{\xi}(x).$$

если интеграл, стоящий в правой части, существует.

По-другому математическое ожидание можно записать в следующем виде:

$$\mathbb{E}\xi \stackrel{\text{def}}{=} \int_{\Omega} \xi(\omega) d\mathbb{P}(\omega),$$

, Можно сказать, что это обобщение понятия среднего для случайной величины, ведь $\int d\mathbb{P}(\omega) = 1$. Кроме того справедливо равенство, которое более удобно для практики: ведь $\int d\mathbb{P}(\omega) = 1$.

Следствие 1. Если ξ — дискретная случайная величина, то

$$\mathbb{E}\xi = \sum_k x_k \mathbb{P}\{\xi = x_k\},$$

где x_1, x_2, \dots — набор значений случайной величины ξ (при условии, что ряд абсолютно сходится; в противном случае говорят, что случайная величина ξ **не имеет математического ожидания**).

Следствие 2. Если ξ — абсолютно непрерывная случайная величина, имеющая плотность распределения $f(x)$, то

$$\mathbb{E}\xi = \int_{\mathbb{R}} x f(x) dx,$$

при условии, что интеграл сходится абсолютно (иначе случайная величина ξ **не имеет математического ожидания**).

Следствие 3. Если $g(x)$ — борелевская функция на прямой, то $\eta = g(\xi)$ — тоже случайная величина и

$$\mathbb{E}g(\xi) = \int_{\Omega} g(\xi(\omega)) d\mathbb{P}(\omega) = \int_{\mathbb{R}} g(x) dF_{\xi}(x) = \int_{\mathbb{R}} x dF_{g(\xi)}(x)$$

Если ξ — дискретная случайная величина, то

$$\mathbb{E}g(\xi) = \sum_k g(x_k) \mathbb{P}\{g(\xi) = g(x_k)\},$$

Если ξ — абсолютно непрерывная случайная величина, имеющая плотность распределения $f(x)$, то

$$\mathbb{E}g(\xi) = \int_{\mathbb{R}} g(x) f(x) dx,$$

Из определения следует, что $\mathbb{E}\xi < \infty \Leftrightarrow \mathbb{E}|\xi| < \infty$. Можно показать, что $\mathbb{E}\xi$ не ограничено или не существует, если, например, $F(x) < 1 - \frac{1}{x}$ для всех достаточно больших x .

Основные свойства математического ожидания следуют из свойств интеграла Лебега.

1. Если a и b постоянные, то $\mathbb{E}[a + b\xi] = a + b\mathbb{E}\xi$.
2. $\mathbb{E}[\xi + \eta] = \mathbb{E}\xi + \mathbb{E}\eta$, если существуют какие-нибудь два участвующих в равенстве математических ожидания. Отметим, что случайные величины совсем не обязательно должны быть независимыми. Данное свойство не совсем очевидно, если задавать математическое ожидание по формулам, которые мы получили в Следствиях 1 и 2, но оно легко получается из определения через интеграл Лебега.
3. Если $a \leq \xi \leq b$, то $a \leq \mathbb{E}\xi \leq b$.
4. $|\mathbb{E}\xi| \leq \mathbb{E}|\xi|$.
5. Если $\xi \geq 0$ и $\mathbb{E}\xi = 0$, то $\xi = 0$ с вероятностью 1. Данное свойство легко получить, если представить интеграл Лебега по Ω в виде суммы интеграла Лебега по A и B , где $A = \{\omega \in \Omega \mid \xi(\omega) = 0\}$, $B = \Omega \setminus A$.
6. $\mathbb{P}\{A\} = \mathbb{E}\mathbb{I}_A$.
7. Если ξ, η — независимые случайные величины, то $\mathbb{E}[\xi\eta] = \mathbb{E}\xi\mathbb{E}\eta$. Данное свойство доказывается через произведение мер и **теорему Фубини**.

Пример 1. Пусть ξ — число выпавших очков на игральном кубике. Тогда

$$\mathbb{E}\xi = \sum_{k=1}^6 k \mathbb{P}\{\xi = k\} = \frac{1}{6} \sum_{k=1}^6 k = \frac{7}{2} = 3,5.$$

Пример 2. Пусть $\xi \sim \text{Be}(p)$. Найдите $\mathbb{E}\xi$.

