Теория вероятностей и математическая статистика

12 марта 2016 г.

С точностью до нормировочного множителя характеристическая функция $h(x) \stackrel{\mathrm{def}}{=} \mathbb{E}\, e^{i\xi x}$ д. с. в. ξ равна преобразованию Фурье ее вероятностной меры.

Характеристическая функция обладает свойствами:

1) $h(0) = 1$, $|h(x)| \le 1$,
2) $h \in C(\mathbb{R})$, то есть $h(x)$ – непрерывная функция. [1],
3) $\sum_{k,j} h(x_k - x_j)\bar{c}_k c_j \ge 0$ для любых $\{x_k\} \in \mathbb{R}$, $\{c_k\} \in \mathbb{C}$, т. е. матрица $\{h(x_k - x_j)\}$ положительно определена.

[1] Более точно, $h(x)$ – равномерно непрерывная функция, так как в силу теоремы Лебега о предельном переходе под знаком интеграла существует предел $|h(x+\delta) - h(x)| \le \mathbb{E}|e^{i\xi\delta} - 1| \to 0$ при $\delta \to 0$. Поскольку $\mathbb{E}|e^{i\xi\delta} - 1|$ не зависит от x, сходимость равномерна.

Очевидно, что произведение двух характеристических функций обладает свойствами 1–2. Произведение характеристических функций вероятностных мер обладает свойством 3, так как распределение суммы независимых случайных величин является сверткой их распределений, а характеристическая функция свертки (т.е. преобразование Фурье свертки) является произведением характеристических функций сомножителей (т.е. их преобразований Фурье). Таким образом, множество характеристических функций вероятностных мер является алгеброй относительно поточечного умножения.

Действительно, для независимых действительных случайных величин ξ, χ выполнены следующие равенства:

$$P_{\xi+\chi}(B) = \int_{\mathbb{R}} P_\xi(B-x)\,P_\chi(dx) = \int_{\mathbb{R}} P_\xi(dx)\,P_\chi(B-x), \tag{1}$$

$$h_{\xi+\chi}(u) = \int\int e^{iu(b-x)}P_\xi(db-x)\,e^{iux}\,P_\chi(dx) =$$
$$= \int_{\mathbb{R}} e^{iub}P_\xi(db)\int_{\mathbb{R}} e^{iux}\,P_\chi(dx) = h_\xi(u)\,h_\chi(u).$$

Поскольку левая часть свертки (1) является вероятностной мерой, ее характеристическая функция положительно определена; с другой стороны, эта характеристическая функция равна произведению $h_{\xi+\chi}(u) = h_\xi(u)\,h_\chi(u)$.

Лемма Шура

Менее очевидно следующее утверждение: *любая функция со свойствами 1–3 определяет вероятностную меру*. Алгебраическую суть дела иллюстрирует *лемма Шура*.

Лемма 1

Пусть $A = \{a_{k,j}\}$ *и* $B = \{b_{k,j}\}$ *– две положительно определенные матрицы. Тогда матрица* $C = \{a_{k,j}b_{k,j}\}$ *также положительно определена*[a].

[a] Утверждение леммы можно усилить: если сомножители A и B – строго положительны, то матрица C также строго положительна.

Доказательство. Неотрицательно определенную матрицу A можно представить в виде суммы проекторов π_m на ее собственные векторы z_m:

$$A = \sum_m \lambda_m \pi_m = \sum_m \lambda_m z_m \bar{z}_m^T, \quad \lambda_k \geq 0, \; z_m \in \mathbb{C},$$

где $\lambda_m \geq 0$ –собственные значения матрицы A, z_m – ортонормированные собственные векторы (векторы-столбцы), $\pi_m = z_m \bar{z}_m^T = |z_m\rangle\langle z_m|$ – эрмитова матрица ранга 1, равная

$$\pi_m = \begin{pmatrix} z_{m,1}\bar{z}_{m,1} & z_{m,1}\bar{z}_{m,2} & \cdots & z_{m,1}\bar{z}_{m,n} \\ z_{m,2}\bar{z}_{m,2} & z_{m,2}\bar{z}_{m,2} & \cdots & z_{m,2}\bar{z}_{m,n} \\ \cdots & \cdots & \cdots & \cdots \\ z_{m,n}\bar{z}_{m,n} & z_{m,n}\bar{z}_{m,2} & \cdots & z_{m,n}\bar{z}_{m,n} \end{pmatrix}.$$

Матрицы π_m обладают свойствами проекторов: $\pi_m = \pi_m^T = \pi_m^2$.

Матрица $C = \{a_{k,j} b_{k,j}\}$ не является композицией матриц A и B и называется *произведением Адамара* матриц A и B:

$$C = A \circ B, \quad c_{k,j} = a_{k,j} b_{k,j}.$$

Необходимое и достаточное условие положительной определенности матрицы C имеет вид $(x, Cx) > 0$ для любого ненулевого вектора $x \in \mathbb{C}$. Поскольку $a_{k,j} = \sum_m \lambda_m \overline{z}_{m,k}\, z_{m,j}$, то

$$(x, Cx) = \sum_{k,j} a_{k,j} b_{k,j} \overline{x}_k\, x_j = \sum_m \lambda_m \sum_{k,j} b_{k,j} \overline{(z_{m,k} x_k)} (z_{m,j} x_j) > 0,$$

так как $\lambda_m > 0$, и из положительной определенности B следует, что $\sum_{k,j} b_{k,j} \overline{(z_{m,k} x_k)} (z_{m,j} x_j) > 0$ при каждом m . □

Из леммы Шура и определения $h(u)$ вытекает

Теорема 1

Функции, обладающие свойствами 1) – 3) образуют алгебру относительно поточечного умножения.

Доказательство. Свойства 1) и 2) для произведения очевидны, а свойство 3) наследуется в силу леммы Шура: для любых $\{x_k\} \in \mathbb{R}$, $\{c_k\} \in \mathbb{C}$

$$\sum_{k,j} h_1(x_k - x_j) h_2(x_k - x_j) \overline{c}_k c_j \geq 0$$

так как поэлементное произведение положительно определенных матриц $M_{kj}^{(n)} = h_n(x_k - x_j)$ положительно определено.

