Приложение
к программе курса "Теория вероятностей",
прочитанного В.В.Сенатовым весной 2009 г.

О СЛАБОЙ СХОДИМОСТИ РАСПРЕДЕЛЕНИЙ

В данном приложении рассматриваются некоторые виды сходимости. В теории вероятностей используются различные множества функций, среди которых множества случайных величин (заданных на некотором вероятностном пространстве), множество функций распределения, множество характеристических функций. Каждое из этих множеств имеет свою специфику, поэтому на каждом из них вводятся свои виды сходимости и возникают вопросы о связях между различными видами сходимостей, определенными на одном и том же множестве функций. Так как упомянутые множества связаны друг с другом, то возникают и вопросы о связях между сходимостями на одном из них со сходимостями на других. Ниже мы рассмотрим только связь между сходимостью по вероятности на множестве случайных величин и слабой сходимостью функций распределения этих случайных величин.

Большая часть этого приложения посвящена вопросам, связанным с различными видами сходимости на множестве функций распределения.

Тот материал, который ниже изложен в замечаниях, является дополнительным и его знание на экзамене не требуется.

Определение 1. Говорят, что действительная функция $F(x)$, определенная для всех $-\infty < x < \infty$, является *функцией распределения*, если

1. $0 \leqslant F(x) \leqslant 1, -\infty < x < \infty$,
2. $F(x)$ не убывает на всей действительной оси,
3. $F(x)$ непрерывна слева для всех x,
4. $F(-\infty) = 0, F(+\infty) = 1$,

где $F(-\infty) = \lim\limits_{x\to-\infty} F(x), F(+\infty) = \lim\limits_{x\to\infty} F(x)$ (эти пределы существуют в силу монотонности и ограниченности $F(x)$).

Множество всех функций распределения далее обозначается $\mathcal{P}$.

Отметим, что функцией распределения действительной случайной величины ξ называется функция

$$F_\xi(x) = \mathsf{P}(\xi < x) = \mathsf{P}\left(\{\omega : \xi(\omega) < x\}\right).$$

При этом имеется в виду, что задано некоторое вероятностное пространство $(\Omega, \mathcal{F}, \mathsf{P})$ и на нем определена измеримая функция $\xi(\omega), \omega \in \Omega$. Функция распределения $F_\xi(x)$ определена для всех действительных x в силу измеримости функции $\xi(\omega)$, то есть в силу того, что для каждого действительного x множество $\{\omega : \xi(\omega) < x\}$ принадлежит σ-алгебре $\mathcal{F}$ подмножеств множества Ω, которая является областью определения вероятностной меры P. Иногда функцию распределения случайной величины определяют, заменяя $\mathsf{P}(\xi < x)$ на $\mathsf{P}(\xi \leqslant x)$, в этом случае функция распределения $F_\xi(x)$ оказывается непрерывной справа. В содержательных утверждениях теории вероятностей безразлично, как определяется $F_\xi(x)$ в точках разрыва: важны лишь положения точек разрыва функции $F_\xi(x)$ (если они есть) на действительной оси и величины приращений $F_\xi(x+0) - F_\xi(x-0) = \mathsf{P}(\xi = x)$ в точках разрыва.

Функции распределения случайных величин обладают свойствами 1. - 4., указанными выше. Обратно, для любой функции $F \in \mathcal{P}$ существует вероятностное пространство $(\Omega, \mathcal{F}, \mathsf{P})$ и измеримая функция ξ на этом пространстве такая, что $F_\xi(x) = F(x)$.

Сходимость в основном. Рассмотрим вначале три вида сходимости на множестве функций распределения $\mathcal{P}$.

Говорят, что последовательность функций распределения $\{F_n\}_{n\geqslant 1}$ сходится поточечно к функции F, если при $n \to \infty$ для любого фиксированного x, $-\infty < x < \infty$, числовая последовательность $\{F_n(x)\}_{n\geqslant 1}$ сходится к числу $F(x)$, то есть если для любого $\varepsilon > 0$ существует $n_\varepsilon(x)$, такое, что $|F_n(x) - F(x)| < \varepsilon$ при всех $n > n_\varepsilon(x)$. Если число $n_\varepsilon(x)$ можно выбрать одним и тем же для всех действительных x, то получается сходимость, которая называется равномерной. Равномерная сходимость метризуется метрикой

$$\rho(F, G) = \sup_{-\infty < x < \infty} |F(x) - G(x)|,$$

иначе говоря, последовательность функций распределения $\{F_n\}_{n\geqslant 1}$ равномерно сходится к функции F (которая необходимо является функцией распределения) тогда и только тогда, когда $\rho(F_n, F) \to 0, n \to \infty$. Наконец, будем говорить, что последовательность $\{F_n\}_{n\geqslant 1}$ функций распределения сходится в среднем к функции F (которая необходимо является функцией распределения), если

$$\kappa(F_n, F) = \int\limits_{-\infty}^{\infty} |F_n(x) - F(x)| dx \to 0, n \to \infty.$$

Величина $\kappa(F, G) = \int\limits_{-\infty}^{\infty} |F(x) - G(x)| dx$ называется средней метрикой или расстоянием в среднем между функциями распределения F и G.