Решение. $\mathbb{E}\xi = 1 \cdot \mathbb{P}\{\xi = 1\} + 0 \cdot \mathbb{P}\{\xi = 0\} = p$

□

Пример 3. Рассмотрим случайную величину ξ с распределением $\mathbb{P}\{\xi = 2^n\} = 2^{-n}, n \geq 1$. Тогда

$$\mathbb{E}\xi = \sum_n 2^n \cdot 2^{-n} = +\infty.$$

У случайной величины η с распределением $\mathbb{P}\{\eta = 2^n\} = \mathbb{P}\{\eta = -2^n\} = 2^{-n-1}$ нет мат. ожидания, так как $\mathbb{E}\xi_+ = \mathbb{E}\xi_- = +\infty$ (про ξ_- и ξ_+ смотри в приложении).

Пример 4. Примером абсолютно непрерывной случайной величины, не имеющей мат. ожидания, является случайная величина $\xi \sim (0, 1)$. Действительно, её плотность $f(x) = \frac{1}{\pi} \cdot \frac{1}{x^2+1}$, а интеграл

$$\int_{\mathbb{R}} \frac{x}{x^2+1} dx$$

не сходится абсолютно.

Упражнение 1. Используя линейность мат. ожидания, найдите мат. ожидания случайной величины $\xi \sim \text{Binom}(n, p)$.

Решение. Как мы знаем, если рассмотреть независимые случайные величины $\eta_1, \eta_2, \dots \sim \text{Be}(p)$, то сумма первых n из них имеет биномиальное распределение $\text{Binom}(n, p)$. Так как мат. ожидание зависит только от функции распределения, то отсюда

$$\mathbb{E}\xi = \mathbb{E}\left[\sum_{k=1}^n \eta_k\right] = \sum_{k=1}^n \mathbb{E}\eta_k = np.$$

□

Свойство линейности математического ожидания бывает очень полезно для поиска мат. ожидания количества чего-нибудь. В таких задачах часто работает следующий приём: представим исходную случайную величину в виде суммы индикаторов каких-то событий (возможно, зависимых) и воспользуемся линейностью мат. ожидания.

Следствие 4. Если у величин совпадают мат. ожидания, то мы ничего не можем сказать, из одного они распределения или нет. Например, если взять биномиальную случайную величину и бернуллиевскую, мы можем так подобрать параметры, что у них будут совпадать.

Упражнение 2. (Задача №79) Имеется n пронумерованных конвертов и n пронумерованных писем. Письма случайным образом раскладываются по конвертам (все $n!$ способов равновероятны). Найдите мат. ожидание числа совпадений номеров письма и конверта (письмо лежит в конверте с тем же номером).

Решение. Пусть ξ — число совпадений номеров письма и конверта. Рассмотрим следующие индикаторные случайные величины:

$$\xi_i = \begin{cases} 1, & i\text{-е письмо попало в } i\text{-й конверт,} \\ 0, & \text{иначе.} \end{cases}$$

Тогда $\xi = \sum_{i=1}^n \xi_i$. Кроме того, $\xi_i \sim \text{Be}(p_i)$, где $p_i = \mathbb{P}\{i\text{-е письмо попало в } i\text{-й конверт}\}$. Для всех $i = 1, 2, \dots, n \hookrightarrow p_i = \frac{(n-1)!}{n!} = \frac{1}{n}$, а значит, $\mathbb{E}\xi_i = p_i = \frac{1}{n}$. Отсюда следует, что

$$\mathbb{E}\xi = \sum_{i=1}^n \mathbb{E}\xi_i = 1.$$

□

Пример 5. Пусть $\xi \sim \text{Exp}(1)$ (т.е. $F_\xi(x) = 1 - e^{-x}$) и $\eta = \frac{e^\xi \sin \xi}{\xi}$. Тогда существует несобственный интеграл Римана-Стильтьеса

$$\int_{-\infty}^{+\infty} \frac{e^x \sin x}{x} dF_\xi(x) = \int_0^{+\infty} \frac{e^x \sin x}{x} e^{-x} dx = \int_0^{+\infty} \frac{\sin x}{x} dx = \frac{\pi}{2},$$

но $\int_0^{+\infty} \left| \frac{\sin x}{x} \right| dx = \infty$, а значит, η не имеет математического ожидания.

Пример 6. Пусть $\xi \sim \mathcal{U}[0, 1]$ и $\eta = R(\xi)$, где

$$R(x) = \begin{cases} x, & x \in \mathbb{R} \setminus \mathbb{Q}, \\ 0, & x \in \mathbb{Q}. \end{cases}$$

Тогда $\mathbb{E}\eta = \mathbb{E}\xi = \frac{1}{2}$, т. к. $\xi = \eta$ почти наверное, что означает, что $\mathbb{P}\{\xi = \eta\} = 1$. Однако интеграла Римана (Римана-Стильтьеса) $\int_0^1 R(x) dx$ не существует.