Упражнение 1

Показать, что если функция $h(z)$ *положительно определена в том смысле, что* $\sum_{k,j} h(z_k - z_j)c_k\overline{c}_j \geq 0$, *то* $h(x-y)\,h(y-x) \leq h(0)^2$ *и* $|h(x)| \leq h(0)$ *для любых* x, y. *В частности, из* $h(0) = 1$ *следует, что* $|h(x)| \leq 1$. Файл e3-1

Упражнение 2

Доказать, что если $h(u)$ *– характеристическая функция, то* $h_1(u) = e^{c\,(h(u)-1)}$, $c > 0$, $h_2(u) = e^{\frac{1}{u}\int_0^u h(z)\,dz - 1}$ *также являются характеристическими функциями.* Файл e3-2

Лемма 2

Функции вида $h(u) = \max\{0, 1 - |u|\}$ *и*

$$h_{a,p}(u) = \sum p_n h(u/a_n), \quad p_n,\ a_n > 0,\ \sum p_n = 1$$

– характеристические функции вероятностных мер.

Доказательство. Достаточно доказать утверждение леммы для функции $h(u) = \max\{0, 1 - |u|\}$. В этом случае

$$p(x) = \frac{1}{2\pi}\int_{-1}^{1} e^{iux}(1 - |u|)\, du = \frac{1}{\pi}\mathrm{Re}\,(1 + i\partial_x)\int_0^1 e^{iux}\, du =$$

$$= \frac{1}{\pi}\mathrm{Re}\,(1 + i\partial_x)\frac{e^{ix} - 1}{ix} = \frac{1 - \cos x}{\pi x^2} \geq 0.$$

Поскольку $h(0) = 1 = \int_{\mathbb{R}} p\, dx$, то из положительности $p(x)$ следует, что p – плотность вероятности.

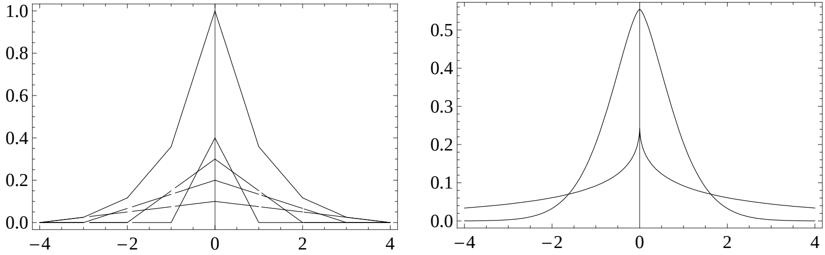

Рис. 1: Слева жирной линией изображен график функции $h_\sigma(u) = \sum_{n=1}^{4} p_n \max\{0, 1 - |u|/a_n\}$, где $a = \{4, 3, 2, 1\}$, $p = \{0.1, 0.2, 0.3, 0.4\}$, а тонкими линиями – графики слагаемых. Справа – выпуклая вниз слева и справа от точки $u = 0$ характеристическая функция $h_{0.5}(u) = e^{-\sqrt{|u|}}$ (негладкий максимум) и $h_{1.5}(u) = e^{-\sqrt{|u|^3}}$ (гладкий максимум)

Следующее утверждение является обобщением леммы .

Теорема 2

Непрерывная, четная, выпуклая вниз слева и справа от точки $u = 0$ *функция* $h_{\mathrm{conv}}(u)$ *является характеристической функцией вероятностной меры*[a], *если* $h_{\mathrm{conv}}(u) \geq 0$, $h_{\mathrm{conv}}(0) = 1$, $h_{\mathrm{conv}}(u) \to 0$ *при* $u \to \infty$*.*

---

[a] Для функций выпуклых вниз используется термин *convex*, а для выпуклых вверх – *concave* (англ.)

Доказательство. Любую ограниченную непрерывную выпуклую вниз симметричную функцию $h_{\mathrm{conv}}(u)$ слева и справа от точки $u = 0$ можно равномерно аппроксимировать функциями вида $h_{a,p}(u)$.

В силу доказываемой в следующей лекции формулы обращения

$$P(a,b) = \lim_{N\to\infty} \frac{1}{2\pi} \int_{-N}^{N} \frac{e^{-iua} - e^{-iub}}{iu} h(u)\, du, \qquad (2)$$

из поточечной сходимости характеристических функций следует слабая сходимость вероятностных мер. Таким образом обосновывается достаточное условие Пойа. $\square$

Условиям этой теоремы удовлетворяют, например, функции

$$h_\alpha(u) = e^{-|cu|^\alpha}, \ \alpha \in (0,1], \ c > 0.$$

Упражнение 3

Показать, что функция $h_q(u) = (1 + q|u|^2)^{-1}$, $q > 0$, *положительно определена. (Указание. Используя теорию вычетов, вычислить фурье-прообраз* $p_q(x)$.*)* Файл e3-3

Отсюда следует положительная определенность функций $e^{-q|u|^2} = \lim_{N\to\infty}(1 + qN^{-1}|u|^2)^{-N}$ и $h(q)^n$, $n \geq 1$, (в силу леммы Шура) и $e^{c\, h_q(u)} = \sum_n \frac{(c\, h_q(u))^n}{n!}$ для любых $c > 0$.

