ЛЕКЦИИ ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ

А.В. Колесников

Содержание

1. Слабая сходимость мер. Меры на полных сепарабельных метрических пространствах. Теорема Прохорова.

Приложения слабой сходимости мер не исчерпываются вероятностными задачами. По своей природе это понятие функционально-аналитическое, поскольку может рассматриваться как *-слабая сходимость на пространстве мер, но оно также широко используется в вариационных задачах и уравнениях в частных производных. Первые фундаментальные результаты о слабой сходимости мер были получены в работах выдающегося геометра А.Д. Александрова, при этом они были мотивированы его работами по теории выпуклых поверхностей.

Пример 1.1. *Рассмотрим δ-образную последовательность функций $f_\delta(x) = \frac{1}{\sqrt{2\pi\delta^2}} e^{-\frac{x^2}{2\delta^2}}$ на прямой. Каждая функция удовлетворяет условию $\int_{\mathbb{R}} f_\delta(x)\ dx = 1$. Очевидно, $\lim_{\delta\to 0} f_\delta(x) = 0$ если $x \ne 0$ и $\lim_{\delta\to 0} f_\delta(0) = +\infty$. Несмотря на то, что $\lim_{\delta\to 0} f_\delta(x) = 0$ почти всюду, функция, тождественно равная нулю, не является разумным пределом $\{f_\delta\}$. Действительно, как хорошо известно из анализа (и нетрудно проверить самостоятельно), $\lim_{\delta\to 0}\int \varphi(x) f_\delta(x)\ dx = \varphi(0)$ для непрерывной ограниченной функции φ. Очевидно, в этом примере не выполняется свойство о предельном переходе под знаком интеграла $\lim_{\delta\to 0}\int \varphi f_\delta(x)\ dx \ne 0$. Естественным решением данной проблемы является следующее: рассматривать $\{f_\delta\}$ как последовательность* **мер** *$\{f_\delta(x)\ dx\}$. Пределом этой последовательности является мера Дирака δ_0, сосредоточенная в точке 0, т.е. мера, определенная условием $\delta(A) = 0$, если $0 \in A$ и $\delta(A) = 1$, если $0 \in A$. Очевидно, $\int f(x)\ d\delta = f(0)$.*

Естественным классом пространств, на которых удобно рассматривать слабую сходимость, является класс так называемых польских пространств.

Определение 1.2. *Польскими пространствами называются полные сепарабельные метрические пространства.*

Мы будем рассматривать борелевские вероятностные меры на польском пространстве X, то есть, меры, определенные на борелевской σ-алгебре $\mathcal{B}(X)$. Метрику X будем обозначать через ρ.

Определение 1.3. *Пусть $\{\mu_n\}$ — последовательность вероятностных мер. Будем говорить, что $\{\mu_n\}$ сходится слабо к вероятностной мере $\{\mu\}$, если для любой непрерывной ограниченной функции $f : X \to \mathbb{R}$ выполнено соотношение*

$$\lim_{n\to\infty} \int_X f\ d\mu_n = \int_X f\ d\mu.$$

Теорема 1.4. *(А.Д. Александров) Последовательность вероятностных мер $\{\mu_n\}$ слабо сходится к вероятностной мере μ тогда и только тогда, когда выполнено одно из условий*

1)
$$\mu(G) \geq \overline{\lim}_n \mu_n(G)$$
для любого замкнутого G

2)
$$\mu(O) \leq \underline{\lim}_n \mu_n(O)$$
для любого открытого O

3) *Для любого множества B, удовлетворяющего условию $\mu(\partial B) = 0$*
$$\mu(B) = \lim_n \mu_n(B).$$

Доказательство. Предположим, что $\mu_n \to \mu$ слабо. Пусть G — замкнутое множество. Рассмотрим функцию $x \to \rho(x, G) = \inf_{y \in G} \rho(x, y)$. Заметим, что эта функция непрерывна (почему?). Положим $f_m(x) = (1 + m\rho(x, G))^{-1}$. Заметим, что $f_m(x) \geq I_G(x)$. По определению слабой сходимости имеем
$$\int f_m(x)\ d\mu = \lim_n \int f_m(x)\ d\mu_n \geq \overline{\lim}_n \mu_n(G).$$
Переходя к пределу по m и пользуясь теоремой Лебега об ограниченной сходимости, получаем $\lim_m \int f_m(x)\ d\mu = \mu(G)$, следовательно
$$\mu(G) \geq \overline{\lim}_n \mu_n(G).$$
Очевидно, свойства 1) и 2) эквивалентны.

Далее, 3) легко следует из 1) + 2), если учесть, что $\mathrm{Int}B = B \setminus \partial B \subset B \subset \overline{B}$ и
$$\mu(\mathrm{Int}B) = \mu(B) = \mu(\overline{B}),$$
если $\mu(\partial B) = 0$.

Для вывода слабой сходимости из свойства 3) заметим, что достаточно доказать свойство
$$\lim_n \int f(x)\ d\mu_n = \int f(x)\ d\mu$$
для произвольной *неотрицательной* функции $f(x)$ Пусть $M = \sup_X f(x)$. Применим известную формулу
$$\int f(x)\ d\mu_n = \int_0^M \mu_n(\{f(x) > t\})\ dt.$$
В силу непрерывности $f(x)$ имеем $\partial\{x : f(x) > t\} = \{x : f(x) = t\}$. Так как множества $\{x : f(x) = t\}$ не пересекаются при различных t, то $\mu(\{x : f(x) = t\}) > 0$ только для не более чем счетного набора значений t. Для всех остальных $\lim_n \mu_n(\{f(x) > t\}) = \mu(\{f(x) > t\})$. Преходя к пределу под знаком интеграла (пользуясь теоремой об ограниченной сходимости), получаем
$$\lim_n \int f(x)\ d\mu_n = \lim_n \int_0^M \mu_n(\{f(x) > t\})\ dt = \int_0^M \mu(\{f(x) > t\})\ dt = \int f(x)\ d\mu.$$
□

Следующий факт известен как теорема Улама.

Теорема 1.5. *Пусть μ — вероятностная борелевская мера на польском пространстве X. Для любого $\varepsilon > 0$ существует такое компактное множество $K \subset X$, что $\mu(X \setminus K) \leq \varepsilon$.*

Доказательство. Для любого ε и любого $n \in \mathbb{N}$ найдем конечное объединение $A_n = \cup_{i=1}^{N(n)} B_{1/n}(x_i)$ шаров радиса $1/n$, удовлетворяющих условию $\mu(X \setminus A_n) < \frac{\varepsilon}{2^n}$. Это можно сделать в силу того, что все пространство является X является счетным объединением таких шаров. Заметим, что A_n обладает конечной $1/n$-сетью (центры шаров из A_n). Пусть $K = \cap_{n=1}^{\infty} A_n$. Так как $K_n \subset A_n$, то K обладает конечной ε - сетью для любого $\varepsilon > 0$. Поэтому K — вполне ограниченное, замкнутое множество в полном метрическом пространстве. Следовательно, оно является компактом. Кроме того, $\mu(X \setminus K) = \mu(\cup_{n=1}^{\infty} A_n^c) \leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon$. □

Определение 1.6. *Слабой топологией на пространстве вероятностных мер называется топология, порожденная множествами вида $O_{\mu,f,\varepsilon} = \big\{\nu : |\int_X f d\nu - \int_X f d\mu| < \varepsilon\big\}$, f — непрерывная ограниченная функция.*

Будем говорить, что множество $\mathcal{M}$ вероятностных мер слабо компактно, если из любой последовательности $\{\mu_n\} \subset \mathcal{M}$ можно выделить слабо сходящуюся подпоследовательности. Строго говоря, это свойство следует называть слабой секвенциальной компактностью. Впрочем (см. задачи), слабая сходимость на метрических пространствах метризуема. Следовательно, в нашей ситуации эти понятия эквивалентны. Семейство мер $\mathcal{M}$ вероятностных мер на польском пространстве X называется плотным, если для любого $\varepsilon > 0$ существует такой компакт $K \subset X$, что $\mu(X \setminus K) \le \varepsilon$ для любой меры $\mu \in \mathcal{M}$.

Следующий результат (теорема Прохорова) играет исключительно важную роль в приложениях теории слабой сходимости мер. Для его доказательства понадобится еще один классический результат функционального анализа.

Теорема 1.7. *(Ф. Рисс) Пусть K—метрический компакт, $C(K)$ — пространство непрерывных функций на K, наделенное равномерной нормой. Для любого непрерывного линейного функционала $L : C(K) \to \mathbb{R}$ существует такая борелевская (в общем случае, знакопеременная) мера μ, что имеет место представление*

$$L(f) = \int_K f \ d\mu.$$

В случае, если L — неотрицательный функционал (т.е. $L(f) \ge 0$ для $f \ge 0$), то μ — неотрицательная мера.