Упражнение 3. Найти математическое ожидание $\xi \sim \text{Poisson}(\lambda)$

Решение. $\mathbb{E}\xi = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} k = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda \cdot \lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{r=0}^{\infty} \frac{\lambda^r}{r!} = \lambda$.

□

Упражнение 4. Найти математическое ожидание $\xi \sim \mathcal{N}(0, 1)$

Решение. Запишем через плотностью случайной величины: $\mathbb{E}\xi = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x e^{-\frac{x^2}{2}} dx = 0$, так как это нечетная функция.

□

Дисперсия и другие числовые характеристики случайных величин

Определение 2. Пусть имеется случайная величина ξ с математическим ожиданием $\mathbb{E}\xi$. **Центрированной** случайной величиной, соответствующей величине ξ , называется случайная величина $\overset{\circ}{\xi}$:

$$\overset{\circ}{\xi} = \xi - \mathbb{E}\xi$$

Определение 3. k -м начальным моментом случайной величины ξ , где $k \in \mathbb{N}$, называется величина :

$$\nu_k = \mathbb{E}(\xi^k)$$

k -м центральным моментом случайной величины ξ , где $k \in \mathbb{N}$, называется величина :

$$\mu_k = \mathbb{E}(\overset{\circ}{\xi})^k = \mathbb{E}((\xi - \mathbb{E}\xi)^k)$$

Для моментов справедливы формулы(они все берутся из следствия 3):

1. Если ξ — дискретная случайная величина, то

$$\nu_k = \sum_i x_i^k \mathbb{P}\{\xi = x_i\}$$

$$\mu_k = \sum_i (x_i - \mathbb{E}\xi)^k \mathbb{P}\{\xi = x_i\}$$

2. Если ξ — абсолютно непрерывная случайная величина, имеющая плотность распределения $f(x)$, то

$$\nu_k = \int_{\mathbb{R}} x^k f(x) dx$$

$$\mu_k = \int_{\mathbb{R}} (x - \mathbb{E}\xi)^k f(x) dx$$

3. Если $g(x)$ — борелевская функция на прямой, то $\eta = g(\xi)$ — тоже случайная величина и если ξ — дискретная случайная величина, то

$$\nu_k = \sum_i g^k(x_i) \mathbb{P}\{g(\xi) = g(x_i)\},$$

если ξ — абсолютно непрерывная случайная величина, имеющая плотность распределения $f(x)$, то

$$\nu_k = \int_{\mathbb{R}} g^k(x) f(x) dx,$$

Определение 4. Дисперсией случайной величины ξ называется второй центральный момент:

$$\mathbb{D}\xi \stackrel{\text{def}}{=} \mathbb{E}[(\xi - \mathbb{E}\xi)^2].$$

Перечислим некоторые свойства дисперсии.

1. $\mathbb{D}\xi \geq 0$. $\mathbb{D}\xi = 0$ тогда и только тогда, когда существует такая константа c , что $\mathbb{P}\{\xi = c\} = 1$.
2. $\mathbb{D}\xi = \mathbb{E}[\xi^2 - 2\xi\mathbb{E}\xi + (\mathbb{E}\xi)^2] = \mathbb{E}[\xi^2] - (\mathbb{E}\xi)^2$.

3. $\mathbb{D}[a + b\xi] = b^2\mathbb{D}[\xi]$, где $a, b = \text{const}$.

4. Если ξ и η — независимые случайные величины, то $\mathbb{D}[\xi + \eta] = \mathbb{D}\xi + \mathbb{D}\eta$. Действительно,

$$\begin{aligned}\mathbb{D}[\xi + \eta] &= \mathbb{E}[(\xi + \eta - \mathbb{E}\xi - \mathbb{E}\eta)^2] \\ &= \mathbb{E}[(\xi - \mathbb{E}\xi)^2] + \mathbb{E}[(\eta - \mathbb{E}\eta)^2] + 2\mathbb{E}[(\xi - \mathbb{E}\xi)(\eta - \mathbb{E}\eta)] \\ &= \mathbb{D}\xi + \mathbb{D}\eta + 2\mathbb{E}[\xi - \mathbb{E}\xi]\mathbb{E}[\eta - \mathbb{E}\eta] \\ &= \mathbb{D}\xi + \mathbb{D}\eta.\end{aligned}$$

Когда ищем дисперсию суммы n случайных величин достаточно их попарной независимости для того, чтобы дисперсия суммы была равна сумме дисперсий.