Эти понятия сходимости пришли в теорию вероятностей из теории функций, они с успехом используются в различных задачах теории вероятностей (метрика ρ прекрасно работает в случаях, когда хотя бы одна из функций распределения F или G является гладкой, а метрика κ с успехом используется в случаях, когда у распределений F и G существуют моменты порядка выше первого), но в общих задачах теории вероятностей, когда изначально не предполагается ни гладкость исходных функций распределения, ни существование каких либо моментов, эти понятия сходимости мало пригодны.

Для того, чтобы получить естественное для теории вероятностей понятие сходимости распределений, пригодное для общих задач теории вероятностей, встанем на следующую точку зрения. В общем курсе теории вероятностей сначала определяются случайные величины, а затем определяются их функции распределения (как характеристики этих случайных величин). Пусть $(\Omega, \mathcal{F}, \mathsf{P})$ некоторое достаточно богатое вероятностное пространство (скажем, интервал $(0, 1)$ с σ-алгеброй борелевских подмножеств и мерой Лебега), на котором существуют случайные величины с любыми функциями распределения из $\mathcal{P}$. Потребуем, чтобы для любой последовательности случайных величин $\{\xi_n\}_{n\geqslant 1}$ из сходимости $\xi_n(\omega) \to \xi(\omega), n \to \infty$, для всех $\omega \in \Omega$ следовала сходимость соответствующих функций распределения.

Убедимся в том, что перечисленные выше виды сходимости функций распределения этому требованию не удовлетворяют.

Рассмотрим вначале среднее расстояние κ. Пусть случайная величина η, определенная на вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P})$ имеет распределение Коши с плотностью $\frac{1}{\pi(1+x^2)}, -\infty < x < \infty$, и пусть F_η - соответствующая функция распределения. Определим случайные величины $\xi_n = \frac{\eta}{n}, n = 1, 2, \ldots$ Функции распределения этих случайных величин $F_{\xi_n}(x) = \mathsf{P}(\xi_n < x) = \mathsf{P}(\frac{\eta}{n} < x) = \mathsf{P}(\eta < nx) = F_\eta(nx)$. Ясно, что $\xi_n(\omega) = \frac{\eta(\omega)}{n} \to \xi(\omega) \equiv 0,\ \omega \in \Omega,\ n \to \infty$, и функция распределения предельной случайной величины есть E_0. Здесь и далее E_c обозначает вырожденную функцию распределения с единственным единичным скачком в точке c. Очевидно, что $E_0(x) = E_0(nx)$ для любого $n > 0$.

Величина

$$\kappa(F_{\xi_n}, F_\xi) = \int\limits_{-\infty}^{\infty} |F_{\xi_n}(x) - E_0(x)|dx = \int\limits_{-\infty}^{\infty} |F_\eta(nx) - E_0(nx)|dx =$$

$$= \frac{1}{n} \int\limits_{-\infty}^{\infty} |F_\eta(x) - E_0(x)| dx.$$

Все интегралы в этой цепочке равенств сходятся или расходятся одновременно и ясно, что последний интеграл расходится, поскольку на отрицательной полуоси

$$F_\eta(x) - E_0(x) = F_\eta(x) = \int\limits_{-\infty}^{x} \frac{1}{\pi(1+u^2)} du \sim \frac{1}{\pi|x|},\ x \to -\infty,$$

а на положительной полуоси

$$|F_\eta(x) - E_0(x)| = 1 - F_\eta(x) = \int\limits_{x}^{\infty} \frac{1}{\pi(1+u^2)} du \sim \frac{1}{\pi x},\ x \to \infty.$$

Таким образом, $\xi_n(\omega) \to \xi(\omega), n \to \infty, \omega \in \Omega$, но расстояния $\kappa(F_{\xi_n}, F_\xi)$ бесконечны.

Замечание. Можно показать, что интеграл $\int\limits_{-\infty}^{\infty} |F_\eta(x) - E_0(x)| dx$ сходится тогда и только тогда, когда у случайной величины η существует математическое ожидание и в этом случае

$$\int\limits_{-\infty}^{\infty} |F_\eta(x) - E_0(x)| dx = \beta_1 = \mathsf{E}|\eta|.$$

Для таких случайных величин η последовательность $\xi_n(\omega) = \frac{\eta(\omega)}{n} \to \xi(\omega) \equiv 0,\ \omega \in \Omega,\ n \to \infty$ и $\kappa(F_{\xi_n}, F_\xi) = \frac{\beta_1}{n} \to 0$, при $n \to \infty$.

Рассмотрим равномерное расстояние ρ. Пусть $\eta(\omega) \equiv 1$ и $\xi_n = \frac{\eta}{n},\ n = 1, 2, \ldots$. Ясно, что $\xi_n(\omega) = \frac{\eta(\omega)}{n} = \frac{1}{n} \to \xi(\omega) \equiv 0, \omega \in \Omega$, $n \to \infty$. Однако, функции распределения случайных величин ξ_n суть $E_{1/n}(x)$, функция распределения ξ есть $E_0(x)$ и $\rho(F_{\xi_n}, F_\xi) \equiv 1$ при всех $n = 1, 2, \ldots$, поскольку $E_0(x) - E_{1/n}(x) = 1$ при всех $x \in (0, 1/n)$.