Теорема 1.8. *(Ю.В. Прохоров) Семейство $\mathcal{M}$ вероятностных мер на польском пространстве X тогда и только тогда слабо компактно, когда оно плотно.*

Доказательство. Доказательство необходимости близко доказательству теоремы Улама. Пусть $\{x_i\}$ — счетное всюду плотное множество в X. Положим $A_{m,r} = \cup_{i \le m} B_r(x_i)$, $B_r(x_i) = \{x : |x - x_i| \le r\}$. Докажем, что

$$\lim_m \inf_{\mu \in \mathcal{M}} \mu(A_{m,r}) = 1$$

для любого $r > 0$. Предположим, что это не так. Зафиксируем $r > 0$. Тогда для некоторого $\varepsilon > 0$ и для любого m можно найти такую меру μ_m, что $\mu_m(A_{m,r}) < 1 - \varepsilon$. Пользуясь слабой компактностью $\mathcal{M}$ и переходя к подпоследовательностям, без ограничения общности можно считать, что $\{\mu_m\}$ сходится слабо к некоторой мере μ. Для меры μ существует такое натуральное число M, что $\mu(A_{M,r}) > 1 - \frac{\varepsilon}{2}$. Но в силу слабой сходимости $1 - \frac{\varepsilon}{2} < \mu(A_{M,r}) \le \underline{\lim}_m \mu_m(A_{M,r}) \le \underline{\lim}_m \mu_m(A_{m,r}) < 1 - \varepsilon$. Мы получили противоречие. Следовательно, для всех $r > 0$ $\lim_m \inf_{\mu \in \mathcal{M}} \mu(A_{m,r}) = 1$. Отсюда немедленно следует, что для любого $\varepsilon > 0$ и $N \in \mathbb{N}$ существует такое число $M(N)$, что

$$\mu(A_{M(N),1/N}) = \mu(\cup_{i \le M(N)} B_{1/N}(x_i)) > 1 - \frac{\varepsilon}{2^N}$$

для любого $\mu \in \mathcal{M}$. Множество

$$\cap_{N \in \mathbb{N}} \cup_{i \le M(N)} B_{1/N}(x_i)$$

является искомым компактом.

Докажем достаточность. Предположим сначала, что $X = K$ — компакт. В пространстве $C(K)$ непрерывных функций на K найдем счетное всюду плотное множество функций $\{f_j\}$ (почему такое существует?). Так как для каждого f_j число

$$\int_K f_j \ d\mu$$

ограничено числом $\sup_K |f_j|$, то, пользуясь диагональным методом, можно найти последовательность $\{\mu_n\}$, т.ч. предел

$$\lim_n \int f_j \ d\mu_n = F(f_j)$$

существует. Заметим, что $|F(f_j)| \le \sup_K |f_j|$ и $F(f_j) \ge 0$ для любой неотрицательной функции f_j. Следовательно, F продолжается по непрерывности до неотрицательного линейного функционала на $C(K)$. По теореме Рисса, существует вероятностная мера μ, обладающая свойством $F(f) = \int_K f \ d\mu$. При этом $\lim_n \int f_j \ d\mu_n = \int f_j \ d\mu$. Найдем произвольную непрерывную функцию f и приблизим ее по равномерной норме некоторой функцией f_j: $|f - f_j| < \varepsilon$. Тогда $\left|\int f \ d\mu - \lim_n \int f \ d\mu_n\right| \le \left|\int f_j \ d\mu - \lim_n \int f_j \ d\mu_n\right| + \left|\int (f - f_j) \ d\mu\right| + \left|\lim_n \int (f - f_j) \ d\mu_n\right| \le 2\varepsilon$. В силу произвольности ε, получаем $\int f \ d\mu = \lim_n \int f \ d\mu_n$.

В случае некомпактного пространства X надо воспользоваться известным из общей топологии фактом, что X гомеоморфно вкладывается в некоторое компактное метрическое пространство Y. В качестве Y можно взять, например, $[0,1]^{\mathbb{N}}$. Таким образом, можно считать, что $X \subset Y$ и любая мера μ продолжается на все Y по формуле $\mu(A) = \mu(A \cap X)$, $A \subset Y$. Можно выделить последовательность мер $\{\mu_n\}$, слабо сходящуюся на Y к мере ν. Возьмем последовательность компактов $K_m \subset X$, т.ч.

$$\sup_{\mu \in \mathcal{M}} \mu(X \setminus K_m) \le \frac{1}{m}.$$

Тогда по свойству 1) теоремы 1.4 $\nu(K_m) \ge \overline{\lim}_n \mu_n(K_m) \ge 1 - \frac{1}{m}$ (так как $K_m \subset X \subset Y$ — компакт, и, следовательно, замкнут в Y). Поэтому $\nu(\cup_m K_m) = 1$ и $\nu(X) = 1$. Докажем, что $\{\mu_n\}$ (рассматриваемые как меры на X) слабо сходятся к $\nu|_X$. Действительно, возьмем открытое множество U в X. Оно является пересечением некоторого открытого множества $V \subset Y$ с X. В силу сходимости $\mu_n \to \nu$ имеем $\underline{\lim}_n \mu_n(U) = \underline{\lim}_n \mu_n(V) \ge \nu(V) = \nu(V \cap X) = \nu(U)$. По свойству 2) теоремы 1.4 $\mu_n \to \nu$ слабо, как меры на X. $\square$

Замечание 1.9. *Достаточность также можно доказывать, используя теорему Банаха-Алаоглу о слабой компактности единичного шара.*

Занятие 1

1) Докажите, что последовательность вероятностных мер $\{\mu_n\}$ на прямой сходится слабо к мере μ тогда и только тогда, когда $F_{\mu_n}(t) \to F_\mu(t)$ для всех точек непрерывности функции $F_\mu(t)$.
2) Докажите, что последовательность мер Дирака $\{\delta_{x_n}\}$ на метрическом пространстве X тогда и только тогда слабо сходится, когда $x_n \to x \in X$. В частности, пространство X топологически вкладывается в пространство вероятностных мер на X, наделенное слабой топологией.
3) Докажите, что любая вероятностная мера на $\mathbb{R}^n$ является слабым пределом мер вида
$$\sum_{i=1}^{n} c_i \cdot \delta_{x_i}$$
(дискретных мер).
4) Пусть $F : X \to Y$ — непрерывное отображение метрических пространств, $\{\mu_n\}$ — слабо сходящаяся последовательность мер на X. Доказать, что последовательность мер-образов $\mu_n \circ F^{-1}$ является слабо сходящейся.
5) Пусть $\{\xi_n\}$ — последовательность случайных векторов, причем $\xi = \lim_n \xi_n$ почти наверное. Используя теорему об ограниченной сходимости, доказать, что $\mu_{\xi_n} \to \mu_\xi$ слабо.

6*) (теорема Скорохода) Пусть $\mu_n \to \mu$ — слабо сходящаяся последовательность вероятностных мер на $\mathbb{R}^n$. Доказать, что существует вероятностное пространство $(\Omega, \mathcal{F}, P)$ и такая последовательность случайных векторов $\{\xi_n\}$, что $\xi_n \to \xi$ п.н. и $\mu_{\xi_n} = \mu_n$, $\mu_\xi = \mu$.

7) Пусть $\{\mu_n\}$ — последовательность вероятностных мер на $\mathbb{R}^{\mathbb{N}}$, причем проекции $\{\mu_n\} \circ x_i^{-1}$ на i-ю координату образуют плотную последовательность мер на прямой. Доказать, что $\{\mu_n\}$ — плотная последовательность мер на $\mathbb{R}^{\mathbb{N}}$.

8*) (метризуемость слабой сходимости) Пусть μ, ν — вероятностные меры на $\mathbb{R}^n$. Положим $d(\mu,\nu) = \inf\{\varepsilon > 0,\ \mu(A) \le \nu(A^\varepsilon) + \varepsilon, A - \text{борелевское}\}$. Докажите, что 1) d — метрика, 2) $d(\mu_n, \mu) \to 0$ тогда и только тогда, когда $\mu_n \to \mu$ слабо.

9) Построить пример слабо сходящейся последовательности мер $\mu_n \to \mu$ на X и такого замкнутого (открытого) множества $A \subset X$, что сужения $\mu_n|_A$ на A не сходятся слабо к $\mu|_A$ на X.
10) Доказать, что семейство мер $\mathcal{M}$ на метрическом пространстве X является плотным, тогда и только тогда, когда существует такая функция $f : X \to [0, +\infty]$, что $\sup_{\mu \in \mathcal{M}} \int f \, d\mu < \infty$ и множества $\{x : f(x) < n\}$ компактны для любого $n \in \mathbb{N}$.

2. Характеристические функции. Теорема единственности. Теорема Бохнера. Характеристические функции и слабая сходимость.

В этом разделе мы будем иметь дело с комплекснозначными случайными величинами. Комплекснозначную случайную величину $\zeta = \xi + i\eta$ можно представлять себе как пару действительнозначных с.в. Будем считать, что $\mathbb{E}\zeta = \mathbb{E}\xi + i\mathbb{E}\eta$.

Определение 2.1. *Характеристической функцией (преобразованием Фурье) случайного вектора $\xi = (\xi_1, \cdots, \xi_n)$ называется функция*

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

Эта же функция называется характеристической функцией меры $\mu_\xi = P \circ \xi^{-1}$.