5. $\mathbb{D}\xi$ минимизирует значение $\mathbb{E}[(\xi - a)^2]$, $a \in \mathbb{R}$. Действительно, расписывая математическое ожидание квадрата, получим

$$\mathbb{E}[(\xi - a)^2] = \mathbb{E}[\xi^2] - 2a\mathbb{E}\xi + a^2.$$

Относительно a записанное выражение является квадратичной функцией, а значит, достигает своего минимума при $a = \mathbb{E}\xi$, что и требовалось доказать.

Упражнение 5. Найти дисперсию случайной величины ξ , если

(a) $\xi \sim \text{Binom}(n, p)$;

(b) $\xi \sim \mathcal{U}[a, b]$.

Решение. (a) Как мы знаем, сумма n независимых бернуллиевских случайных величин $\xi_1, \dots, \xi_n \sim \text{Be}(p)$ имеет биномиальное распределение $\text{Binom}(n, p)$. Дисперсия бернуллиевской случайной величины: $\mathbb{D}\xi_1 = \mathbb{E}[\xi_1^2] - (\mathbb{E}\xi_1)^2 = p - p^2 = p(1 - p)$. Так как ξ_1, \dots, ξ_n независимы, то

$$\mathbb{D}\xi = \sum_{k=1}^n \mathbb{D}\xi_k = np(1 - p).$$

(b) Математическое ожидание случайной величины, имеющей равномерное распределение на отрезке $[a, b]$: $\mathbb{E}\xi = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{a+b}{2}$. Второй момент: $\mathbb{E}[\xi^2] = \frac{1}{b-a} \int_a^b x^2 dx = \frac{a^2+ab+b^2}{3}$. В итоге получаем, что

$$\mathbb{D}\xi = \frac{a^2 + ab + b^2}{3} - \frac{a^2 + 2ab + b^2}{4} = \frac{(b-a)^2}{12}.$$

□

Упражнение 6. Найти дисперсию случайной величины $\xi \sim \mathcal{N}(0, 1)$

Решение. Запишем по определению: $\mathbb{D}\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2 = \mathbb{E}\xi^2$. Распишем этот интеграл поподробнее: $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}} dx = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x d\left(e^{-\frac{x^2}{2}}\right) = -\frac{1}{\sqrt{2\pi}} x e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 0 + 1 = 1$.

□

Следствие 5. Как известно, если $X \sim \mathcal{N}(\mu, \sigma^2)$, то $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$. Можем выразить $X = \mu + \sigma\xi$, где $\xi \sim \mathcal{N}(0, 1)$. Тогда $\mathbb{E}X = \mu, \mathbb{D}X = \sigma^2$.

Упражнение 7. 100 паровозов выехали из города по однопутной железной дороге, каждый с постоянной скоростью. Когда движение установилось, образовалось несколько караванов (групп, движущихся рядом, со скоростью лидера каравана). Найти мат.ожидание и дисперсию их числа. Функция распределения скорости непрерывна и строго возрастает на положительной полуоси, а скорости различных паровозов независимы.

Решение. Для начала используем важный факт в этой задаче, а именно то, что функция распределения скорости непрерывна и строго возрастает на положительной полуоси, а скорости различных паровозов независимы. Из этого следует, что с нулевой вероятностью скорости двух паровозов совпадут (аналогия: два раза попасть в одну и ту же точку отрезка).

Каждому паровозу с номером i поставим в соответствие случайную величину ξ_i , равную 1, если он является лидером каравана (возможно, состоящего только из самого этого паровоза), и равную 0 в противном случае.

$$\mathbb{P}\{\xi_i = 1\} = \mathbb{P}\{\text{паровоз с } i \text{ номером медленнее всех, кто идет раньше его}\} = \frac{1}{i}$$

Наша итоговая случайная величина $\eta = \sum_{i=1}^{100} \xi_i$.

Посчитаем мат.ожидание, воспользовавшись линейностью:

$$\mathbb{E}\eta = \sum_{i=1}^{100} \mathbb{E}\xi_i = \sum_{i=1}^{100} \frac{1}{i}$$

Легко заметить, что ξ_i попарно независимы: является ли i -ое число самым большим среди i первых чисел никак не зависит от того является ли j -ое самым большим среди первых j чисел, в силу независимости скоростей.