Упражнение 4

Показать, что для любого $a \in (0, 1)$ *имеет место тождество*

$$|u|^{2a} = C_a \int_0^\infty \left(1 - \frac{1}{1 + q|u|^2}\right) \frac{dq}{q^{a+1}}, \quad \frac{1}{C_a} = \int_0^\infty (1 + q)^{-1} \frac{dq}{q^a}.$$

Файл e3-4

Теперь заметим, что

$$h_a(u) = e^{C_a \int_0^\infty ((1+q|u|^2)^{-1}-1)\frac{dq}{q^{a+1}}} = e^{-C_a \int_0^\infty H(q|u|^2)\frac{dq}{q^{a+1}}}$$
$$= e^{-|u|^{2a} C_a \int_0^\infty H(q)\frac{dq}{q^{a+1}}} = e^{-|u|^{2a}},$$

где $H(q) = \frac{q}{1+q}$, а в показателе экспоненты стоит интегральная сумма положительно определенных функций. Отсюда следует положительная определенность функций вида $e^{-c|u|^{2a}}$ для любых $a \in (0,1)$, $c > 0$. $\square$

Функция $h_P(u) = e^{a(e^{iub}-1)}$ является х. ф. для пуассоновского распределения $p(\xi = bn) = \frac{a^k}{k!}e^{-a}$, а функция $h_G(u) = e^{-\frac{(u\sigma)^2}{2}+iuc}$ является х. ф. нормального распределения $\mathcal{N}(c,\sigma)$:

$$p(x) = (2\pi\sigma^2)^{-1/2}\exp\{-(x-c)^2/2\sigma^2\},$$

В силу леммы Шура, их произведение также является х. ф.:

$$h(u) = h_G(u)h_P(u) = e^{\psi(u)}, \ \psi(u|a,b,c) = a(e^{iub}-1) - \frac{1}{2}(u\sigma)^2 + iuc.$$

Классу х. ф. принадлежат также функции вида $h(u) = e^{\Phi(u)}$, где

$$\Phi(u) = iuc - \frac{(u\sigma)^2}{2} + \int_{\mathbb{R}} \left(e^{iup} - 1 - \frac{iup}{1+p^2}\right)\frac{1+p^2}{p^2}\mu(dp) \qquad (3)$$

обладают свойством *безграничной делимости*:

$\forall n > 0: \quad h_{c,\sigma^2,\mu}(u) = h^n_{c/n,\sigma^2/n,\mu/n}(u)$. Под интегралом ( ) -функция, равномерно ограниченная по p при каждом u и положительно определенная при каждом p, а μ -любая конечная неотрицательная мера.

Распределение Хольцмарка

Эмпирические данные с распределениями, характеристические функции которых имеют вид $h_\xi(u) \stackrel{\mathrm{def}}{=} \mathbb{E}\, e^{i\xi\, u} = e^{-c|u|^a}$, $0 < a \leq 2$, встречаются в статистиках, имеющих физическое и экономическое происхождение (распределения напряжений в кристаллических решётках, температуры в ядерном реакторе, магнитного поля ансамблей заряженных спиновых частиц).

Распределение Хольцмарка является модельным распределением флуктуаций гравитационного поля во Вселенной. Пусть звездная материя распределена случайным образом: случайными являются координаты и массы звезд, а также их плотность во Вселенной. Сформулируем предположения Хольцмарка.

1) Флуктуации числа звезд n_V в объеме V имеют пуассоновское распределение:
$P(n_V = N) = \frac{(V\rho)^N}{N!} e^{-\rho V}$, где ρ – средняя плотность.
2) Сила притяжения, действующая на единичную массу в точке $x \in \mathbb{R}^3$ со стороны звезды массы m, находящейся в точке y, описывается законом Ньютона: $F = \gamma m \frac{x-y}{|x-y|^3}$.
3) Масса звезд является случайной величиной, имеющей конечный момент порядка $a = \frac{3}{2}$:
$\int_0^\infty m^{3/2} p(dm) = \mu_{3/2} < \infty$. Этот момент входит в распределение случайной силы.
4) Случайные величины n_V, y, m независимы; точка y равномерно распределена в любом содержащем ее объеме: $P(y \in d^3y|V) = V^{-1}d^3y$.

Вычислим характеристическую функцию случайной силы F, создаваемой n_V звездами в объеме V, действующей на единичную массу, помещенную в начало координат. Предполагается, что число звезд n_V в объеме V имеет пуассоновское распределение, а х. ф. $\mathbb{E}\, e^{iuF_V}$ вклада объема V в случайную силу F равна

$$h_F(u,V) = \mathbb{E}\, e^{i\gamma \sum_{n=0}^{n_V} m_n \frac{(u,y_n)}{|y_n|^3}} = \sum_{N=0}^{\infty} \frac{(V\rho)^N\, e^{-\rho V}}{N!} \prod_{n=1}^{N} \mathbb{E}\, e^{i\gamma m_n \frac{(u,y_n)}{|y_n|^3}}$$

$$= \sum_{N=0}^{\infty} \frac{(V\rho)^N}{N!} e^{-\rho V} \left(\mathbb{E}\, e^{i\gamma m \frac{(u,y)}{|y|^3}} \right)^N = e^{\rho\, \mathbb{E} \int_V \left(e^{i\gamma m \frac{(u,y)}{|y|^3}} - 1 \right) d^3 y}.$$

Последнее математическое ожидание вычисляется по распределению массы.

Упражнение 5

Убедитесь, что $\left(r^2\sin\frac{1}{r^2}-1\right)r^2=0(r^{-2})$ *при* $r\to\infty$ *и* $\int_0^\infty\left(r^2\sin\frac{1}{r^2}-1\right)r^2dr=-\frac{2\sqrt{2\pi}}{15}$.

Файл e3-5H

Поэтому существует предел таких интегралов при $V\to\infty$.

Вычислим интеграл $I(m)=\int_{\mathbb{R}^3}\left(e^{i\gamma m\frac{(u,y)}{|y|^3}}-1\right)d^3y$ в полярным координатах:

$$I=2\pi\int_0^\infty r^2dr\int_0^\pi\sin\phi\,d\phi\left(e^{i\gamma m\frac{|u|\cos\phi}{r^2}}-1\right)=$$

$$=4\pi\int_0^\infty\left(\frac{r^2}{\gamma m|u|}\sin\frac{\gamma m|u|}{r^2}-1\right)r^2dr=$$

$$=4\pi(\gamma m|u|)^{\frac{3}{2}}\int_0^\infty\left(r^2\sin\frac{1}{r^2}-1\right)r^2dr=-\frac{4}{15}(2\pi\gamma m|u|)^{\frac{3}{2}}.$$

Вычисляя математическое ожидание $\mathbb{E}\, I(m)$ по распределению массы, находим характеристическую функцию случайной силы:

$$\boxed{h_F(u) = e^{-\frac{4\mu_{3/2}}{15}(2\pi\gamma|u|)^{\frac{3}{2}}}}$$

Фурье-прообраз этой функции неотрицателен и его норма в $L_1(\mathbb{R}\,)$ равна $h_F(0) = 1$, однако явное аналитическое выражение для плотности вероятности через элементарные функции не выражается.