Если ξ имеет плотность f_ξ, то

$$\varphi_\xi(t) = \int e^{i(t,x)}\ f_\xi(x)\ dx.$$

Заметим, что для дискретных с.в. характеристическая функция совпадает с производящей с точностью до замены переменных.

Пример 2.2. *Для бернуллиевской с.в. имеем*

$$\varphi_\xi(t) = pe^{it} + q.$$

Пример 2.3. *Пусть* $\xi \sim \mathcal{N}(0,1)$. *Тогда*

$$\varphi_\xi(t) = e^{-\frac{t^2}{2}}.$$

Если $\xi \sim \mathcal{N}(a,\sigma^2)$, *то* $\varphi_\xi(t) = e^{iat-\frac{t^2\sigma^2}{2}}$.

Упражнение 2.4. *Докажите это соотношение. Указание: разложите* e^{it} *в ряд Тейлора и найдите полученный ряд, применив формулу для гауссовских моментов* $\mathbb{E}\xi^{2n} = (2n-1)!!$. *Также этот результат можно получить методами теории функций комплексного переменного.*

Теорема 2.5. *(Единственность) Если* $\varphi_\xi = \varphi_\eta$, *то распределения* ξ *и* η *совпадают.*

Доказательство. Достаточно показать (почему?), что для любой непрерывной функции f с компактным носителем, содержащемся в кубе $K = [-\pi k, \pi k]^n$, и такой, что $\sup_{x\in\mathbb{R}^n} |f(x)| \le 1$, имеем:

$$\int_{\mathbb{R}^n} f\ d\mu_\xi = \int_{\mathbb{R}^n} f\ d\mu_\eta.$$

По теореме Вейерштрасса-Стоуна найдем тригонометрический многочлен f_ε вида

$$f_\varepsilon(x) = \sum_{j=0}^{N(m)} c_j \exp(i(y_j,x)),$$

т.ч. $\sup_K |f - f_\varepsilon| < \varepsilon$. В силу периодичности $\sup_{\mathbb{R}^n} |f_\varepsilon(x)| < 1 + \varepsilon$. Так как

$$\int f_\varepsilon\ d\mu_\xi = \int f_\varepsilon\ d\mu_\eta,$$

то

$$\begin{aligned}\Big|\int f d\mu_\xi - \int f d\mu_\eta\Big| &\le \int |f - f_\varepsilon|\ d\mu_\xi + \int |f - f_\varepsilon|\ d\mu_\eta \\ &= \int_K |f - f_\varepsilon|\ d\mu_\xi + \int_K |f - f_\varepsilon|\ d\mu_\eta + \int_{K^c} |f - f_\varepsilon|\ d\mu_\xi + \int_{K^c} |f - f_\varepsilon|\ d\mu_\eta \\ &\le 2\varepsilon + (1+\varepsilon)(\mu_\xi(K^c) + \mu_\eta(K^c)).\end{aligned}$$

Выбирая достаточно большой куб K, можно добиться, что $\mu_\xi(K^c) + \mu_\eta(K^c) \le \frac{\varepsilon}{1+\varepsilon}$. Тогда $\Big|\int f d\mu_\xi - \int f d\mu_\eta\Big| \le 3\varepsilon$. В силу произвольности ε получаем $\int f d\mu_\xi = \int f d\mu_\eta$. □

Справедливы различные формулы обращения для характеристических функций. Например, если $n = 1$, то верна следующая формула

$$F_\xi(b) - F_\xi(a) = \lim_{s\to\infty} \frac{1}{2\pi} \int_{-s}^{s} \frac{e^{-isa} - e^{-itb}}{it} \varphi_\xi(t)\ dt.$$

Предположим, что распределение случайного вектора ξ непрерывно и $|\varphi_\xi| \in L^1(\mathbb{R}^n)$. В курсе анализа доказывается формула обращения

$$f_\xi(x) = \frac{1}{(2\pi)^n} \int e^{-i(t,x)} \varphi_\xi(t)\ dt.$$

Следующие свойства проверяются элементарными выкладками.

Теорема 2.6. *(Основные свойства характеристических функций)*

1) $|\varphi_\xi(t)| \le |\varphi_\xi(0)| = 1$.
2) φ *равномерно непрерывна по* t.

3) $\varphi(t) = \overline{\varphi(-t)}$.

Теорема ниже сформулирована для простоты изложения в одномерном случае. Читатель без труда сформулирует многомерный аналог.

Теорема 2.7. *(Дифференцируемость и конечность моментов) Если* $\mathbb{E}|\xi|^n < \infty$*, то* φ *дифференцируема* n *раз, причем*

$$\varphi^{(n)}_\xi(t) = \int (ix)^n e^{ixt}\, d\mu_\xi.$$

Доказательство. Заметим, что

$$\frac{\varphi(t+h) - \varphi(t)}{h} = \int e^{itx} \frac{e^{ixh} - 1}{h}\, d\mu_\xi(x).$$

В силу того, что $\bigl|\frac{e^{ixh}-1}{h}\bigr| \le |x|$, получаем

$$\Bigl|\frac{\varphi(t+h) - \varphi(t)}{h}\Bigr| \le \int |x|\, d\mu_\xi(x) < \infty.$$

В силу теоремы о мажорированной сходимости существует

$$\lim_{h\to 0} \frac{\varphi(t+h) - \varphi(t)}{h} = \lim_{h\to 0} \int e^{itx} \frac{e^{ixh} - 1}{h}\, d\mu_\xi(x) = i \int x e^{itx}\, d\mu_\xi(x).$$

Для $n > 1$ применяем индукцию. □

Замечание 2.8. *Верно утверждение, которое можно считать обратным. Если существует* $\varphi^{(2n)}(0)$*, то* $\mathbb{E}\xi^{2n} < \infty$*. Доказательство см. в* [4] *(т.1, 2(13)),* [2]*. Интересно отметить, что существование* $\varphi'(0)$ *не влечет, вообще говоря, конечность первого момента* ξ *(см. пример в* [2]*).*

Теорема 2.9. *(Независимость и характеристические функции). Компоненты случайного вектора* $\xi = (\xi_1, \cdots, \xi_n)$ *независимы тогда и только тогда, когда* φ_ξ *разлагается в произведение одномерных характеристических функций.*

$$\varphi_\xi(t_1, \cdots, t_n) = \prod_{i=1}^n \varphi_{\xi_i}(t_i).$$

Если с.в. ξ *является суммой независимых с.в.* $\xi = \sum_{i=1}^n \xi_i$*, то ее характеристическая функция является произведением характеристических функций этих с.в.*

$$\varphi_\xi(t) = \prod_{i=1}^n \varphi_{\xi_i}(t).$$

Пример 2.10. *Распределение Коши* $f_\xi = \frac{1}{\pi}\frac{1}{1+x^2}$ *имеет характеристическую функцию*

$$e^{-|t|}.$$

Это можно проверить методами комплексного анализа. Следовательно, если $\eta = \frac{\sum_{i=1}^n \xi_i}{n}$*, где* ξ_i *независимы и имеют распределение Коши, то* $\varphi_\eta = (e^{-|t|})^n(t/n) = e^{-|t|}$*. Следовательно,* $f_\eta = f_\xi$*.*

Теорема Бохнера-Хинчина

Комплекснозначная функция φ называется неотрицательно определенной, если

$$\sum_{k,l=1}^n \varphi(t_k - t_l)\lambda_k \overline{\lambda_l} \ge 0$$

для любых действительных $t_k \in \mathbb{R}^n$ и комплексных $\lambda_k \in \mathbb{C}$. Любая характеристическая функция является неотрицательно определенной, так как

$$\sum_{k,l} \mathbb{E}e^{i(t_k - t_l, x)}\lambda_k\overline{\lambda_l} = \mathbb{E}\Bigl|\sum_{k=1}^n \lambda_k e^{it_k x}\Bigr|^2 \ge 0.$$

Теорема 2.11. *Пусть* φ *— непрерывная функция и* $\varphi(0) = 1$*. Для того, чтобы* φ *была характеристической функцией некоторого случайного вектора, необходимо и достаточно, чтобы* φ *была неотрицательно определенной.*

Неравенство Чебышева

Теорема 2.12. *Любая с.в. ξ удовлетворяет следующим неравенствам*

$$P(|\xi|^p \geq t) \leq \frac{\mathbb{E}(|\xi|^p)}{t^p}, \ p \geq 1.$$

В частности, $P(|\xi - \mathbb{E}\xi| \geq t)) \leq \frac{D\xi}{t^2}$.

Слабая сходимость и характеристические функции

Теорема 2.13. *Последовательность вероятностных мер $\{\mu_n\}$ на $\mathbb{R}^d$ тогда и только тогда слабо сходится к вероятностной мере μ, когда соответствующие характеристические функции $\varphi_n(y) = \int_{\mathbb{R}^d} e^{i(x,y)}\ d\mu_n$ сходятся поточечно к характеристической функции $\varphi(y) = \int_{\mathbb{R}^d} e^{i(x,y)}\ d\mu$.*

Лемма 2.14. *Для двух вероятностных мер μ и ν имеет место равенство*

$$\int \varphi_\mu d\nu = \int \varphi_\nu d\mu.$$

Доказательство. Вытекает из теоремы Фубини. □

Из предыдущей леммы и неравенства Чебышева получаем следствие.