Наконец, рассмотрим поточечную сходимость. Пусть η - случайная величина с функцией распределения F_η и пусть $\xi_n = \eta + \frac{(-1)^n}{n}, n = 1, 2, \ldots$ Очевидно, что

$$F_{\xi_n}(x) = \mathsf{P}\left(\eta + \frac{(-1)^n}{n} < x\right) = \mathsf{P}\left(\eta < x - \frac{(-1)^n}{n}\right) = F_\eta\left(x - \frac{(-1)^n}{n}\right).$$

Ясно, что для любого действительного x при четных n величины $F_{\xi_n}(x) \to F_\eta(x-0),\ n \to \infty$, а для нечетных n величины $F_{\xi_n}(x) \to F_\eta(x+0),\ n \to \infty$.

Если x - точка непрерывности функции $F_\eta(x)$, то $F_{\xi_n}(x) \to F_\eta(x)$, $n \to \infty$, однако ни в одной точке разрыва функции $F_\eta(x)$ такой сходимости нет, поскольку для таких x последовательность $\{F_{\xi_n}(x)\}_{n\geqslant 1}$ содержит две подпоследовательности, сходящиеся к различным пределам $F_\eta(x-0)$ и $F_\eta(x+0)$.

Последний пример наводит на мысль о том, что при определении вида сходимости, естественного для общих задач теории вероятностей, целесообразно сделать исключение для точек разрыва предельной функции.

Определение 2. Говорят, что последовательность функций распределения $\{F_n\}_{n\geqslant 1}$ сходится в основном к монотонной функции $F(x), -\infty < x < \infty$, если для любой фиксированной точки $x \in \mathbb{C}(F)$

$$F_n(x) \to F(x),\ n \to \infty,$$

где $\mathbb{C}(F)$ - множество точек непрерывности предельной функции F. Эту сходимость мы будем обозначать $F_n(x) \xrightarrow{\text{осн.}} F(x),\ n \to \infty$.

В этом определении ничего не говорится о точках разрыва предельной функции и мы можем и будем считать, что предельная функция $F(x)$ в точках разрыва непрерывна слева.

Легко проверить, что во всех примерах, рассмотренных выше, $F_n(x) \xrightarrow{\text{осн.}} F(x),\ n \to \infty$.

Замечание. В этом определении условие монотонности предельной функции потребовалось для того, чтобы избежать "тривиальных несуразностей" вроде следующей. Пусть функции $F_n(x)$, $n = 1, 2, \ldots$, таковы, что $F_n(x) = 0,\ x \leqslant 0,\ F_n(x) = 1,\ x > 1$, для всех $n = 1, 2, \ldots$, а функция $F(x) = 0,\ x \leqslant 0,\ F(x) = 1,\ x > 1$, и $F(x)$ при $x \in (0, 1]$ совпадает с функцией Дирихле. Если в определении 2 опустить требование монотонности предельной функции F, то получится, что последовательность $\{F_n\}_{n\geqslant 1}$ сходится к F как бы ни вели себя функции распределения $F_n(x)$ при $x \in (0, 1]$.

Очевидно, что предельная функция в определении 2 является неубывающей и что это определение корректно в том смысле, что если $F_n(x) \xrightarrow{\text{осн.}} F(x),\ n \to \infty$, то предельная функция единственна.

Определение 2 - лишь первый шаг к определению слабой сходимости, но прежде, чем идти дальше, необходимо рассмотреть некоторые свойства сходимости в основном.

Прежде всего заметим, что в определении 2 предельная функция не обязана быть функцией распределения. Простейший пример, показывающий это, связан с вырожденными функциями распределения $E_n(x),\ n = 0, \pm 1, \pm 2, \ldots$ Очевидно, что $E_n(x) \to 0$ для любого фиксированного x, поскольку все функции $E_n(x)$ равны нулю при $n > x$, то есть последовательность $\{E_n\}_{n\geqslant 1}$ при $n \to \infty$ сходится в основном к функции, тождественно равной нулю, а эта функция не является функцией распределения. Аналогично, последовательность $\{E_{-n}\}_{n\geqslant 1}$ сходится в основном к функции, тождественно равной единице. Эти последовательности не являются экзотическими. Такие же примеры дают последовательности функций распределения

$$G_n(x) = \mathsf{P}\left(\xi_1 + \cdots + \xi_n < x\right)$$

(ненормированных!) сумм независимых одинаково распределенных случайных величин $\xi_1, \xi_2, \ldots$. Пределом последовательности функций распределения $\{G_n\}_{n\geqslant 1}$ может быть либо функция, тождественно равная нулю (если $\mathsf{E}\xi_1 > 0$), либо функция, тождественно равная единице (если $\mathsf{E}\xi_1 < 0$), либо функция, тождественно равная $1/2$ (если $\mathsf{E}\xi_1 = 0$).

Нетрудно понять, что функция, предельная для последовательности функций распределения в смысле сходимости в основном, обладает лишь свойствами 1. - 3. из определения 1. Обозначим $\overline{\mathcal{P}}$ множество функций, обладающих этими свойствами. Это множество является замыканием множества $\mathcal{P}$ в смысле сходимости в основном. Действительно, каждая предельная (для последовательности функций распределения) функция принадлежит $\overline{\mathcal{P}}$ и любая функция F из $\overline{\mathcal{P}}$ является предельной для некоторой последовательности функций распределения, поскольку если $n \to \infty$, то последовательность функций распределения

$$F_n(x) = \begin{cases} 0, & x \leqslant -n, \\ F(x), & -n < x \leqslant n, \\ 1, & x > n \end{cases}$$

сходится в основном (и даже при каждом действительном x) к функции F. Функции из множества $\overline{\mathcal{P}} \setminus \mathcal{P}$ иногда называют несобственными функциями распределения.