Теорема Бохнера–Хинчина

Важность леммы Шура в том, что она не использует предположения о том, что функции со свойствами *1– 3* являются характеристическими функциями мер. Это свойство алгебры характеристических функций устанавливает *теорема Бохнера–Хинчина*.

Теорема 3

Любая ограниченная измеримая функция $h(u)$, обладающая свойствами $1)–3)$, является характеристической функцией некоторой вероятностной меры.

Доказательство. Докажем теорему сначала для интегрируемых функций $h(u) \in L_1(\mathbb{R}^d) \cap L_2(\mathbb{R}^d)$, обладающих свойствами 1–3. Для таких функций корректно определено преобразование Фурье:

$$p(x) \stackrel{\text{def}}{=} \frac{1}{(2\pi)^d} \int e^{-ixu} h(u)\, du \in L_2(\mathbb{R}^d). \tag{4}$$

Непрерывность $p(x)$ следует из теоремы Лебега о предельном переходе под знаком интеграла для функций из $h \in L_1(\mathbb{R}^d)$. Покажем, что из положительной определенности $h(u)$ следует $p(x) \geq 0$, а из условия $h(0) = 1$ следует условие нормировки: $\int p(x)\, dx = 1$.

Рассмотрим двойной интеграл от непрерывной функции h, допускающий аппроксимацию положительными конечными суммами по непересекающимся разбиениям d-мерного куба $Q_N = [-N, N]^d = \bigsqcup_{k=1}^{K} \Delta_k$. Пусть $m_k = \mathrm{mes}\, \Delta_k$, тогда

$$J_N(x) \overset{\text{def}}{=} \frac{1}{(2N)^d} \int_{Q_N} \int_{Q_N} e^{i(p',x)-i(p,x)} h(p-p')\, dp\, dp'$$

$$= \frac{1}{(2N)^d} \lim_{M\to\infty} \sum_{j,k=1}^{M} e^{i(p_k,x)-i(p_j,x)} h(p_j - p_k)\, m_j\, m_k$$

$$= \frac{1}{(2N)^d} \lim_{M\to\infty} \sum_{j,k=1}^{M} h(p_j - p_k)\, \bar{z}_j z_k \geq 0, \quad z_k = m_k e^{i(p_k,x)}.$$

Поэтому $J_N(x) \geq 0$ для любой положительно определенной функции h.

Интегрируя функцию $e^{i(p',x)-i(p,x)}h(p-p')$ по d-мерному кубу Q_N, перейдем к новым переменным $u=p-p'\in[-2N,2N]$, $v=p'+p\in[-2N,2N]$:

$$J_N(x)=\frac{1}{(2N)^d}\int_{Q_N}\int_{Q_N}e^{i(p',x)-i(p,x)}h(p-p')\,dp\,dp'$$

$$=\frac{1}{(4N)^d}\Bigg(\int_{u:\,u_k\in[0,2N]}e^{-i(u,x)}h(u)du\int_{v:\,v_k\in[-2N+u_k,2N-u_k]}dv$$

$$+\int_{u:\,u_k\in[-2N,0]}e^{-i(u,x)}h(u)du\int_{v:\,v_k\in[-2N-u_k,2N+u_k]}dv\Bigg)$$

$$=\int_{\mathbb{R}^d}I_{Q_N}(u)\,e^{-i(u,x)}h(u)\prod_{k=1}^{d}\left(1-\frac{|u_k|}{2N}\right)du.$$

Совершая замену $p - p' = u$, $p + p' = v$ и разбивая область интегрирования на области $D_1 = \{u_k \in [0, 2N], v_k \in [-2N + u_k, 2N - u_k]\}$ и $D_2 = \{u_k \in [-2N, 0], v_k \in [-2N - u_k, 2N + u_k]\}$, получаем новую область интегрирования. Геометрическая интерпретация состоит в том, что при замене $(p, p') \to (u, v)$

$$p - p' = u, \quad p + p' = v,$$
$$p = \frac{u + v}{2}, \quad p' = \frac{v - u}{2}.$$

сторона квадрата меняется в $\sqrt{2}$ раза по каждой переменной.

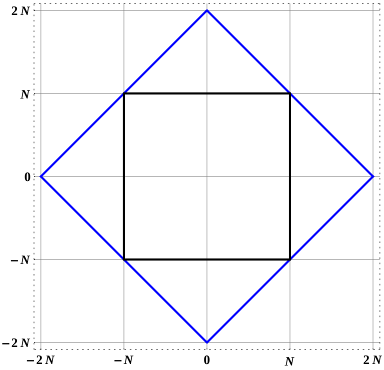

Рис. 2: Области интегрирования в переменных (p, p') (черная граница) и (u, v) (синяя граница)

Поскольку $|h(u)| \in L_1(\mathbb{R}^d)$, $0 \le I_{Q_N}(u) \uparrow 1$ и $0 \le 1 - \frac{|u|}{2N} \uparrow 1$ при $N \to \infty$ в каждой точке, то теорема Лебега о предельном переходе под знаком интеграла позволяет вычислить предел:

$$\lim_{N\to\infty} J_N(x) = \int_{\mathbb{R}^d} e^{-iux} h(u) \lim_{N\to\infty} I_{Q_N}(u) \prod_k \left(1 - \frac{|u_k|}{2N}\right) du =$$
$$= \int_{\mathbb{R}^d} e^{-iux} h(u)\, du = (2\pi)^d p(x) \ge 0.$$

Таким образом, *Фурье-прообразы положительно определенных функций положительны в обычном смысле*:

$$p(x) = \frac{1}{(2\pi)^d} \int e^{-ixu} h(u)\, du = \frac{1}{(2\pi)^d} \lim_{N\to\infty} \int_{Q_N} e^{-ixu} h(u)\, du \quad (5)$$
$$= \frac{1}{(2\pi)^d} \lim_{N\to\infty} J_N(x) \ge 0.$$