Следствие 2.15. *Если φ_ν действительна, то*

$$\mu(\varphi_\nu \leq t) \leq \frac{1}{1-t} \int \big(1 - \varphi_\mu(y)\big)\ d\nu(y).$$

Следствие 2.16. *Для всякой вероятностной борелевской меры μ на $\mathbb{R}^n$ справедливо неравенство*

$$\mu(x : |x| \geq t) \leq \frac{\sqrt{e}}{\sqrt{e}-1} \int_{\mathbb{R}^n} \big[1 - \varphi_\mu(y/t)\big]\ d\gamma(y),$$

где $\gamma = \frac{1}{(2\pi)^{n/2}} e^{-\frac{|x|^2}{2}}\ dx$ — стандартная гауссовская мера на $\mathbb{R}^n$.

Доказательство. Применим предыдущее следствие к мере $\nu = \gamma_t$, где γ_t — образ меры γ при отображении $x \to \frac{x}{t}$. Имеем $\varphi_\nu = e^{-\frac{x^2}{2t^2}}$. Тогда

$$\begin{aligned}\mu(|x| \geq t) = \mu(\varphi_{\gamma_t} \leq e^{-1/2}) &\leq \frac{1}{1-e^{-1/2}} \int_{\mathbb{R}^n} \big[1 - \varphi_\mu(y)\big]\ d\gamma_t(y) \\ &= \frac{\sqrt{e}}{\sqrt{e}-1} \int_{\mathbb{R}^n} \big[1 - \varphi_\mu(y/t)\big]\ d\gamma(y).\end{aligned}$$

□

Доказательство теоремы 2.13:

Доказательство. Очевидно, слабая сходимость влечет поточечную сходимость характеристических функций. Для доказательства обратного утверждения покажем, что последовательность мер μ_n плотна. В силу поточечной сходимости и теоремы о мажорированной сходимости

$$\lim_n \mu_n(|x| \geq t) \leq \lim_n \frac{\sqrt{e}}{\sqrt{e}-1} \int \big[1 - \varphi_{\mu_n}(y/t)\big]\ d\gamma(y) = \frac{\sqrt{e}}{\sqrt{e}-1} \int \big[1 - \varphi_\mu(y/t)\big]\ d\gamma(y).$$

Кроме того,

$$\lim_{t\to\infty} \int \big[1 - \varphi_{\mu_n}(y/t)\big]\ d\gamma(y) = 0$$

для любого n. Из этих двух свойств несложно извлечь, что

$$\lim_t \sup_{n>N} \mu_n(|x| \geq t) = 0$$

(что и означает плотность). Действительно, для $\varepsilon > 0$ найдем сначала такое t, что $\mu(|x| \geq t) \leq \frac{\sqrt{e}}{\sqrt{e}-1} \int \big[1 - \varphi_\mu(y/t)\big]\ d\gamma(y) \leq \varepsilon$. Следовательно, $\lim_n \mu_n(|x| \geq t) \leq \varepsilon$ и $\mu_n(|x| \geq t) \leq 2\varepsilon$ для достаточно больших $n > N$. Остается подобрать такое $T > t$, что $\mu_n(|x| > T) < \varepsilon$ при $n \leq N$. Следовательно,

$\sup_{n>N} \mu_n(|x| \geq T) \leq 2\varepsilon$ и плотность доказана. Отсюда следует, что $\mu_n \to \mu$ слабо. Действительно, иначе можно было бы найти подпоследовательность $\{\mu_{n_m}\}$, слабо сходящуюся к другой мере $\tilde{\mu}$ (мы пользуемся плотностью исходной последовательности и теоремой Прохорова). Но тогда $\lim_m \varphi_{n_m} = \varphi_{\tilde{\mu}}$. Это противоречит тому, что $\lim_m \varphi_{n_m} = \varphi_\mu$ и теореме единственности. □

Замечание 2.17. *1) Проверьте, что утверждение теоремы допускает следующее усиление: если $\varphi_n \to \varphi$ поточечно и $\varphi(y)$* **непрерывна в нуле***, то меры μ_n сходятся слабо к некоторой мере μ и φ является характеристической функцией μ.*

2) Требование непрерывности в нуле важно. Например, последовательность характеристических функций $\cos^n x$ сходится поточечно к разрывной функции, не являющейся характеристической.

Занятие 2

1) Доказать, что если φ — характеристическая функция, то $\overline{\varphi}, \varphi^2, |\varphi|^2, Re(\varphi)$ — характеристические функции. Построить пример, показывающий, что $|\varphi|$ и $Im(\varphi)$ могут не быть характеристическими функциями.
2) Найти характеристические функции распределений: биномиальное, пуассоновское, геометрическое, равномерное, гамма-распределение, хи-квадрат, Коши.
3) Пусть ξ_i — независимые стандартные гауссовские с.в. Найти характеристические функции распределений $\xi_1 \cdot \xi_2$, ξ_1/ξ_2, $\xi_1 \cdot \xi_2 + \xi_3 \cdot \xi_4$. Найти плотность распределения последней с.в.
4) Как по характеристической функции вычислить дисперсию с.в.? Корреляцию компонент случайного вектора?
5) Доказать, что если $\varphi(t) = 1 + o(t^2)$ при $t \to 0$, то $\varphi = 1$. Доказать, что e^{-t^p} не является характеристической функцией при $p > 2$.
6) Доказать, что если φ — характеристическая функция, то $e^{\lambda(\varphi-1)}$ тоже является характеристической функцией, $\lambda > 0$.
7) Доказать, что $\varphi(t) = (1-|t|)I_{|t|\leq 1}$ является характеристической функцией. Указание: найти преобразование Фурье φ и воспользоваться теоремой об обращении.
8) Из предыдущей задачи вывести **теорему Пойа**: четная функция φ, выпуклая и убывающая на $[0, +\infty)$, $\lim_{t\to\infty} \varphi(t) = 0$, $\varphi(0) = 1$, является характеристической. Выведите отсюда, что функция $\frac{1}{1+|t|}$ является характеристической.
9) Найти пример двух зависимых с.в., для которых произведение характеристических функций равно характеристической функции суммы. Указание: рассмотрите случайный вектор с плотностью
$$f(x, y) = \frac{1}{4}(1 + xy(x^2 - y^2)) \cdot I_{(|x|<1,|y|<1)}.$$
10) (Характеризация гауссовских с.в.) Пусть X, Y — независимые, одинаково распределенные с.в. с нулевым средним и единичной дисперсией. Доказать, что если распределение
$$\frac{1}{\sqrt{2}}(X + Y)$$
совпадает с распределением X и Y, то X и Y являются нормальными с.в.

3. Гауссовские распределения. Центральная предельная теорема.

Гауссовские распределения

Определение 3.1. *Случайный вектор ξ называется гауссовским, если его характеристическая функция имеет вид*
$$\varphi_\xi(t) = \exp\Big(i\langle t, a\rangle - \frac{1}{2}\langle At, t\rangle\Big),$$
где $a \in \mathbb{R}^n$ и A — симметричная невырожденная матрица.

В случае невыпожденной матрицы A можно явно указать плотность ξ.

Теорема 3.2. *Если A невырождена, то ξ имеет плотность*
$$f_\xi(x) = \sqrt{\frac{\det Q}{(2\pi)^d}} \exp\Big(-\frac{\langle Q(x-a), (x-a)\rangle}{2}\Big),$$
где $Q = A^{-1}$.

Доказательство. Переходя к новому ортогональному базису, можно считать, что Q диагональна. Тогда теорема легко следует из соответствующей одномерной теоремы. □

В общем случае надо приблизить матрицу A по норме невырожденными матрицами и воспользоваться соотношением между слабой сходимостью мер и поточечной сходимостью функционалов.