Замечание. О последовательностях $\{F_n\}_{n\geqslant 1}$, таких, что $F_n(x) \xrightarrow{\text{осн.}} F(x),\ n \to \infty,\ F \notin \mathcal{P}$, иногда говорят, что они "уносят ненулевую вероятностную массу на бесконечность". Так, о последовательностях функций распределения $\{G_n\}_{n\geqslant 1}$ (они определены выше) таких, что $G_n(x) \to 0,\ n \to \infty$, для любого действительного x, говорят, что они уносят единичную вероятностную массу на бесконечность, о последовательностях $\{G_n\}_{n\geqslant 1}$, таких, что $G_n(x) \to 1,\ n \to \infty$, для любого действительного x, говорят, что они уносят единичную вероятностную массу на $-\infty$, о последовательностях $\{G_n\}_{n\geqslant 1}$, таких, что $G_n(x) \to 1/2,\ n \to \infty$, для любого действительного x, говорят, что они уносят половину вероятностной массы, связанной с распределениями сумм $\xi_1, \ldots, \xi_n$ на ∞, а вторую половину - на $-\infty$.

Оказывается, что множество $\overline{\mathcal{P}}$ обладает одним важным свойством - оно является компактом (для сходимости в основном).

Теорема 1. *Из любой последовательности* $\{F_n\}_{n\geqslant 1} \subset \overline{\mathcal{P}}$ *можно извлечь подпоследовательность* $\{F_n^{(n)}\}_{n\geqslant 1}$ *и указать функцию* $F \in \overline{\mathcal{P}}$ *такие, что*

$$F_n^{(n)}(x) \xrightarrow{\text{осн.}} F(x),\ n \to \infty.$$

Определение сходимости в основном для последовательности $\{F_n\}_{n\geqslant 1} \subset \overline{\mathcal{P}}$ дословно повторяет определение 2.

Доказательство теоремы 1 использует стандартную диагональную процедуру и опирается на сепарабельность действительной прямой.

Доказательство теоремы 1. Пусть $\{F_n\}_{n\geqslant 1}$ последовательность функций из $\overline{\mathcal{P}}$ и $\{r_k\}_{k\geqslant 1}$ счетное всюду плотное множество на прямой (например, множество всех рациональных чисел). Рассмотрим числовую последовательность $\{F_n(r_1)\}_{n\geqslant 1}$. Так как эта последовательность ограничена, то из нее можно выделить сходящуюся подпоследовательность

$$F_n^{(1)}(r_1) \to F(r_1),\ n \to \infty,$$

где мы обозначили $F(r_1)$ число, являющееся пределом подпоследовательности $\{F_n^{(1)}(r_1)\}_{n\geqslant 1}$.

Числовая последовательность $\{F_n^{(1)}(r_2)\}_{n\geqslant 1}$ также является ограниченной и из нее можно выделить сходящуюся подпоследовательность

$$F_n^{(2)}(r_2) \to F(r_2),\ n \to \infty,$$

где мы обозначили $F(r_2)$ число, являющееся пределом последовательности $\{F_n^{(2)}(r_2)\}_{n\geqslant 1}$. Очевидно, что

$$F_n^{(2)}(r_k) \to F(r_k),\ n \to \infty,\ k = 1, 2.$$

Для $k = 1$ это верно в силу того, что $\{F_n^{(2)}\}_{n\geqslant 1}$ является подпоследовательностью последовательности $\{F_n^{(1)}\}_{n\geqslant 1}$, а при $k = 2$ это верно по построению последовательности $\{F_n^{(2)}(r_2)\}_{n\geqslant 1}$.

Продолжая эту процедуру, мы на m-м шаге построим последовательность $\{F_n^{(m)}\}_{n\geqslant 1}$, для которой

$$F_n^{(m)}(r_k) \to F(r_k),\ n \to \infty,\ k = 1, 2, \ldots, m.$$

Для $k = 1, 2, \ldots, m-1$ это верно в силу того, что каждая следующая последовательность является подпоследовательностью предыдущей, а для $k = m$ - по построению $\{F_n^{(m)}(r_m)\}_{n\geqslant 1}$.

Таким образом, на всюду плотном множестве $\{r_k\}_{k\geqslant 1}$ действительной прямой мы определили функцию $F(x),\ x \in \{r_k\}_{k\geqslant 1}$. Очевидно, что для диагональной последовательности $\{F_n^{(n)}\}_{n\geqslant 1}$

$$F_n^{(n)}(r_k) \to F(r_k),\ n \to \infty,\ \text{для любого } k = 1, 2, \dots$$

Так как функции $F_n^{(n)}(x)$ не убывают на всей действительной прямой, то функция $F(x)$ не убывает на своей область определения $\{r_k\}_{k\geqslant 1}$, и мы можем для любого $x \notin \{r_k\}_{k\geqslant 1}$ доопределить функцию F, например следующим образом

$$F(x) = \lim_{k\to\infty} F(r'_k),$$

где $\{r'_k\}_{k\geqslant 1}$ произвольная монотонно возрастающая подпоследовательность последовательности $\{r_k\}_{k\geqslant 1}$, стремящаяся к x. В силу монотонности функции $F(x)$ на множестве $\{r_k\}_{k\geqslant 1}$ последний предел не зависит от подпоследовательности $\{r'_k\}_{k\geqslant 1}$.