Тогда дисперсия:

$$\mathbb{D}\eta = \sum_{i=1}^{100} \mathbb{D}\xi_i = \sum_{i=1}^{100} \frac{1}{i} - \frac{1}{i^2}$$

□

Пространство L_2

Случайные величины с конечным вторым моментом образуют линейное функциональное пространство $L_2(\Omega, \mathcal{F}, \mathbb{P})$. Кроме того, отождествим случайные величины, которые отличаются на множестве вероятностной меры ноль. Это делается для того, чтобы ввести скалярное произведение:

$$\langle \xi, \eta \rangle \stackrel{\text{def}}{=} \mathbb{E}[\xi\eta].$$

Действительно, равенство $\langle \xi, \xi \rangle = 0$ означает только, что $\xi = 0$ почти наверное, но не для любого $\omega \in \Omega$. Но мы их отождествили, поэтому все аксиомы скалярного произведения выполнены. Таким образом, мы получили евклидово пространство. Норма вводится естественным образом: $\|\xi\| \stackrel{\text{def}}{=} \sqrt{\langle \xi, \xi \rangle}$. Отсюда следует, что

$$\mathbb{D}\xi = \|\xi - \mathbb{E}\xi\|^2.$$

Для скалярного произведения выполняется **неравенство Коши-Буняковского-Шварца**:

$$|\mathbb{E}[\xi\eta]| \stackrel{\text{def}}{=} |\langle \xi, \eta \rangle| \leq \|\xi\| \cdot \|\eta\| \stackrel{\text{def}}{=} \sqrt{\mathbb{E}[\xi^2]\mathbb{E}[\eta^2]}.$$

Определение 5. Ковариацией случайных величин ξ, η называется число

$$\text{cov}(\xi, \eta) \stackrel{\text{def}}{=} \mathbb{E}[(\xi - \mathbb{E}\xi)(\eta - \mathbb{E}\eta)] = \mathbb{E}[\xi\eta] - \mathbb{E}\xi\mathbb{E}\eta.$$

Из неравенства КБШ следует, что для ковариации выполняется неравенство:

$$|\text{cov}(\xi, \eta)| \leq \sqrt{\mathbb{D}\xi\mathbb{D}\eta}.$$

Нетрудно видеть, что если ξ, η — независимые, то $\text{cov}(\xi, \eta) = 0$. Обратное неверно: если ковариация величин равна 0, из этого не следует их независимость. Тем не менее, ковариация часто используется, как некоторая мера зависимости величин, т.к. ей удобно пользоваться.

Пример 7. Привести пример двух случайных величин таких, что и ковариация равна 0, но они не являются независимыми.

Решение. Пусть случайная величина ξ принимает значения 0, $\pi/2$, π , каждое с вероятностью $1/3$. Тогда $\cos \xi$ будет принимать значения -1, 0 и 1, каждое с вероятностью $1/3$, а $\mathbb{P}\{\sin \xi = 1\} = 1/3$, $\mathbb{P}\{\sin \xi = 0\} = 2/3$, $\mathbb{P}\{\sin \xi = -1\} = 0$. Тогда $\text{cov}(\sin \xi, \cos \xi) = 0$, но $\mathbb{P}(\sin \xi = 1, \cos \xi = 1) \neq \mathbb{P}(\cos \xi = 1)\mathbb{P}(\sin \xi = 1)$. \square

Определение 6. Коэффициентом корреляции случайных величин ξ, η называется число

$$r_{\xi\eta} \stackrel{\text{def}}{=} \frac{\text{cov}(\xi, \eta)}{\sqrt{\mathbb{D}\xi\mathbb{D}\eta}}.$$

Он принимает значения из $[-1, 1]$ и лучше отражает зависимость случайных величин ξ и η . Заметим также, что если $r_{\xi\eta} = 0$, то $\mathbb{D}[\xi + \eta] = \mathbb{D}\xi + \mathbb{D}\eta$.