Теорема 3.3. *Линейный образ гауссовского распределения является гауссовским распределением.*

Определение 3.4. *Ковариационной матрицей случайного вектора ξ называется матрица C, где*

$$C_{ij} = \mathbb{E}\Big(\big(\xi_i - \mathbb{E}(\xi_i)\big)\big(\xi_j - \mathbb{E}(\xi_j)\big)\Big).$$

Упражнение 3.5. *Докажите, что C симметрична и неотрицательно определена.*

Теорема 3.6. *Для гауссовского случайного вектора $a = \mathbb{E}\xi$ и*

$$C = Q.$$

Доказательство. Указание: продифференцируйте дважды характеристическую функцию в нуле. □

Следствие 3.7. *У гауссовского случайного вектора компоненты независимы тогда и только тогда, когда они некоррелированы (C диагональна).*

Характеристические функции : теорема Бохнера-Хинчина

Доказательство: Необходимость доказана выше. Докажем достаточность при дополнительном предположении, что φ интегрируема. Положим

$$f(y) = \frac{1}{(2\pi)^n} \int e^{-i(x,y)} \varphi(x)\, dx.$$

Заметим, что f ограничена и непрерывна. Докажем, что f — неотрицательная интегрируемая функция. Из определения неотрицательно определенных функций несложно извлечь, что $\varphi(z)e^{i(z,x)}$ неотрицательно определена для любого z. Положим $p_t(x) = \frac{1}{(2\pi t)^{n/2}} \exp(-\frac{x^2}{2t})$ (напомним, что это вероятностная плотность). Так как p_t интегрируема, то корректно определена свертка p_t и f. Пользуясь теоремой Фубини и явным видом характеристических функций гауссовых распределений, получаем

$$\begin{aligned} p_t * f(x) &= \int_{\mathbb{R}^n} f(y) p_t(x-y)\, dy = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-i(y,z)} \varphi(z) p_t(x-y)\, dy\, dz \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} e^{-i(y,z)} p_t(x-y)\, dy \Big) \varphi(z)\, dz = (2\pi)^{-n} \int_{\mathbb{R}^n} \varphi(z) \exp\Big(-i(x,z) - \frac{t|z|^2}{2}\Big)\, dz. \end{aligned}$$

Так как $\varphi(t)\exp(-i(y,z))$ неотрицательно определена, то из данного равенства и определения неотрицательной определенности вытекает, что $p_t * f$ неотрицательна.

При $t \to \infty$ в силу непрерывности f получаем $f(x) = \lim_{t\to\infty} p_t * f(x) \geq 0$ (см. пример 1.1). Неотрицательность доказана. Докажем интегрируемость. Если в равенстве положить $x = 0$, получим

$$(2\pi t)^{n/2} \int_{\mathbb{R}^n} f(y) p_t(y)\, dy \leq (2\pi t)^{n/2} (2\pi)^{-n} \int_{\mathbb{R}^n} \varphi(z) \exp\Big(-\frac{t|z|^2}{2}\Big)\, dz \leq (2\pi)^{-n} (2\pi t)^{n/2} \int_{\mathbb{R}^n} \varphi(0) \exp\Big(-\frac{t|z|^2}{2}\Big)\, dz.$$

Поледнее выражение равно $\varphi(0)$. Устремляя t к бесконечности, по теореме Фату получаем $\int_{\mathbb{R}^n} f(y)\, dy < \infty$. По теореме обращения характеристическая функция f совпадает с φ.

В общем случае надо рассмотрим функции $\varphi_k(x) = \varphi(x)\exp(-\frac{|x|^2}{2k})$. Нетрудно проверить, что эти функции неотрицательно определены и интегрируемы. Им соответсвуют некоторые вероятностные меры $\{\mu_k\}$. Очевидно, $\lim_{k\to\infty} \varphi_k(x) = \varphi$. В силу замечания 2.17 μ_k сходятся к некоторой вероятностной мере μ, причем φ — характеристическая функция μ.

Центральная предельная теорема

Теорема 3.8. *(Центральная предельная теорема). Пусть $\{\xi_n\}$ — последовательность независимых одинаково распределенных с.в. с с $\mathbb{E}\xi_i^2 < \infty$, $a = \mathbb{E}(\xi_i)$, $\sigma^2 = D\xi_i$. Тогда*

$$\lim_{n\to\infty} P\Big(\frac{\xi_1 + \cdots + \xi_n - na}{\sqrt{n}\sigma} < t\Big) = \Phi(t),$$

где $\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}}\, dx.$

Доказательство. Достаточно доказать, что характеристическая функция $\frac{\xi_1+\dots+\xi_n-na}{\sqrt{n}\sigma}$ сходится поточечно к $e^{-\frac{t^2}{2}}$. Так как $\frac{\xi_1+\dots+\xi_n-na}{\sqrt{n}\sigma} = \frac{1}{\sqrt{n}}\sum_{i=1}^n \eta_i$, $\eta_i = \frac{\xi_i-a}{\sigma}$, η_i — независимые с.в. со средним ноль и единичной дисперсией, то

$$\varphi_{\frac{\xi_1+\dots+\xi_n-na}{\sqrt{n}\sigma}}(t) = \varphi^n_{\eta_i}(t/\sqrt{n}).$$

По свойствам характеристических функций φ_{η_i} дважды дифференцируема, таким образом $\varphi_{\eta_i}(t) = 1 - \frac{t^2}{2} + o(t^2)$ и

$$\varphi_{\frac{\xi_1+\dots+\xi_n-na}{\sqrt{n}\sigma}}(t) = \Big(1 - \frac{t^2}{2n} + o(t^2/n)\Big)^n \to e^{-\frac{t^2}{2}}.$$

Теорема доказана. □

Замечание 3.9. *Важным частным случаем (и ее классической версией) ЦПТ является формула оценки числа успехов S_n в схеме Бернулли (формула Муавра-Лапласа)*

$$P\Big(k_1 \le S_n \le k_2\Big) \approx \Phi\Big(\frac{k_2 - np}{\sqrt{npq}}\Big) - \Phi\Big(\frac{k_1 - np}{\sqrt{npq}}\Big).$$

Пример 3.10. *(Формула Стирлинга) Пусть ξ_i — независимые показательные с.в.с $\lambda = 1$. Несложно показать по индукции, что*

$$f_{\xi_1+\dots+\xi_n} = \frac{x^n}{n!}e^{-x}I_{\{x\ge 0\}}.$$

Тогда

$$f_{\frac{\xi_1+\dots+\xi_n-n}{\sqrt{n}}}(x) = \sqrt{n}\frac{(n+\sqrt{n}t)^n}{n!}e^{-n-\sqrt{n}t} = \sqrt{n}n^n\frac{(1+\frac{t}{\sqrt{n}})^n}{n!}e^{-n-\sqrt{n}t} \sim \sqrt{n}n^n\frac{e^{-n-\frac{t^2}{2}}}{n!}.$$

С другой стороны, по центральной предельной теореме

$$f_{\frac{\xi_1+\dots+\xi_n-n}{\sqrt{n}}}(t) \to \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}.$$

Следовательно,

$$\sqrt{n}n^n\frac{e^{-n-\frac{t^2}{2}}}{n!} \sim \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$$

и мы получаем формулу Стирлинга

$$n! \sim \sqrt{2\pi n}\Big(\frac{n}{e}\Big)^n.$$

Занятие 3

1) Три стрелка поочередно ведут стрельбу по одной и той же мишени. Вероятности попадания в мишень при одном выстреле для каждого стрелка равны $0.4; 0.6; 0.8$. Оценить вероятность того, что число попаданий при 450 выстрелах будет заключаться в пределах от 239 до 390 (≥ 0.9).
2) Вероятность попадания в мишень при одном выстреле равна $0,2$. Найти вероятность того, что в результате 500 выстрелов промахов окажется от 410 до 430 (0.13).
3) Кольцо спрятано в одном из ящиков с номерами $\dots -2, -1, 0, 1, 2 \cdots$. Вам дают подсказку. При подкидывании монеты 100 раз вы узнаете сумму номера ящика с кольцом и числа выпавших гербов. Оно равно 75. 1) Какие 11 ящиков следует проверить? 2) оцените вероятность найти кольцо в этих 11 ящиках (0.36).
4) (Теорема Пуассона. Закон редких явлений). Оценить вероятность того, что из 256 человек ровно двое родились в Новый Год $(0, 122)$.
5) Пусть X, Y — независимые пуассоновские с.в. с параметром n. Найти предел
$$\lim_{n\to\infty} P(X - Y \le \sqrt{2n}x).$$
6) Вывести формулу Стирлинга способом примера 3.10, но используя пуассоновские распределения вместо показательных.

4. Закон больших чисел. Лемма Бореля-Кантелли. Закон нуля и единицы. Сходимость сумм независимых случайных величин.

Основные виды сходимости

Для усвоения материала этого раздела надо владеть следующими понятиями: сходимость почти наверное (почти всюду), сходимость по вероятности (по мере), в среднем (в $L^p(\mu)$, $p \geq 1$), по распределению (слабая сходимость мер) и знать соотношения между ними (см. [4], глава 2(10)):

$$\text{сходимость п.н.} \to \text{сходимость по вероятности} \to \text{сходимость по распределению.}$$

$$\text{сходимость в среднем} \to \text{сходимость по вероятности.}$$

Из сходимости по вероятности не следует сходимости п.н., но верна следующая теорема.

Теорема 4.1. *(Рисс) Из последовательности с.в $\{\xi_n\}$, сходящейся по вероятности, можно выделить почти наверное сходящуюся подпоследовательность.*

Чрезвычайно важными являются следующие теоремы о сходимости под знаком мат. ожидания

Теорема 4.2. *Пусть $\{\xi_n\}$ — последовательность с.в. Тогда*

1) *(Лебег). Если $\xi_n \to \xi$ п.н. и $|\xi_n| < \eta$, $\eta \in L^1(P)$ то*
$$\lim_n \mathbb{E}(\xi_n) = \mathbb{E}(\xi).$$
2) *(Фату). Если $\xi_n \geq \eta$ и $\mathbb{E}(\eta) > -\infty$, то*
$$\underline{\lim}_n \mathbb{E}(\xi_n) \geq \mathbb{E}(\underline{\lim}_n \xi_n).$$

Закон больших чисел

В этом разделе мы кратко обсудим вопросы сходимости почти наверное для сумм независимых случайных величин, составляющие обширный раздел теории вероятностей (предельные теоремы). Классический результат на эту тему — закон больших чисел, открытый еще Я. Бернулли в 1713 г. (для схемы Бернулли).