Таким образом, мы определили на всей действительной прямой некоторую неубывающую функцию $F(x)$. Пусть x - произвольное действительное число и подпоследовательности $\{r'_k\}_{k\geqslant 1}$ и $\{r''_k\}_{k\geqslant 1}$ последовательности $\{r_k\}_{k\geqslant 1}$ таковы, что $r'_k \uparrow x,\ r''_k \downarrow x,\ k \to \infty$. Тогда для любых n и k

$$F_n^{(n)}(r'_k) \leqslant F_n^{(n)}(x) \leqslant F_n^{(n)}(r''_k).$$

В правой и в левой частях этой цепочки неравенств мы можем переходить к пределу по $n \to \infty$. Мы не можем переходить к пределу по n в средней части этой цепочки, поскольку он может не существовать для некоторых x, но к верхнему и к нижнему пределу переходить можно всегда. Зафиксируем число k и перейдем в правом неравенстве цепочки к верхнему пределу по $n \to \infty$, а в левом - к нижнему. Так мы получим цепочку равенств и неравенств

$$F(r'_k) = \lim_{n\to\infty} F_n^{(n)}(r'_k) = \varliminf_{n\to\infty} F_n^{(n)}(r_k) \leqslant \varliminf_{n\to\infty} F_n^{(n)}(x) \leqslant \varlimsup_{n\to\infty} F_n^{(n)}(x) \leqslant$$

$$\leqslant \varlimsup_{n\to\infty} F_n^{(n)}(r''_k) = \lim_{n\to\infty} F_n^{(n)}(r''_k) = F(r''_k).$$

Теперь, переходя к пределу по $k \to \infty$ мы видим, что

$$F(x-0) \leqslant \varliminf_{n\to\infty} F_n^{(n)}(x) \leqslant \varlimsup_{n\to\infty} F_n^{(n)}(x) \leqslant F(x+0).$$

Если x - точка непрерывности предельной функции, то $F(x-0) = F(x+0) = F(x)$, но это означает, что существует $\lim F_n^{(n)}(x)$, который равен $F(x)$.

Для завершения доказательства осталось заметить, что определенная нами функция $F(x)$ не обязана быть непрерывной слева в точках разрыва, но в этих точках мы можем ее переопределить так, чтобы она стала непрерывной слева.

Компактностью множества $\overline{\mathcal{P}}$ в смысле сходимости в основном мы неоднократно воспользуемся в дальнейшем, а сейчас заметим, что в подавляющем большинстве задач теории вероятностей утверждение о том, что последовательность функций распределения сходится к функции, которая не является функцией распределения, бессодержательны. Поэтому понятие сходимости в основном целесообразно усилить с тем, чтобы получить тот вид сходимости, который играет фундаментальную роль в теории вероятностей.

Определение 3. Говорят, что последовательность $\{F_n\}_{n\geqslant 1} \subset \overline{\mathcal{P}}$ *слабо сходится* к функции $F \in \overline{\mathcal{P}}$, если

1. $F_n(x) \xrightarrow{\text{осн.}} F(x),\ n \to \infty$,

2. $F_n(-\infty) \to F(-\infty),\ F_n(\infty) \to F(\infty),\ n \to \infty$.

Эту сходимость мы будем обозначать $F_n \Rightarrow F,\ n \to \infty$.

Таким образом, в определении слабой сходимости мы дополнительно требуем, чтобы предельные значения функций F_n сходились предельным значениям предельной функции.

Отметим, что если в определении 3 условия принадлежности множеству $\overline{\mathcal{P}}$ заменить на условия принадлежности множеству $\mathcal{P}$, то полученное определение будет совпадать с определением сходимости в основном (для функций распределения), поскольку условие 2. окажется выполненным автоматически, то есть, если мы рассматриваем т о л ь к о функции распределения, то понятия слабой сходимости и сходимости в основном совпадают.

Замечание. Различать понятия сходимости в основном и сходимости $\Rightarrow$ нужно только в том случае, когда предельной функции распределения для последовательности $\{F_n\}_{n\geqslant 1}$ функций распределения не существует или неясно, существует предельная функция распределения или нет. В этом случае обращение к множеству $\overline{\mathcal{P}}$ и поиск предельных функций для последовательности $\{F_n\}_{n\geqslant 1}$ или для ее подпоследовательностей может оказаться очень полезным.