Неравенства, связанные с моментами

Неравенства Коши-Буняковского:

$$|\mathbb{E}\xi\eta| \leq \sqrt{\mathbb{E}\xi^2\mathbb{E}\eta^2}$$

$$|\text{cov}(\xi, \eta)| \leq \sqrt{\mathbb{D}\xi\mathbb{D}\eta}$$

Неравенство Йенсена:

$$f(\mathbb{E}\xi) \leq \mathbb{E}f(\xi), \quad f - \text{выпуклая функция}$$

Неравенство Юнга:

$$\mathbb{E}|\xi\eta| \leq \frac{\mathbb{E}|\xi|^p}{p} + \frac{\mathbb{E}|\eta|^q}{q}, \quad \frac{1}{p} + \frac{1}{q} = 1, \quad p, q \geq 0$$

Неравенство Гёльдера:

$$\mathbb{E}|\xi\eta| \leq (\mathbb{E}|\xi|^p)^{\frac{1}{p}} (\mathbb{E}|\eta|^q)^{\frac{1}{q}}, \quad \frac{1}{p} + \frac{1}{q} = 1, \quad p, q \geq 0$$

Неравенство Минковского:

$$(\mathbb{E}|\xi + \eta|^p)^{\frac{1}{p}} \leq (\mathbb{E}|\xi|^p)^{\frac{1}{p}} + (\mathbb{E}|\eta|^p)^{\frac{1}{p}}, \quad p \geq 1$$

Неравенство Ляпунова:

$$(\mathbb{E}|\xi|^p)^{\frac{1}{p}} \leq (\mathbb{E}|\xi|^q)^{\frac{1}{q}}, \quad 0 < p \leq q$$

Упражнение 8. Пусть ξ и η случайные величины такие, что $\mathbb{E}\xi = 0$, $\mathbb{E}\eta = 0$, $\mathbb{D}\xi = 1$, $\mathbb{D}\eta = 1$, $\text{cov}(\xi, \eta) = \rho$. Доказать, что

$$\mathbb{E} \max\{\xi^2, \eta^2\} \leq 1 + \sqrt{1 - \rho^2}$$

Решение.

$$\mathbb{E}(\xi - \eta)^2 = \mathbb{E}\xi^2 + \mathbb{E}\eta^2 - 2\mathbb{E}\xi\eta = 2 - 2\rho$$

$$\mathbb{E}(\xi + \eta)^2 = \mathbb{E}\xi^2 + \mathbb{E}\eta^2 + 2\mathbb{E}\xi\eta = 2 + 2\rho$$

Получаем, используя неравенство Гелдера:

$$\begin{aligned} \mathbb{E} \max\{\xi^2, \eta^2\} &= \mathbb{E}\left(\frac{1}{2}(\xi^2 + \eta^2) + \frac{1}{2}|\xi^2 - \eta^2|\right) = 1 + \frac{1}{2}\mathbb{E}(|\xi + \eta| \cdot |\xi - \eta|) \leq \\ &\leq 1 + \frac{1}{2}\sqrt{\mathbb{E}(\xi - \eta)^2 \mathbb{E}(\xi + \eta)^2} = 1 + \frac{1}{2}\sqrt{4(1 - \rho)(1 + \rho)} = 1 + \sqrt{1 - \rho^2} \end{aligned}$$

□

Упражнение 9. Дивергенцией Кульбака-Лейблера между абсолютно непрерывными распределениями с плотностями p, q называется величина

$$KL(p||q) \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} \ln \frac{p(x)}{q(x)} \cdot p(x) dx$$

Она часто возникает в теории информации и мат. статистике как мера близости двух распределений (“истинного” распределения p и некоторого распределения q). Покажите, что она неотрицательна.

Решение. Пусть у нас есть случайная величина ξ с плотностью $p(x)$. Рассмотрим борелевскую функцию $\ln \frac{p(x)}{q(x)}$. Тогда воспользуемся неравенством Йенсена:

$$KL(p||q) = \mathbb{E}(\ln \frac{p(x)}{q(x)}) = \mathbb{E}(-\ln \frac{q(x)}{p(x)}) \geq -\ln \mathbb{E} \frac{q(x)}{p(x)} = -\ln \int_{-\infty}^{+\infty} \frac{q(x)}{p(x)} \cdot p(x) dx = 0$$

□

Приложение. Интеграл Лебега по вероятностной мере и интеграл Стильтьеса

Пусть задано вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$. Интеграл Лебега по вероятностной мере (далее в этом разделе будем называть это просто интегралом) от произвольной измеримой функции определяется в 3 этапа. Начнём с определения для **простых функций**.

Определение 7. Случайная величина $\xi(\omega)$ называется **простой**, если множество её значений конечно (частным случаем простых являются дискретные).