Теорема 4.3. *(Закон больших чисел). Пусть $\{\xi_n\}$ — последовательность независимых, одинаково распределенных с.в. с $\mathbb{E}\xi_i^2 < \infty$, $a = \mathbb{E}(\xi_i)$, $\sigma^2 = D\xi_i$. Тогда*
$$\lim_{n\to\infty} P\Big(\Big|\frac{\xi_1 + \cdots + \xi_n}{n} - a\Big| > \varepsilon\Big) = 0$$
для любого $\varepsilon > 0$.

Доказательство. В силу неравенства Чебышева достаточно оценить $\mathbb{E}\Big|\frac{\xi_1+\cdots+\xi_n}{n} - a\Big|^2$. Пользуясь независимостью и одинаковой распределенностью, получаем
$$\mathbb{E}\Big|\frac{\xi_1 + \cdots + \xi_n}{n} - a\Big|^2 = \frac{1}{n^2}\mathbb{E}\Big|\sum_{i=1}^n (\xi_i - a)\Big|^2 = \frac{1}{n}\mathbb{E}\Big|(\xi_1 - a)\Big|^2 = \frac{\sigma^2}{n} \to 0$$
. □

Отметим, что на самом деле верно значительное усиление этого результата.

Теорема 4.4. *(Усиленный закон больших чисел (А.Н. Колмогоров)). Пусть $\{\xi_n\}$ — последовательность независимых, одинаково распределенных с.в. с $\mathbb{E}|\xi_n| < \infty$. Тогда имеет место сходимость почти наверное*
$$\frac{S_n}{n} = \frac{\xi_1 + \cdots + \xi_n}{n} \to a \text{ п.н.}.$$

Мы докажем ниже усиленный з.б.ч. при более сильном предположении конечности четвертого момента.

Важным техническим средством для доказательства теорем о суммах независимых с.в. является лемма Бореля-Кантелли.

Лемма 4.5. *(Борель-Кантелли). Пусть $\{A_n\}$ — последовательность событий. Событие A состоит в том, что события A_n случаются бесконечно часто*

$$A = \bigcap_{n=1}^{\infty} \bigcup_{m \ge n} A_m.$$

1) *Если $\sum_{m=1}^{\infty} P(A_m) < \infty$, то $P(A) = 0$.*
2) *Если $\sum_{m=1}^{\infty} P(A_m) = \infty$ и события $\{A_m\}$ независимы, то $P(A) = 1$.*

Доказательство. a) $P(A) \le \lim_n P(\bigcup_{m \ge n} A_m) \le \lim_n \sum_{m \ge n} P(A_m) = 0$.

b) В силу независимости $P(\cap_{m=n}^{\infty} A_m^c) = \prod_{m=n}^{\infty}(1 - P(A_m))$. Заметим, что

$$\ln P(\cap_{m=n}^{\infty} A_m^c) = \sum_{m=n}^{\infty} \ln(1 - P(A_m)) \le -\sum_{m=n}^{\infty} P(A_m) = -\infty.$$

Следовательно, $P(\cap_{m=n}^{\infty} A_m^c) = 0$. Тогда $0 = P(\cup_{n=1}^{\infty} \cap_{m=n}^{\infty} A_m^c) = P(A^c) = 1 - P(A)$. □

Теорема 4.6. *Пусть $\{\xi_i\}$ — последовательность независимых с.в. с конечным четвертым моментом. Предположим, что*

$$\mathbb{E}(\xi_n - \mathbb{E}\xi_n)^4 < C.$$

Тогда

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

Идея доказательства: Без ограничения общности $\mathbb{E}\xi_n = 0$. По лемме Бореля-Кантелли достаточно доказать, что $\sum_n P\left(\left|\frac{S_n}{n}\right| \ge \varepsilon\right) < \infty$. По неравенству Чебышева достаточно доказать сходимость $\sum_n \mathbb{E}\left(\frac{S_n}{n}\right)^4$. Далее надо оценить $\mathbb{E}S_n^4$ через четвертые моменты ξ_i, пользуясь независимостью и неравенством Коши-Буняковского.

Закон нуля и единицы

Определение 4.7. *Пусть $\{\xi_n\}$ — последовательность с.в.. Обозначим через $\mathcal{F}^{(n)}$ σ-алгебру, порожденную с.в. $\{\xi_n, \xi_{n+1}, \cdots\}$. Положим*

$$\mathcal{F}^{\infty} = \cap_{n=1}^{\infty} \mathcal{F}^{(n)}.$$

Система множеств $\mathcal{F}^{\infty}$ является пересечением σ-алгебр, следовательно, сама является σ-алгеброй.

Пример 4.8. *Следующие события принадлежат $\mathcal{F}^{\infty}$.*

$$\Big\{\overline{\lim}_n \xi_n \le c\Big\}, \quad \Big\{\frac{S_n}{n} \text{ сходится}\Big\}.$$

Теорема 4.9. *(Закон нуля и единицы. А.Н. Колмогоров.) Пусть $\{\xi_i\}$ — независимые с.в. с $\mathbb{E}\xi_i = 0$. Тогда для любого события $A \in \mathcal{F}^{\infty}$ имеем $P(A) = 0$ либо $P(A) = 1$.*

Доказательство. Так как A принадлежит σ-алгебре $\mathcal{F}^{(n)}$, порожденной $\{\xi_{n+1}, \xi_{n+2}, \cdots\}$, то A не зависит от σ-алгебры $\mathcal{F}_n$, порожденной $\{\xi_1, \cdots, \xi_n\}$. Но объединение $\mathcal{F}_n$ порождает всю σ-алгебру цилиндрических множеств. Значит, A не зависит от самого себя. Следовательно $P(A) = P(AA) = P^2(A)$. Отсюда вытекает утверждение теоремы. □

Отсюда следует, что события $\{\sum_{n=1}^{\infty} \xi_n$ сходится$\}$, $\{\lim_n \xi_n$ существует$\}$ имеют вероятность ноль либо единица для независимых с.в. $\{\xi_n\}$.

Случайные ряды

Завершим этот раздел теоремой из теории случайных рядов. Как мы увидим в дальнейшем, здесь применяется техника так называемых мартингалов, важного класса случайных процессов.

Теорема 4.10. *Пусть $\{\xi_n\}$ — последовательность независимых с.в. Если $\sum_{n=1}^{\infty} \mathbb{E}(\xi_n^2) < \infty$, то $\sum_{n=1}^{\infty} \xi_i$ сходится почти всюду.*

Доказательство. Положим $S_n = \sum_{i=1}^n \xi_i$. Докажем сперва следующее неравенство.

$$P(\max_{1\le i\le n} |S_i| \ge \varepsilon) \le \frac{\mathbb{E}S_n^2}{\varepsilon^2}.$$

Обозначим $A = \{\max_{1\le i\le n} |S_i| \ge \varepsilon\}$, $A_k = \{1 \le i \le k, |S_i| < \varepsilon; |S_k| \ge \varepsilon\}$. Заметим, что A есть дизъюнктное объединение A_k и

$$\mathbb{E}S_n^2 \ge \mathbb{E}S_n^2 I_A = \sum_{k=1}^n \mathbb{E}S_n^2 I_{A_k}.$$

Далее, в силу независимости

$$\mathbb{E}S_n^2 I_{A_k} = \mathbb{E}(S_k + \sum_{i=k+1}^n \xi_i)^2 I_{A_k} = \mathbb{E}(S_k^2) I_{A_k} + 2\mathbb{E}S_k(\sum_{i=k+1}^n \xi_i) I_{A_k} + \mathbb{E}(\sum_{i=k+1}^n \xi_i)^2 I_{A_k} \ge \mathbb{E}(S_k^2) I_{A_k}.$$

Следовательно,

$$\mathbb{E}S_n^2 \ge \sum_{k=1}^n \mathbb{E}(S_k^2) I_{A_k} \ge \varepsilon^2 \sum_{k=1}^n P(A_k) = \varepsilon^2 P(A).$$

Отсюда следует сходимость п.н. Действительно, достаточно доказать, что $\{S_n\}$ почти всегда образует фундаментальную последовательность.

$$P(\omega : \{S_n(\omega)\} \text{ не фундаментальна}) = P(\bigcup_m \bigcap_n A_{m,n}),$$

где $A_{m,n} = \{\omega : \sup_{k\ge 1} |S_{n+k} - S_n|(\omega) > \frac{1}{m}\}$. Поэтому достаточно доказать, что $P(\bigcap_n A_{m,n}) = 0$. Но по доказанным неравенствам

$$P(\bigcap_n A_{m,n}) \le \lim_{n\to\infty} P(A_{m,n}) \le \lim_{n\to\infty} m^2 \sum_{k=n}^\infty \mathbb{E}\xi_k^2 = 0.$$