Учитывая, что из $h \in L_2(\mathbb{R}^d)$ следует $p \in L_2(\mathbb{R}^d)$, вычислим преобразование Фурье:

$$\begin{aligned} \mathbb{E}\, e^{i(x,u)} &= \int e^{i(x,u)} p(x)\, dx \\ &= \int h(v)\, dv \frac{1}{(2\pi)^d} \int e^{i(x,v)-i(x,u)}\, dx = h(v). \end{aligned} \tag{6}$$

В силу непрерывности h, при $u = 0$ имеем

$$\int p(x)\, dx = h(0) = 1,$$

следовательно, $p(x) \in L_1(\mathbb{R}^d)$ – плотность вероятности. $\square$

Заметим, что функция $h_\varepsilon(x) = e^{-\varepsilon u^2}$ является характеристической функцией нормального распределения с нулевым средним и дисперсией $\frac{1}{2\varepsilon}$. Если функция $h(u)$ не является абсолютно интегрируемой функцией из L_2, рассматривается семейство функций $h_\varepsilon(u) = e^{-\varepsilon u^2} h(u) \in L_1 \cap L_2$, обладающих указанными выше свойствами и являющихся характеристическими функциями вероятностных мер $P_\varepsilon[a, b) = \int_a^b p_\varepsilon(\xi) d\xi$.

Можно доказать, что из сходимости $h_\varepsilon(u) \to h(u)$ в каждой точке $u \in \mathbb{R}$ при $\varepsilon \to 0$ следует слабая сходимость вероятностных мер P_ε к вероятностной мере $P: \ P(A) = \int_A p(x)dx$ на любом ограниченном измеримом множестве A. Действительно, для любой гладкой финитной функции $f(x)$, имеющей гладкий, быстро убывающий фурье-образ $\widetilde{f}(u)$, выполнено *равенство Парсеваля*:

$$\int f(x)\, p_\varepsilon(x)\, dx = \int f(x')\, p_\varepsilon(x)\, dx' dx\, \frac{1}{(2\pi)^d} \int du\, e^{iu(x-x')} = \int \widetilde{f}(-u)\, h_\varepsilon(u)\, du, \tag{7}$$

где $\quad \widetilde{f}(u) \stackrel{\text{def}}{=} \frac{1}{(2\pi)^d} \int dx\, e^{iux} f(x).$

Напомним, что из $f \in C_0^\infty$ следует, что $\widetilde{f} \in L_1$, поэтому в силу равномерной оценки $|h_\varepsilon(u)| \leq 1$ и теоремы Лебега, в правой части равенства (6) можно совершить предельный переход при $\varepsilon \to 0$. Он используется как определение левой части при $\varepsilon = 0$, а правая часть допускает равномерную по ε оценку:

$$\left| \int f(x) p_\varepsilon(x)\, dx \right| \leq \sup_x |f(x)| \stackrel{\text{def}}{=} ||f||_C,$$

определяющую линейный ограниченный функционал

$$p(f) = \int f(x) p(x)\, dx$$

на множестве непрерывных ограниченных функций f, обращающихся в нуль на бесконечности.

Его можно продолжить на множество ограниченных функций, содержащее индикаторные функции открытых и замкнутых ограниченных борелевских множеств B:

$$p(I_B) = \sup_{f(x)<I_B(x)} p(f), \quad p(I_{\overline{B}}) = \inf_{f(x)>I_B(x)} p(f).$$

В итоге каждой функции со свойствами 1–3 сопоставляется вероятностная мера $P(B) = p(I_B)$, для которой в каждой точке непрерывности имеет место формула обращения (4).
Если $h \in L_1(\mathbb{R})$, то вероятностная мера имеет плотность абсолютно непрерывную относительно меры Лебега, а формула обращения имеет простой вид:

$$\boxed{P(a,b) = \int I_{(a,b)}(x)\,\frac{dx}{2\pi}\int h(u)e^{-iux}\,du = \frac{1}{2\pi}\int \frac{e^{-iua}-e^{-iub}}{iu}\,h(u)\,du} \tag{8}$$

Если вероятностная мера содержит вклады от дельта-функций, являющихся непрерывными линейными функционалами на $C_0^\infty(\mathbb{R}^d)$, то формула (7) обобщается

$$\boxed{P(\otimes_{k=1}^d [a_k, b_k]) = \lim_{N\to\infty} \frac{1}{(2\pi)^d} \int_{Q_N} \prod_{k=1}^d \frac{e^{-iu_k a_k} - e^{-iu_k b_k}}{iu_k} h(u)\, du}$$

где Q_N –компакт в $\mathbb{R}^d$ и $Q_N \uparrow \mathbb{R}^d$ при $N \to \infty$.
Рассмотрим пример. Пусть $d = 1$, $\{p_k\} : \sum_k p_k = 1$, $h(u) = \sum_k p_k e^{iuy_k}$ -равномерно ограниченная функция, тогда

$$P(a, b) = \lim_{N\to\infty} \frac{1}{2\pi} \int_{Q_N} \sum_k p_k \frac{e^{-iu(a-y_k)} - e^{-iu(b-y_k)}}{iu} du.$$

Учитывая, что вклад $\cos u(a-y_k)$ в интеграл равен нулю, а

$$\lim_{N\to\infty}\frac{1}{\pi}\int_{-N}^{N}\frac{\sin xu}{u}\,du=\lim_{\epsilon\to+0}\frac{1}{\pi}\int_0^x dy\int_{\mathbb{R}}\cos yu\,e^{-\frac{u^2\epsilon}{2}}\,du=\operatorname{sign}x,$$

в точках непрерывности получим

$$P(a,b)=\sum_k\frac{p_k}{2}(\operatorname{sign}(b-y_k)-\operatorname{sign}(a-y_k))=\sum_k p_k I_{(a,b)}(y_k),$$

т. е. $p(x)=\sum_k p_k\delta(x-y_k)$. Такие меры корректно определены на множестве непрерывных равномерно ограниченных или финитных функций, а вероятностные меры, абсолютно непрерывные относительно меры Лебега, корректно определены на более широком множестве ограниченных измеримых функций.