Определение 8. Индикатором множества $A \in \mathcal{F}$ называется простая функция, заданная на элементах Ω следующим образом:

$$\mathbb{I}_A(\omega) = \begin{cases} 1, & \text{если } \omega \in A, \\ 0, & \text{иначе.} \end{cases}$$

Заметим, что любая простая функция $\xi(\omega)$ представима в виде

$$\xi(\omega) = \sum_{k=1}^n a_k \mathbb{I}_{A_k}(\omega),$$

где $a_k, k = 1, \dots, n$ — **различные** значения, принимаемые ξ , а $A_k = \{\omega \mid \xi(\omega) = a_k\}$. Множества A_k попарно не пересекаются и $\bigcup_{k=1}^n A_k = \Omega$.

Определение 9. Интегралом от простой случайной величины $\xi(\omega)$ называется число

$$\int \xi d\mathbb{P} \stackrel{\text{def}}{=} \int \xi(\omega) d\mathbb{P}(\omega) = \sum_{k=1}^n a_k \mathbb{P}\{A_k\}.$$

Интегралом по множеству $A \in \mathcal{F}$ от простой измеримой функции $\xi(\omega)$ называется

$$\int_A \xi d\mathbb{P} \stackrel{\text{def}}{=} \int \xi(\omega) \mathbb{I}_A(\omega) d\mathbb{P}(\omega).$$

Теперь определим интеграл от положительной случайной величины.

Лемма 1. Пусть случайная величина $\xi(\omega) \geq 0$. Тогда существует последовательность $\{\xi_n(\omega)\}_{n=1}^\infty$ измеримых простых функций такая, что $\xi_n \uparrow \xi$ при $n \rightarrow \infty$ (сходимость поточечная, т. е. для каждого $\omega \hookrightarrow \xi(\omega) = \lim_{n \rightarrow \infty} \xi_n(\omega)$).

Определение 10. Интегралом неотрицательной функции $\xi(\omega)$ называется число

$$\int \xi d\mathbb{P} = \lim_{n \rightarrow \infty} \int \xi_n d\mathbb{P},$$

где ξ_n — последовательность простых измеримых функций таких, что $\xi_n \uparrow \xi$ при $n \rightarrow \infty$. Будем говорить, что интеграл $\int \xi d\mathbb{P}$ **существует**, а ξ **интегрируема**, если $\int \xi d\mathbb{P} < \infty$.

Интеграл неотрицательной функции корректно определён, т. к. его значения не зависят от выбора последовательности простых функций ξ_n монотонно сходящихся к ξ .

Определение 11. Интегралом от произвольной измеримой функции $\xi(\omega)$ называется число

$$\int \xi d\mathbb{P} \stackrel{\text{def}}{=} \int \xi^+ d\mathbb{P} - \int \xi^- d\mathbb{P}, \quad \xi^\pm = \max\{0, \pm \xi\},$$

если хотя бы одно из значений $\int \xi^\pm d\mathbb{P}$ конечно. В противном случае говорят, что **интеграла не существует**, а функция ξ **не интегрируема**.

Нетрудно показать, что $\int \xi d\mathbb{P} < \infty$ тогда и только тогда, когда $\int |\xi| d\mathbb{P} < \infty$. Кроме того, если существует $\int \xi d\mathbb{P}$, то существует и $\int_A \xi d\mathbb{P} = \int \xi \mathbb{I}_A d\mathbb{P}$ для любого $A \in \mathcal{F}$.

Выполняются привычные свойства интегралов.

1. Если множества $A_i \in \mathcal{F}$ попарно не пересекаются и $\bigcup_i A_i = \Omega$, то

$$\int \xi d\mathbb{P} = \sum_i \int_{A_i} \xi d\mathbb{P}.$$

2. $\int (\xi + \eta) d\mathbb{P} = \int \xi d\mathbb{P} + \int \eta d\mathbb{P}.$

3. Если a — произвольная постоянная, то

$$\int a\xi d\mathbb{P} = a \int \xi d\mathbb{P}.$$

4. Если $\xi \leq \eta$, то $\int \xi d\mathbb{P} \leq \int \eta d\mathbb{P}.$

5. $|\int \xi d\mathbb{P}| \leq \int |\xi| d\mathbb{P}.$

6. Если $c_1 \leq \xi \leq c_2$, то $c_1 \leq \int \xi d\mathbb{P} \leq c_2$ (здесь активно используется тот факт, что \mathbb{P} — вероятностная мера).