Теорема доказана. □

Занятие 4

1) В поселке 2500 жителей. Каждый из них независимо от других примерно 6 раз в месяц ездит на поезде в город. Поезд едет раз в сутки. Какой наименьшей вместительностью должен обладать поезд, чтобы он переполнялся в среднем не чаще одного раза в сто дней?
2) Плотность вероятности двумерного случайного вектора (X, Y) имеет вид
$$f(x, y) = ae^{-4x^2-6xy-9y^2}.$$
Найти a, корреляцию X и Y, и условные математические ожидания
$$\mathbb{E}(g(X, Y)|X),\ \mathbb{E}(g(X, Y)|Y).$$
3)* Доказать, что условные меры гауссовских распределений относительно координатных функций являются гауссовскими.
4) Закончить доказательство теоремы 4.6.
5) Построить пример последовательности с.в., сходящейся в $L^1(P)$, но не сходящейся почти всюду.
6) (Равномерная интегрируемость). Докажите, что если $\xi_n \to \xi$ п.н. и $\sup_n \mathbb{E}(|\xi_n|^p) < \infty$ для $p > 1$, то $\mathbb{E}|\xi_n - \xi| \to 0$.
7) Доказать, что для независимых одинаково распределенных случайных величин ξ_n, $\mathbb{E}(\xi_n) = 0$, $\mathbb{E}(\xi_n^2) = 1$ выполнено $\overline{\lim}_n \frac{S_n}{\sqrt{n}} = +\infty$ п. н..
8) Пусть $\{\xi_i\}$ — последовательность независимых одинаково распределенных с.в., $\mathbb{E}\xi_i^2 < \infty$ Доказать, что если $n^{-1/2}\sum_{i=1}^n \xi_i$ сходится п.в. к с.в. ξ, то $\xi_i = 0$ п.в.
9) Используя метод характеристических функций, докажите, что закон больших чисел верен в предположении конечности первого момента.
10) Пусть $\{\xi_i\}$ — последовательность независимых одинаково распределенных случайных величин с $\mathbb{E}\xi_i = a$, $D\xi_i = \sigma^2$. Докажите, что $\frac{\sum_{i=1}^d \xi_i}{\sum_{i=1}^d \xi_i^2}$ сходится по вероятности и найти предел.
11*) (Сингулярность бесконечномерных мер) Пусть $\gamma_\sigma = \prod_{i=1}^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x_i^2}{2\sigma^2}} dx_i$ — счетное произведение гауссовских мер на $\mathbb{R}^\infty$. Доказать, что γ_{σ_1} и γ_{σ_2} сингулярны для $\sigma_1 \ne \sigma_2$, т.е. $\gamma_{\sigma_1}(S) = 1, \gamma_{\sigma_2}(S) = 0$ для некоторого множества S.

12*) (Лемма Пуанкаре) Пусть μ_n — равномерное распределение на сфере радиуса $\sqrt{n}$ в пространстве $\mathbb{R}^n$. Зафиксируем m-мерное подпространство $\mathbb{R}^m = \{x_i = 0,\ i > m\}$ и рассмотрим последовательность распределений ν_n, состоящую из ортогональных проекций μ_n на $\mathbb{R}^m$. Доказать, что $\{\nu_n\}$ слабо сходится к стандартному нормальному распределению на $\mathbb{R}^m$.

13) Доказать, что если последовательность независимых с.в. $\{\xi_i\}$ сходится п.в. к с.в. ξ, то ξ постоянна п.в.

14) Пусть $\{X_n\}$ — независимые с.в. с
$$P(X_n = n) = P(X_n = -n) = \frac{1}{2n\ln n},\quad P(X_n = 0) = 1 - \frac{1}{n\ln n}.$$
Докажите, что последовательность S_n/n сходится по вероятности, но не п.н.

15) (Теорема Вейерштрасса о приближении многочленами). Пусть $f : [0,1] \to \mathbb{R}$ — непрерывная функция. Доказать, что
$$\lim_{n\to\infty} \sup_{0\le x\le 1} \Bigl| f(x) - \sum_{k=0}^{n} f(k/n) C_n^k x^k (1-x)^{n-k} \Bigr| = 0.$$

16) Пусть $\{\xi_i\}$ — последовательность с.в. Доказать, что если для некоторого $r > 0$ выполнено $\sum_{i=1}^{\infty} P(|\xi_i| > r) < \infty$, то $P(\sup_i |\xi_i| < \infty) = 1$. Доказать, что обратное верно для независимых с.в.

5. Безгранично делимые распределения. Формула Леви-Хинчина. Устойчивые распределения.

Классы безгранично делимых и устойчивых распределений являются естественными классами распределений, который могут выступать в роли предельных для сумм независимых с.в. Типичное устойчивое распределение не обладает элементарной функцией плотности. Кроме этого, это распределение с "тяжелым хвостом" — у него нет моментов порядка выше второго. Несмотря на это, в современной теории вероятностей это весьма популярный объект исследования. В финансовой математике устойчивые распределения составляют конкуренцию гауссовским.

Определение 5.1. *Безгранично делимым распределением ξ на прямой называется такое распределение, что ее характеристическая функция φ удовлетворяет условию: для любого $n \in \mathbb{N}$ существует такая характеристическая функция φ_n, что $\varphi = \bigl(\varphi_n\bigr)^n$.*

Замечание 5.2. *Заметим, что φ_n определена однозначно. Действительно, пусть существует пара таких характеристических функций $\varphi_n^{(1)}$ и $\varphi_n^{(2)}$, что $(\varphi_n^{(1)})^n = (\varphi_n^{(2)})^n$. Тогда, если $\varphi(x_0) \ne 0$, то, в силу непрерывности обеих функций, они отличаются в окрестности x_0 постоянным множителем вида $t = \sqrt[n]{1}$. Но в силу теоремы Бохнера $t = 1$ (в противном случае нарушается условие неотрицательной определенности для одной из функций). В дальнейшем будем обозначать $\varphi_n = \varphi^{1/n}$.*

Замечание 5.3. *Указанное свойство можно было бы сформулировать иначе: найдутся такие независимые, одинаково распределенные с.в. ξ_i, что*
$$f_{\xi_1+\dots+\xi_n} = f_\xi.$$
Однако, существует некоторая разница в определениях. Не всегда на заданном вероятностном пространстве можно найти такие случайные величины. Мы будем предполагать в дальнейшем, что наше вероятностное пространство достаточно богато и такие с.в. существуют.

Приведем примеры безгранично делимых с.в. Во всех этих примерах соответствующая характеристическая функция $\varphi_n = \varphi_\xi^{1/n}$ имеет такую же форму, но отличается параметрами.

Пример 5.4. 1) *Вырожденное распределение*
$$\varphi(t) = e^{ita},\ \varphi_n(t) = e^{ita/n}$$

2) *Распределение Пуассона*
$$\varphi(t) = \exp\Bigl(\lambda(e^{it}-1)\Bigr),\ \varphi_n(t) = \exp\Bigl(\frac{\lambda}{n}(e^{it}-1)\Bigr)$$

3) *Нормальное распределение*
$$\varphi(t) = e^{ita - \frac{\sigma^2 t^2}{2}},\ \varphi_n(t) = e^{it\frac{a}{n} - \frac{\sigma^2 t^2}{2n}}$$

4) *Распределение Коши*

$$\varphi(t) = e^{ita-\theta|t|},\ \varphi_n(t) = e^{it\frac{a}{n}-\frac{\theta}{n}|t|}$$

5) *Гамма-распределение*

$$\varphi(t) = [1-(it/\theta)]^{-\lambda},\ \varphi_n(t) = [1-(it/\theta)]^{-\lambda/n}.$$

Лемма 5.5. *Характеристическая функция безгранично делимой с.в. не имеет нулей на действительной оси.*

Доказательство. Пусть φ — безгранично делимая с.в. Тогда $|\varphi|^{2/n} = \varphi^{1/n}\cdot\overline{\varphi}^{1/n}$ — характеристическая функция. Перейдем к пределу $n\to\infty$. Функция $g(t) = \lim_n |\varphi|^{2/n}$ принимает значения 0 или 1. Так как $\varphi(0) = 1$, то $g = 1$ в некоторой окрестности нуля. В силу замечания 2.17 меры, соответствующие характеристическим функциям $\varphi^{1/n}$, сходятся слабо и g — характеристическая функция предельной меры. Следовательно, g — непрерывная функция. В силу того, что $g(0) = 1$, то g тождественно равна 1 и f не может принимать нулевые значения. □

Теорема 5.6. *(Леви-Хинчин) Пусть ξ — безгранично делимая случайная величина. Тогда существует такая конечная неотрицательная мера μ и такое число b, что характеристическая функция ξ имеет вид*

$$\varphi(t) = \exp\Big(ibt + \int\big(e^{itx} - 1 - it\frac{x}{1+x^2}\big)\frac{1+x^2}{x^2}d\mu\Big).$$