Центральная предельная теорема

Центральная предельная теорема описывает *слабую* сходимость распределений к нормальному закону, то есть сходимость математических ожиданий для достаточно большого множества функций случайной величины:

$$P_n \xrightarrow{w} P \quad \text{или} \quad \xi_n \xrightarrow{w} \xi, \quad \text{если}$$

$$\lim_{n\to\infty} \mathbb{E}_n f(\xi_n) = \lim_{n\to\infty} \int_X f(x) P_n(dx) = \int_X f(x) P(dx) = \mathbb{E} f(\xi)$$

$$\forall f \in \mathbb{C}_0(X).$$

Рассматриваемые ниже слабо сходящиеся последовательности являются примерами *плотных* семейств вероятностных распределений: “*плотность*” семейства распределений $\{P_n\}$ в топологическом пространстве X означает существование измеримых компактов $K_\epsilon \subset X$, таких, что

$$\sup_n P_n(X \setminus K_\epsilon) \leq \epsilon \qquad \forall \epsilon > 0.$$

Замечание 1

Семейство мер в полном метрическом пространстве плотно, если абсолютные моменты какого-либо порядка $a \geq \epsilon > 0$ *равномерно ограничены:* $\sup_n \mathbb{E}_n \rho_\xi^a < \infty$, *где* ρ_ξ *– расстояние от* $\xi \in X$ *до точки* $0 \in X$.

Действительно, из неравенства Чебышева следует, что в качестве компактов K_ϵ для равномерной по n оценки моментов и доказательства плотности можно использовать шары достаточно большого радиуса:

$$P_n(\rho_\xi > K) \leq \frac{\mathbb{E}\rho_\xi^a}{K^a} = \frac{m_a}{K^a} \leq \epsilon \qquad \forall K > K_\epsilon = (m_a/\epsilon)^{1/a}.$$

Семейство вероятностных мер $\{P_n\}$ называется *относительно слабо компактным*, если любая подпоследовательность мер содержит слабо сходящуюся подпоследовательность.
В весьма общей ситуации плотные семейства относительно слабо компактны: теорема Прохорова утверждает, что в полных метрических сепарабельных пространствах *плотность* и *относительная компактность* эквивалентны.
В случае центральной предельной теоремы единственность предельной точки оказывается очевидным фактом. Как мы увидим далее, доказательство ЦПТ сводится к проверке плотности семейства выборочных распределений.

В конечномерном случае *равномерная непрерывность* семейства характеристических функций в нуле влечет *плотность* соответствующего семейства вероятностных мер. Докажем это утверждение в одномерном случае.

Лемма 3

Пусть $a \stackrel{\text{def}}{=} 1/(1-\sin 1) = \mathrm{const}$ *и* $h_\xi(u) = \mathbb{E}\, e^{i\xi u}$ *– характеристическая функция случайной величины* ξ *с распределением* $P_\xi(dx)$*. Тогда*

$$P(|\xi| > K) = \int_{|x|\geq K} P_\xi(dx) \leq aK \int_0^{1/K} (1 - \mathrm{Re}\, h_\xi(u))\, du. \qquad (9)$$

Доказательство. Для вещественной части характеристической функции $\mathrm{Re}\, h_\xi(u) = \int \cos\, ux\, P_\xi(dx)$ справедливы оценки:

$$K \int_0^{1/K} (1 - \mathrm{Re}\, h_\xi(u))\, du = K \int_0^{1/K} \int_{\mathbb{R}} (1 - \cos\, ux)\, du\, P_\xi(dx) =$$

$$= \int_{\mathbb{R}} \left(1 - \frac{\sin\, x/K}{x/K}\right) P_\xi(dx) \geq \int_{|x| \geq K} \left(1 - \frac{\sin\, x/K}{x/K}\right) P_\xi(dx)$$

$$\geq \inf_{y \geq 1} \left(1 - \frac{\sin y}{y}\right) P_\xi(|x/K| \geq 1) \geq \frac{1}{a} P_\xi(|x| \geq K),$$

где $\inf_{y \geq 1}\left(1 - \frac{\sin y}{y}\right) = 1 - \sin 1 = \frac{1}{a} \approx 0.1585$. Отсюда следует (8). □

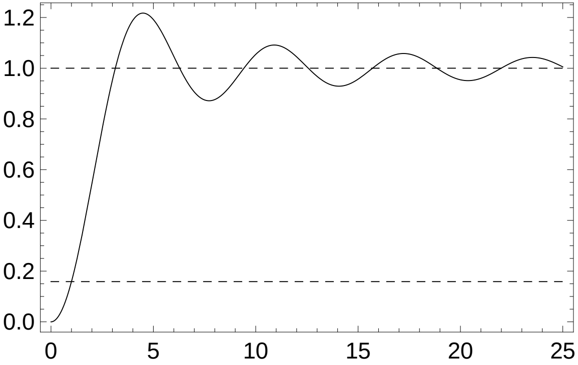

Рис. 3: Непрерывная линия изображает $1 - \frac{\sin y}{y}$, а нижняя пунктирная - постоянную $1 - \sin 1$.

Следствие 1

Если семейство характеристических функций $h_n(u)$ равностепенно непрерывно в нуле, то есть для любого $\epsilon > 0$ существует открытая окрестность нуля U_ϵ такая, что

$$\sup_n |1 - h_n(u)| \leq \epsilon \qquad \forall u \in U_\epsilon,$$

то соответствующее семейство мер P_n слабо компактно.