7. Если $\mathbb{P}\{\xi = \eta\} = 1$ и $\int \xi d\mathbb{P}$ существует, то $\int \xi d\mathbb{P} = \int \eta d\mathbb{P}.$

Теорема 1. Пусть $g(x)$ — борелевская функция на прямой \mathbb{R} . Определим случайную величину $\eta = g(\xi(\omega))$. Если $\int \eta d\mathbb{P}$ существует, то

$$\int \eta d\mathbb{P} = \int_{\mathbb{R}} g(x) d\mathbb{P}_\xi(x).$$

Интеграл в правой части может быть записан также в форме

$$\int g(x) dF_\xi(x).$$

В таком виде он называется **интегралом Лебега-Стильтьеса** от функции $g(x)$ по распределению F_ξ . Более того, если $g(x)$ — непрерывная функция, то интеграл Лебега-Стильтьеса совпадает с интегралом **Римана-Стильтьеса**, равным по определению

$$\int g(x) dF_\xi(x) = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow +\infty}} \lim_{N \rightarrow \infty} \sum_{k=0}^{N-1} g(\tilde{x}_k) [F_\xi(x_{k+1}) - F_\xi(x_k)],$$

где предел в правой части берётся при мелкости разбиения $\max_k (x_{k+1} - x_k) \xrightarrow[N \rightarrow \infty]{} 0$, где $a = x_0 < x_1 < \dots < x_N = b$ и $\tilde{x}_k \in [x_k, x_{k+1})$. Можно показать, что предел не зависит от выбора разбиения. Более того, в силу Теоремы 1 мы знаем, что все свойства интеграла Лебега по вероятностной мере сохраняются и для интеграла Римана-Стильтьеса в случае непрерывной $g(x)$.

Перед тем как двигаться дальше рассмотрим следующее упражнение.

Пусть $F(x)$ — произвольная функция распределения. Тогда её можно представить в виде суммы $F(x) = F_H(x) + F_D(x)$ непрерывной и дискретных компонент. Действительно, в силу теоремы Лебега мы знаем, что любую вероятностную меру на прямой можно представить единственным способом в виде суммы дискретной, сингулярной и абсолютно непрерывных мер. В частности, для функции распределения $F(x)$ рассмотрим вероятностную меру на прямой, которую она задаёт, и обозначим её $\mu = \mu_d + \mu_s + \mu_{ac}$. Как мы помним из предыдущего семинара, и сингулярная, и абсолютно непрерывные меры являются атомарными, т. е. $\forall x \in \mathbb{R} \hookrightarrow \mu_s\{x\} = \mu_{ac}\{x\} = 0$, а значит, функция $F_H(x) \stackrel{\text{def}}{=} \mu_s\{(-\infty, x)\} + \mu_{ac}\{(-\infty, x)\}$ является непрерывной. Функция $F_D \stackrel{\text{def}}{=} \mu_d\{(-\infty, x)\}$ является кусочно-постоянной и имеет не более чем счётное число точек разрыва. Поэтому в силу равенства $F(x) = \mu\{(-\infty, x)\}$ мы получаем $F(x) = F_H(x) + F_D(x)$. Пусть x_1, x_2, \dots — точки разрывов $F_D(x)$:

$$p_k = F_D(x_k + 0) - F_D(x_k) > 0.$$

Тогда по определению

$$\int g(x) dF(x) = \sum_k p_k g(x_k) + \int g(x) dF_H(x).$$

Рассмотрим два важнейших частных случая.

1. **Дискретное распределение.** Из определения интеграла Стильтьеса получаем, что в случае дискретного распределения интеграл превращается в сумму (функция распределения F является ступенчатой функцией):

$$\int g(x) dF(x) = \sum_k g(x_k) (F(x_k + 0) - F(x_k)) = \sum_k g(x_k) \mathbb{P}\{\xi = x_k\},$$

где x_1, x_2, \dots — точки скачков $F(x)$.

2. **Абсолютно непрерывное распределение.** В случае абсолютно непрерывного распределения $F(x) = \int_{-\infty}^x f(t) dt$ интеграл Стильтьеса превращается в интеграл Римана:

$$\int g(x) dF(x) = \int g(x) f(x) dx.$$

Некоторые свойства интеграла Стильтьеса, вытекающие из определения:

$$1) \int_a^b dF = F(b) - F(a);$$

$$2) \int_a^b g dF = \int_a^c g dF + \int_c^b g dF;$$

$$3) \int (\alpha g + \beta h) dF = \alpha \int g dF + \beta \int h dF, \text{ где } a, b = \text{const.}$$