Рассмотрим вопрос о слабой компактности семейств функций распределения (слово "слабой" означает, что речь идет о слабой сходимости). Простейший нетривиальный пример слабо компактного множества функций распределения дает множество $\mathcal{P}_{a,b}$, -$\infty < a < b < \infty$, состоящее из функций распределения $F(x)$, для которых $F(a) = 0$

и $F(b+0) = 1$, то есть носители этих функций распределения принадлежат отрезку $[a,b]$. (Напомним, что носителем функции распределения F (он обозначается $\mathrm{supp}F$) называется множество точек роста функции распределения F, то есть множество точек x, таких, что $F(x+\varepsilon) - F(x-\varepsilon) > 0$ для любого $\varepsilon > 0$.) Действительно, пусть $\{F_n\}_{n\geqslant 1} \subset \mathcal{P}_{a,b}$. Так как $\mathcal{P}_{a,b} \subset \overline{\mathcal{P}}$, а множество $\overline{\mathcal{P}}$ компактно в смысле сходимости в основном, то из последовательности $\{F_n\}_{n\geqslant 1}$ можно выделить подпоследовательность $\{F_{n_k}\}_{k\geqslant 1}$ и указать функцию $F \in \overline{\mathcal{P}}$ такие, что $F_{n_k} \xrightarrow{\text{осн.}} F,\ k \to \infty$. Но так как для всех функций $F_{n_k}(x)$ величины $F_{n_k}(x) = 0$ при всех $x \leqslant a$ и величины $F_{n_k}(x) = 1$ при всех $x > b$, то $F(x) = 0$ для всех точек непрерывности x функции F, лежащих левее a и $F(x) = 1$ во всех точках x непрерывности функции F, лежащих правее b. Поэтому предельная функция F принадлежит $\mathcal{P}_{a,b}$.

Далее нас будут интересовать слабо относительно компактные семейства $\mathcal{H}$ функций распределения.

Определение 4. Множество $\mathcal{H} \subset \mathcal{P}$ называется *слабо относительно компактным*, если из любой последовательности $\{F_n\}_{n\geqslant 1} \subset \mathcal{H}$ можно выделить подпоследовательность, слабо сходящуюся к функции распределения.

Мы не требуем, чтобы предельная функция распределения принадлежала множеству $\mathcal{H}$ (поэтому мы и говорим об относительной компактности).

Нетрудно установить критерий слабой относительной компактности. В этом критерии используются функции

$$h^-(x;\mathcal{H}) = \sup\{F(-x) : F \in \mathcal{H}\}, h^+(x;\mathcal{H}) = \sup\{1-F(x) : F \in \mathcal{H}\}, x > 0,$$

и

$$h(x;\mathcal{H}) = h^-(x;\mathcal{H}) + h^+(x;\mathcal{H}),\ x > 0.$$

Ясно, что функции h^- и h^+ не возрастают, функция $h^-(-x;\mathcal{H})$ является мажорантой левых хвостов $F(x), x < 0$, для всех функций F из множества $\mathcal{H}$, а функция $h^+(x;\mathcal{H})$ является мажорантой правых хвостов $1 - F(x), x > 0$, для всех функций F из множества $\mathcal{H}$

Очевидно, что если множество $\mathcal{H}$ состоит из конечного числа функций (такие множества заведомо компактны), то $h(x;\mathcal{H}) \to 0,\ x \to \infty$. Понятно также, что семейство $\{E_n\}_{n=-\infty}^{\infty}$ является некомпактным и для него $h(x;\mathcal{H}) \equiv 2$. Оказывается, что по поведению функции $h(x;\mathcal{H})$ можно судить о том, является семейство $\mathcal{H}$ слабо относительно компактным или нет.

Теорема 2. *Множество $\mathcal{H} \subset \mathcal{P}$ является слабо относительно компактным тогда и только тогда, когда $h(x;\mathcal{H}) \to 0,\ x \to \infty$.*

Д о к а з а т е л ь с т в о. Пусть $h(x;\mathcal{H}) \to 0,\ x \to \infty$, и пусть $\{F_n\}_{n\geqslant 1}$ - произвольная последовательность функций распределения из $\mathcal{H}$. Так как $\{F_n\}_{n\geqslant 1} \subset \overline{\mathcal{P}}$, а множество $\overline{\mathcal{P}}$ компактно в смысле сходимости в основном, то из этой последовательности можно выделить подпоследовательность $\{F_{n_k}\}_{k\geqslant 1}$ и указать функцию $F \subset \overline{\mathcal{P}}$, такие, что $F_{n_k} \xrightarrow{\text{осн.}} F,\ n \to \infty$. Очевидно, что для каждой точки $x > 0, x \in \mathbb{C}(\mathbb{F})$,

$$1 - F(x) = \lim_{k\to\infty}(1 - F_{n_k}(x)) \leqslant h^+(x;\mathcal{H}) \leqslant h(x;\mathcal{H}),$$

и если $h(x;\mathcal{H}) \to 0,\ x \to \infty$, то $F(\infty) = 1$. Аналогично доказывается, что $F(-\infty) = 0$, то есть $F \in \mathcal{P}$.

Таким образом, если $h(\infty;\mathcal{H})$, то из любой последовательности функций распределения $\{F_n\}_{n\geqslant 1} \subset \mathcal{H}$ можно выделить подпоследовательность $\{F_{n_k}\}_{k\geqslant 1}$, сходящуюся в основном, а потому и слабо к функции распределения F, то есть множество $\mathcal{H}$ слабо относительно компактно.

Покажем теперь, что если $h(x;\mathcal{H}) \not\to 0$ при $x \to \infty$, то в множестве $\mathcal{H}$ существует последовательность, сходящаяся в основном к некоторой функции, не являющейся функцией распределения, поэтому из этой последовательности нельзя извлечь подпоследовательность, слабо сходящуюся к функции распределения.