Доказательство. Из (8) имеем

$$\sup_n P_n(|x| \geq K) \leq aK \int_0^{1/K} \sup_n (1 - \mathrm{Re}\, h_n(u))\, du \leq a\,\epsilon(K),$$

$$\text{где } \epsilon(K) = \sup_{|u| \leq 1/K} \sup_n (1 - \mathrm{Re}\, h_n(u)) \to 0$$

при $K \to \infty$, поэтому семейство P_n плотно. Cor В силу теоремы Прохорова оно слабо компактно . □

Лемма 4

Для любых $x \in \mathbb{R}$ и $K \geq 0$ погрешность разложения экспоненты в ряд Тейлора допускает оценку:

$$\rho_K(x) \stackrel{\text{def}}{=} e^{ix} - \sum_{k=0}^{K-1} \frac{(ix)^k}{k!} = \frac{(ix)^K}{K!}\,\theta_K(x), \quad |\theta_K(x)| \leq 1. \qquad (10)$$

Доказательство. Нетрудно проверить тождества

$$\rho_1(x) = e^{ix} - 1 = i\int_0^x e^{iy}\,dy = ix\,\theta_1(x), \quad \rho_{k+1}(x) = i\int_0^x \rho_k(y)dy,$$

где $\theta_n(x) = \int_0^1 dy_1 \cdots \int_0^{y_{n-1}} dy_n e^{ixy_n}$, $|\theta_n(x)| \leq 1$. По индукции

$$|\rho_{n+1}(x)| \leq \int_0^x |\rho_n(y)|\,dy \leq \int_0^x \frac{y^n|\theta_n(y)|}{n!}\,dy \leq \frac{x^{n+1}}{(n+1)!}.$$

Пусть $\{x_n\}$ – одинаково распределенные независимые случайные величины в $\mathbb{R}^m$ с распределением $P(dx)$, средним $\mu_1 = \mathbb{E}\, x \in \mathbb{R}^m$. Обозначим через $\xi = x - \mathbb{E}\, x$ скомпенсированную случайную величину и через $h(u) = \mathbb{E}\, e^{i(\xi,u)}$ её х. ф. Ковариационная матрица $R = \mathbb{E}\, \xi \otimes \xi$ предполагается невырожденной.

В силу независимости случайных величин ξ_n х. ф. $h_N(u)$ суммы $\{\frac{\xi_n}{\sqrt{N}}\}_{n=1}^N$ равна произведению их х. ф.:

$$\chi_N = \frac{1}{\sqrt{N}} \sum_1^N \xi_n, \quad h_N(u) = \mathbb{E}\, e^{i\left(u, \frac{1}{\sqrt{N}} \sum_1^N \xi_n\right)} = \left(\mathbb{E}\, e^{iu\frac{\xi}{\sqrt{N}}}\right)^N.$$

Покажем, что семейство х. ф. $h_N(u) = h\left(\frac{u}{\sqrt{N}}\right)^N$ равностепенно непрерывно в нуле, что влечет относительную компактность мер μ_n с этими х. ф. Обозначим через $S_N \subset \mathbb{R}^m$ – шар радиуса $r_N = o(N^{1/2})$ с центром в начале координат и пусть $R = \mathbb{E}\,\xi \otimes \xi$. Тогда

$$\mathbb{E}\,(\xi, u)^2 I_{S_N}(\xi) = (u, \mathbb{E}\,\xi \times \xi\, u)\, I_{S_N}(\xi) = (u, Ru) + \epsilon_N(u), \tag{11}$$

при $N \to \infty$ по теореме Лебега о предельном переходе под знаком интеграла равномерно на ограниченных по u множествах, так как $\mathbb{E}(\xi, u)^2 < \infty$ и $(\xi, u)^2 I_{S_N}(\xi) \uparrow (\xi, u)^2$ в каждой точке ξ. С другой стороны,

$$\mathbb{E}|(\xi, u)|^3 I_{S_N}(\xi) \leq \max_{\xi \in S_N} |(\xi, u)|\, \mathbb{E}|(\xi, u)|^2 I_{S_N}(\xi) \leq r_N |u| (u, Ru). \tag{12}$$

Используя оценки (9) представим $e^{i(u,\xi)}$ в виде суммы

$$\begin{aligned}
e^{i(u,\xi)} =& e^{i(u,\xi)}\left(I_{S_N}(\xi) + I_{\mathbb{R}\setminus S_N}(\xi)\right) \\
=& \left(1 + i(u,\xi) + \frac{(i(u,\xi))^2}{2!} + \rho_3((u,\xi))\right) I_{S_N}(\xi) \\
&+ \left(1 + i(u,\xi) + \rho_2((u,\xi))\right) I_{\mathbb{R}\setminus S_N}(\xi) \\
=& 1 + i(u,\xi) - \frac{(u,\xi)^2}{2!} I_{S_N}(\xi) - \frac{i(u,\xi)^3}{3!}\theta_3\left((u,\xi)\right) I_{S_N}(\xi) \\
&- \frac{(u,\xi)^2}{2!}\theta_2\left((u,\xi)\right) I_{\mathbb{R}\setminus S_N}(\xi), \qquad (13)
\end{aligned}$$

где $\mathbb{E}\xi = 0$, $|\theta_2(\xi)|$, $|\theta_3(\xi)| \leq 1$ и для матожиданий последних двух слагаемых имеются оценки (10)-(11).

Полагая $u \to \frac{u}{\sqrt{N}}$ в (12) и учитывая $\mathbb{E}\xi = 0$, вычислим $\mathbb{E}\, e^{i\frac{(u,\xi)}{\sqrt{N}}}$:

$$\begin{aligned}
\mathbb{E}\, e^{i\left(\frac{u}{\sqrt{N}},\xi\right)} =& 1 - \mathbb{E}\frac{(u,\xi)^2}{2!N} I_{S_N}(\xi) - \mathbb{E}\frac{i(u,\xi)^3}{3!N^{3/2}}\,\theta_3\left(\frac{(u,\xi)}{\sqrt{N}}\right) I_{S_N}(\xi) \\
& - \mathbb{E}\frac{(u,\xi)^2}{2!N}\,\theta_2\left(\frac{(u,\xi)}{\sqrt{N}}\right) I_{\mathbb{R}\setminus S_N}(\xi) \\
=& 1 - \frac{(u,Ru)}{2N} + \frac{|u|^2\, O(\epsilon_N)}{N} + \mathbb{E}\,\frac{(u,\xi)^3}{N^{3/2}}\, I_{S_N}(\xi)\, O(1) \\
=& 1 - \frac{(u,Ru)}{2N} + |u|^2\,\frac{O(\epsilon_N)}{N} + |u|^3\,\mathbb{E}|\xi|^2\,\frac{O(r_N)}{N^{3/2}},
\end{aligned}$$

где $\frac{\epsilon_N}{2N} = o(N^{-1})$, $r_N = o(N^{1/2})$, $\frac{O(r_N)}{N^{3/2}} = o(N^{-1})$ равномерно по u в любой компактной окрестности начала координат.