Доказательство. В силу непрерывности $\varphi(t) = \mathbb{E}e^{it\xi} = \int e^{itx}\,dF(x)$ и того факта, что φ не обращается в нуль, можно рассмотреть функцию на прямой $g(t) = \log\varphi(t)$. Функция g непрерывна и $g(0) = 0$. Заметим,что $\varphi^{1/n}(t) = e^{\frac{1}{n}g(t)}$. Действительно, функция $\xi = e^{\frac{1}{n}g(t)}$ удовлетворяет условиям $\xi^n = \varphi$ и $\xi(0) = 1$. Следовательно, $\xi = \varphi^{1/n}$ в силу непрерывности. Имеем

$$g(t) = \lim_{n\to\infty} n(\varphi^{1/n}(t) - 1).$$

В силу того, что $\varphi^{1/n}$ является безгранично делимой, получаем следующее представление

$$n(\varphi^{1/n} - 1) = n\int_{-\infty}^{\infty}(e^{itx} - 1)dF_n,$$

где F_n — некоторая функция распределения. Проинтегрируем это выражение по t. Получим

$$\lim_n n\int_{\mathbb{R}}\big(1 - \frac{\sin xh}{xh}\big)dF_n = -\frac{1}{2h}\int_{-h}^{h} g(t)\ dt.$$

В силу того, что g непрерывна и $g(0) = 0$, получаем, что правая часть может быть сделана сколь угодно малой при малых h. Заметим также, что $1 - \frac{\sin xh}{xh} \ge \frac{1}{2}$ при $|xh| \ge 2$. Следовательно, для любого ε найдется такое h, что

$$n\int_{|x|\ge 2/h} dF_n \le \varepsilon.$$

Далее, заметим, что существует константа c т.ч.

$$1 - \frac{\sin x}{x} \ge c\frac{x^2}{1+x^2}.$$

Тогда, положив $h = 1$, получаем

$$n\int_{\mathbb{R}}\frac{x^2}{1+x^2}dF_n \le C.$$

для всех n и некоторой константы C. Из доказанных неравенств следует, что семейство мер

$$\mu_n(dx) = \frac{nx^2}{1+x^2}F_n(dx)$$

плотно. Перейдя к подпоследовательности, получаем, что

$$\mu_{n_k} \to \mu,$$

слабо, где μ — некоторая конечная мера. Положим:

$$f(t,x) = \big(e^{itx} - 1 - it\frac{x}{1+x^2}\big)\frac{1+x^2}{x^2},\ x\ne 0,\quad f(x,0) = -\frac{t^2}{2}.$$

Заметим, что $f(x,t)$ непрерывна и ограничена. Поэтому можно перейти к пределу

$$g(t) = n\int_{-\infty}^{\infty}(e^{itx}-1)dF_n = \lim_{n_k\to\infty}\int f(t,x)\ d\mu_{n_k} + \lim_{n_k} itn_k\int \frac{x}{1+x^2}\ F_{n_k}(dx).$$

Получаем

$$g(t) = \int f(t,x)\ d\mu + itb.$$

Предел

$$b = \lim_{n_k} n_k \int \frac{x}{1+x^2}\ F_{n_k}(dx)$$

существует в силу существования остальных пределов. Теорема доказана. □

Теорема 5.7. *Константа b и мера μ в представлении Леви-Хинчина определена единственным образом.*

Другое (эквивалентное) представление характеристической функции носит название представления Леви

$$\varphi(t) = \exp\Big(ibt - \frac{t^2\sigma^2}{2} + \int\big(e^{itx}-1-it\frac{x}{1+x^2}\big)\frac{1+x^2}{x^2}d\Lambda\Big), \tag{1}$$

где Λ — конечная неотрицательная мера с $\Lambda(\{0\})=0$. При $\Lambda=0$ получаем гауссовское распределение.

Теорема 5.8. *Случайная величина ξ может быть пределом по распределению сумм вида $T_n = \sum_{i=1}^n \xi_{n,i}$, где $\{\xi_{n,i}\}$ — независимые, одинаково распределенные с.в. тогда и только тогда, когда она безгранично делима.*

Идея доказательства: Если с.в. ξ безгранично делима, то, очевидно, ее можно представить как предел по распределению таких сумм. Для доказательства обратного утверждения рассмотрим приближающие суммы вида $T_{nk} = \sum_{i=1}^{nk}\xi_{nk,i},\ k,n\in\mathbb{N}$. Положим $T_{nk} = \eta_1(k)+\cdots+\eta_n(k)$, где $\eta_i(k) = \sum_{j=1}^k \xi_{nk,k(i-1)+j}$. Так как $T_{nk}\to\xi$ при $k\to\infty$ по распределению, то $\lim_{N\to\infty}\sup_k P(T_{nk} > N) = 0$. Далее, в силу независимости

$$P(T_{nk} > N) \geq P(\min_i \eta_i(k) > N/n) = P^n(\eta_1(k) > N/n).$$

Таким образом $\lim_N \sup_k P(\eta_1(k) > N/n)$ =0. Аналогичное соотношение выполнено для $P(\eta_i < -N/n)$. Используя теорему Прохорова, несложно показать, что подпоследовательность $\eta_i(k)$ сходятся по распределению к с.в. η . Далее, надо доказать, что распределение ξ совпадает с распределением $\eta_1+\cdots+\eta_n$, где $\{\eta_j\}$ — n независимых копий η.

Устойчивые распределения

Важным классом безгранично делимых распределений являются так называемые устойчивые распределения. Распределение называются устойчивым, если для любого n его характеристическая функция φ имеет представление

$$\varphi^n(t) = e^{ib_nt}\varphi(a_nt),$$

где a_n, c_n — некоторые константы.

Любое устойчивое распределение является безгранично делимым, но не наоборот. Для устойчивых распределений существует представление, аналогичное формуле Леви-Хинчина.

$$\log\varphi = it\beta - c|t|^\alpha\big(1+i\theta\frac{t}{|t|}G(t,\alpha)\big),$$

где $0<\alpha\leq 2, c\geq 0, |\theta|\leq 1$ и $G(t,\alpha) = \tan\frac{\pi\alpha}{2}$ при $\alpha\neq 1$ ($G(t,\alpha) = \frac{2}{\pi}\ln|t|, \alpha\neq 1$). Классическим примером устойчивых распределений являются симметрические устойчивые распределения (распределения Леви) с характеристической функцией

$$\varphi(t) = e^{-c|t|^\alpha},\ 0<\alpha\leq 2.$$

При $\alpha=2$ получаем гауссовское распределение, $\alpha=1$ — распределение Коши. Известен явный вид плотности распределения при $\alpha=1/2$ (см. упражнения).

Теорема 5.9. *Устойчивые распределения и только они являются пределами по распределению сумм вида*

$$\frac{\sum_{i=1}^{n} X_i}{C_n} - D_n,$$

где C_n, D_n — числа, а $\{X_i\}$ — последовательность независимых, одинаково распределенных с.в.

Занятие 5

1) Сколько опытов с бросанием монеты нужно произвести, чтобы с вероятностью $0,98$ можно было ожидать отклонение частоты выпадения герба от теоретической вероятности $0,5$ по абсолютной величине меньше, чем $0,91$ процент? ($n > 7657$).
2) Для n-мерного стандартного нормального вектора ξ найти распределение с.в. $|\xi|$.
3) Доказать, что слабый предел безгранично делимых распределений безгранично делим.
3) Доказать, что равномерное распределение не является распределением суммы двух независимых с.в.
4) Доказать, что произведение двух безгранично делимых характеристических функций является безгранично делимой х.ф.
5) Показать, что биномиальное распределение не является безгранично делимым.
6*) (Устойчивое распределение с элементарной функцией плотности). Доказать, что распределение с плотностью
$$f(x) = \frac{1}{\sqrt{2\pi x^3}} e^{-\frac{1}{2x}} I_{\{x>0\}}.$$
является устойчивым.
7) Найти явно меру Леви для устойчивых симметричных распределений и распределения Пуассона.
8) Доказать, что если φ — характеристическая функция, то $\exp(p(\varphi - 1))$ — безгранично делимая характеристическая функция.
9*) Доказать, что если функция φ имеет вид (1), то она является характеристической функцией безгранично делимого распределения. Докажите, что функции вида $e^{-|t|^\alpha}$ являются характеристическими для $0 < \alpha \leq 2$. Указание: приблизьте интеграл дискретными суммами и представьте функцию как предел произведений характеристических функций пуассоновских распределений.

Список литературы

[1] Богачев В.И., Основы теории меры, тт. 1-2, Москва-Ижевск, НИЦ Регулярная и хаотическая динамика, 2006.
[2] Лукач Е. Характеристические функции. — М. Наука, 1979.
[3] Прохоров А.В., Ушаков В.Г., Ушаков Н.Г., Задачи по теории вероятностей. М. Наука, 1986.
[4] Ширяев А.Н., Вероятность. тт. 1-2. М. МЦНМО, 2007.
[5] Ширяев А.Н., Задачи по теории вероятностей. М. МЦНМО, 2006.
[6] Grimmet G., Stirzaker D., One thousand exercises in probability, Oxford University Press, 2001.
[7] Stirzaker D., Elementary probability, Cambrige University Press, 2003.