Пусть множество $\mathcal{H} \subset \mathcal{P}$ таково, что $h(x;\mathcal{H}) \not\to 0, x \to \infty$. В силу монотонности функций $h^-(x;\mathcal{H})$ и $h^+(x;\mathcal{H})$ это означает существование $\varepsilon > 0$ такого, что верно хотя бы одно из двух неравенств $h^-(\infty;\mathcal{H}) > \varepsilon$ или $h^+(\infty;\mathcal{H}) > \varepsilon$. Пусть, для определенности $h^+(+\infty;\mathcal{H}) > \varepsilon$. В силу монотонности функции $h^+(x;\mathcal{H})$ для любого натурального n

$$h^+(n;\mathcal{H}) = \sup\{1 - F(n) : F \in \mathcal{H}\} > \varepsilon.$$

По свойству точной верхней грани, для любого n в $\mathcal{H}$ существует функция F, для которой $1-F(n) > \varepsilon$. Обозначим F_n любую из этих функций. Нетрудно понять, что в силу монотонности функций $F_n(x)$ для любого $x > 0$

$$1 - F_n(x) > \varepsilon$$

для всех n, начиная с некоторого (заведомо для всех $n > x$).

Так как $\{F_n\}_{n\geqslant 1} \subset \overline{\mathcal{P}}$, то в $\{F_n\}_{n\geqslant 1}$ существует подпоследовательность $\{F_{n_k}\}_{k\geqslant 1}$ и функция $F \in \overline{\mathcal{P}}$, такие, что $F_{n_k} \xrightarrow{\text{осн.}} F,\ k \to \infty$. Для любой фиксированной точки $x > 0,\ x \in \mathbb{C}(F)$,

$$1 - F(x) = \lim_{k\to\infty}(1 - F_{n_k}(x)) \geqslant \varepsilon.$$

Отсюда следует, что $F(\infty) \leqslant 1 - \varepsilon < 1$, то есть $F \notin \mathcal{P}$. Очевидно, что из последовательности $\{F_{n_k}\}_{k\geqslant 1}$ нельзя извлечь подпоследовательность, слабо сходящуюся к функции распределения. Если бы такая подпоследовательность существовала и слабо сходилась к функции $G \in \mathcal{P}$, то она сходилась бы в основном как к F, так и к G, что невозможно.

Случай $h^-(\infty; \mathcal{H}) > 0$ рассматривается аналогично.

Теорема доказана.

Метризуемость слабой сходимости. Метрика Леви.

Определение 5. *Расстоянием (метрикой) Леви* между функциями $F, G \in \overline{\mathcal{P}}$ называется число

$$L(F,G) = \inf\{\varepsilon > 0 : F(x-\varepsilon)-\varepsilon \leqslant G(x) \leqslant F(x+\varepsilon)+\varepsilon \text{ для всех} -\infty < x < \infty\}.$$

Легко видеть, что для данного $\varepsilon > 0$ график функции, стоящей в правой части правого неравенства из определения $L(F, G)$ получится, если мы сдвинем график функции $F(x)$ влево на величину ε и поднимем его на величину ε, а график функции, стоящей в левой части левого неравенства можно получить, сдвинув график функции $F(x)$ вправо на величину ε и опустив его на величину ε. Таким образом в множестве $\overline{\mathcal{P}}$ получается некая ε-окрестность функции F и, грубо говоря, величина $L(F, G)$ - это минимальное ε, при котором график $G(x)$ лежит в ε-окрестности функции F. (При $\varepsilon = 1$ все множество $\overline{\mathcal{P}}$ лежит в ε-окрестности функции F. То есть множество чисел $\varepsilon > 0$, нижняя грань которого есть $L(F, G)$, непусто.)

Нетрудно понять, что расстояние Леви между функциями F и G совпадает со стороной максимального квадрата, который можно вставить между замкнутыми графиками функций F и G , то есть между графиками, к которым добавляются вертикальные отрезки, соединяющие предельные значения функции в ее точках разрыва. Далее, говоря о графиках функций распределения мы всегда имеем в виду замкнутые графики. Здесь и далее, говоря о квадратах, которые можно вставить между графиками функций F и G , мы рассматриваем квадраты со сторонами, параллельными осям координат.

Легко проверить, что $L(F, G)$ является метрикой.

Действительно, величина $L(F, G)$ неотрицательна. Ясно, что если $F(x) = G(x)$, то для любого ε справедливы неравенства

$$F(x-\varepsilon) - \varepsilon \leqslant G(x) \leqslant F(x+\varepsilon) + \varepsilon.$$

и точная нижняя грань таких чисел ε равна нулю, то есть $L(F, G) = 0$. Обратно, если $L(F, G) = 0$, то по свойству точной нижней грани существует последовательность $\{\varepsilon_n\}_{n \geqslant 1}$ такая, что $\varepsilon_n \to 0$ при $n \to \infty$ и для всех действительных x

$$F(x-\varepsilon_n)-\varepsilon_n \leqslant G(x) \leqslant F(x+\varepsilon_n)+\varepsilon_n.$$

Переходя в этих неравенствах к пределу при $n \to \infty$ при любом фиксированном $x \in \mathbb{C}(F)$, мы видим, что $F(x) = G(x)$ при $x \in \mathbb{C}(F)$ Отсюда следует совпадение $F(x)$ и $G(x)$ при всех действительных x.