Используя разложение логарифма в ряд Тейлора, $\ln(1+x) = x + O(x^2)$, мы видим, что имеет место равномерная сходимость любой подпоследовательности $\{h_{N_k}(u)\}_{k=1}^{\infty}$ к пределу

$$h_N(u) = \left(\mathbb{E}\, e^{i\frac{(u,\xi)}{\sqrt{N}}}\right)^N = e^{N\log\left(1-\frac{(u,Ru)}{2N}+|u|^2\left(\frac{O(\epsilon_N)}{N}+|u|\frac{O(r_N)}{N^{3/2}}\right)\right)}$$

$$= e^{-\frac{(u,Ru)}{2}+|u|^2\left(O(\epsilon_N)+o_N(1)|u|\right)} \to e^{-\frac{(u,Ru)}{2}}, \quad N\to\infty,$$

причем семейство $h_N(u)$ непрерывно по u в любой компактной окрестности начала координат *равномерно* относительно N так как $\epsilon_N \to 0$ и $r_N/\sqrt{N} \to 0$ при $N \to \infty$.

Согласно оценке (8), равномерно непрерывному семейству $\{h_N(u)\}$ соответствует плотное семейство вероятностных мер, которое слабо сходится в силу теоремы Прохорова. Поэтому из поточечной сходимости этих характеристических функций $h_N(u)$ к х. ф. нормального распределения с корреляционной матрицей R следует, что распределение суммы $\chi_N = \frac{\sum_1^N \xi_n}{\sqrt{N}}$ в слабо сходится к соответствующему нормальному распределению.

$\square$

Отметим, что приведенное доказательство не предполагает существования абсолютных моментов выше второго.

Рассмотрим формулировку ЦПТ, обосновывающую применение метода Монте–Карло для вычисления математических ожиданий скалярных величин.

Теорема 4

Пусть ξ_n *– i.i.d.r.v. в* $\mathbb{R}$ *со средним* $\mu = \mathbb{E}\xi = \langle\,\xi\,\rangle$ *и дисперсией* $\sigma^2 < \infty$*. Тогда выборочное среднее* $\langle\,\xi\,\rangle_N = \frac{1}{N}\sum_{n=1}^{N}\xi_n$ *сходится к* $\langle\,\xi\,\rangle$ *со скоростью* $c\sqrt{\frac{\sigma^2}{N}}$*, где* c *определяет надежность оценки погрешности* $\epsilon = |\langle\,\xi\,\rangle_N - \langle\,\xi\,\rangle|$*. Более точно,*

$$\lim_{N\to\infty} P\left(|\langle\,\xi\,\rangle_N - \langle\,\xi\,\rangle| > x\sqrt{\frac{\sigma^2}{N}}\right) = \int_{|y|>x} \frac{e^{-y^2/2}}{\sqrt{2\pi}}\,dy.$$

Центральная предельная теорема в форме Ляпунова и Линдеберга

Приведем формулировку ЦПТ, принадлежащую Ляпунову, ученику Чебышева, предполагающую существование третьего абсолютного момента для случайных величин x_n с *различными* средними μ_n и дисперсиями σ_n^2.

Теорема Чебышева допускает рост дисперсии $\sigma^2(n)$ и третьего абсолютного момента $\mu_3(n)$ при условии, что

$$\frac{\sqrt[3]{\sum_{n=1}^{N} \mu_3(n)}}{\sqrt[2]{\sum_{n=1}^{N} \sigma^2(n)}} \to 0 \ \text{при} \ \ n \to \infty$$

Теорема 5

Пусть x_n *– независимые действительные случайные величины,* $D_N^2 \stackrel{\mathrm{def}}{=} \sum_{n=1}^N \sigma_n^2$, $r_N^3 \stackrel{\mathrm{def}}{=} \sum_{n=1}^N \mathbb{E}\left|x_n - \mu_n\right|^3$. *Если* $r_N/D_N \to 0$ *при* $N \to \infty$, *то*

$$\lim_{N\to\infty} P(X_N \in B) = \frac{1}{\sqrt{2\pi}} \int_B e^{-x^2/2}\, dx, \quad X_N \stackrel{\mathrm{def}}{=} \sum_{n=1}^N \frac{x_n - \mu_n}{D_N}.$$

Условие Ляпунова является достаточным для сходимости к нормальному закону. Линдеберг нашел более слабые достаточные условия, которые в ряде важных случаев являются *необходимыми*.

Теорема 6

Пусть x_n *– независимые д. с. в. с распределениями* $P_n(dx)$, *средним* μ_n, *дисперсией* σ_n *и* $D_N^2 \stackrel{\text{def}}{=} \sum_{n=1}^N \sigma_n^2$,

$$\delta_N(\epsilon) \stackrel{\text{def}}{=} \frac{1}{D_N^2} \sum_{n=1}^{N} \left(\int_{-\infty}^{\mu_n - \epsilon D_N} + \int_{\mu_n + \epsilon D_N}^{\infty} \right) (x - \mu_n)^2 \, P_n(dx).$$

Если $\delta_N(\epsilon) \to 0$ *при* $N \to \infty$ *для* $\forall \epsilon > 0$, *то для последовательности д. с. в.* $X_N \stackrel{\text{def}}{=} \sum_{n=1}^N \frac{x_n - \mu_n}{D_N}$ *существует предел*

$$\lim_{N \to \infty} P(X_N \in B) = \frac{1}{\sqrt{2\pi}} \int_B e^{-x^2/2} \, dx.$$