Ясно, что для любого $\varepsilon > L(F,G)$ при всех действительных x справедливо неравенство $G(x) \leqslant F(x+\varepsilon)+\varepsilon$ то есть $G(x)-\varepsilon \leqslant F(x+\varepsilon)$. Положив $y = x+\varepsilon$, мы видим, что $G(y-\varepsilon)-\varepsilon \leqslant F(y)$ при всех действительных y. Аналогично получается неравенство $F(y) \leqslant G(y+\varepsilon)+\varepsilon$ при всех действительных y. Отсюда легко получается равенство $L(F,G) = L(G,F)$.

Пусть $F, G, H \in \overline{\mathcal{P}}$ и числа $\varepsilon_1, \varepsilon_2$ таковы, что $\varepsilon_1 > L(F,H),\ \varepsilon_2 > L(H,G)$, тогда для всех действительных x

$$F(x) \leqslant H(x+\varepsilon_1)+\varepsilon_1 \leqslant (G(x+\varepsilon_1+\varepsilon_2)+\varepsilon_2)+\varepsilon_1 \leqslant G(x+\varepsilon_1+\varepsilon_2)+\varepsilon_1+\varepsilon_2.$$

Теперь понятно, что для всех x

$$G(x-\varepsilon_1-\varepsilon_2)-\varepsilon_1-\varepsilon_2 \leqslant F(x) \leqslant G(x+\varepsilon_1+\varepsilon_2)+\varepsilon_1+\varepsilon_2.$$

Отсюда следует, что $L(F,G) \leqslant \varepsilon_1+\varepsilon_2$ и переходя к пределу при $\varepsilon_1 \downarrow L(F,H)$ и $\varepsilon_2 \downarrow L(H,G)$ мы приходим к неравенству $L(F,G) \leqslant L(F,H)+L(H,G)$.

Таким образом, $L(F,g)$ действительно является метрикой.

Теорема 3. *На множестве $\overline{\mathcal{P}}$ слабая сходимость и сходимость в метрике Леви эквивалентны.*

Д о к а з а т е л ь с т в о. Пусть $\{F_n\}_{n\geqslant 1} \subset \overline{\mathcal{P}}, F \in \overline{\mathcal{P}}$ и $\varepsilon_n = L(F_n,F) \to 0$ при $n \to \infty$. Ясно, что для всех действительных x и всех натуральных n

$$F(x-2\varepsilon_n)-2\varepsilon_n \leqslant F_n(x) \leqslant F(x+2\varepsilon_n)+2\varepsilon_n \tag{1}$$

Так как функции $F(x)$ и $F_n(x)$ монотонны и ограничены, мы можем переходить к пределу по $x \to \infty$, что дает неравенства

$$F(\infty)-2\varepsilon_n \leqslant F_n(\infty) \leqslant F(\infty)+2\varepsilon_n$$

Отсюда следует, что $F_n(\infty) \to F(\infty)$ при $n \to \infty$. Аналогично доказывается, что $F_n(-\infty) \to F(-\infty)$ при $n \to \infty$.

Перейдем теперь в неравенствах (1) к пределу при $n \to \infty$. Переход к такому пределу для левой и правой частей цепочки (1) затруднений не вызывает, так как функция F монотонна и ограничена. Переход к пределу в средней части цепочки (1) нужно проводить аккуратнее, так как этот предел может не существовать для некоторых чисел x, однако, нижний и верхний пределы существуют всегда. Переходя при фиксированном x в левой части (1) к нижнему пределу, а в правой - к верхнему, и учитывая, что нижний предел не превосходит верхнего, мы получим

$$F(x-0) \leqslant \varliminf_{n\to\infty} F_n(x) \leqslant \varlimsup_{n\to\infty} F_n(x) \leqslant F(x+0).$$

Отсюда следует, что

$$F_n(x) \to F(x),\ n \to \infty,\ \text{для любого } x \in \mathbb{C}(F).$$

Таким образом, из сходимости $L(F_n, F) \to 0$ следует $F_n \Rightarrow F$ при $n \to \infty$.

Для того, чтобы доказать, что из $F_n \Rightarrow F$, $n \to \infty$, следует $L(F_n, F) \to 0$, предположим противное, то есть предположим, что существует последовательность $\{F_n\}_{n\geqslant 1} \subset \overline{\mathcal{P}}$ и функция $F \in \overline{\mathcal{P}}$ такие, что

$$F_n \Rightarrow F,\ \text{но } L(F_n, F) \not\to 0 \text{ при } n \to \infty. \tag{2}$$

Далее мы неоднократно будем выделять из последовательности $\{F_n\}_{n\geqslant 1}$ подпоследовательности, обладающие дополнительными свойствами и поскольку речь идет о существовании нужной последовательности, мы всегда можем считать, что сама последовательность $\{F_n\}_{n\geqslant 1}$ обладает этими дополнительными свойствами.

Итак, пусть выполнено (2), тогда существует подпоследовательность $\{F_{n_k}\}_{k\geqslant 1}$ и число $\varepsilon > 0$ такие, что $F_{n_k} \Rightarrow F$, $k \to \infty$, и $L(F_{n_k}, F) \geqslant \varepsilon$ при всех $k \geqslant 1$ и мы можем считать, что сама последовательность $\{F_n\}_{n\geqslant 1}$ такова, что $L(F_n, F) \geqslant \varepsilon$ при всех $n \geqslant 1$, то есть между графиками функций F_n и F можно вставить квадрат $K_\varepsilon(x_n, y_n)$ со стороной ε, где (x_n, y_n) - координаты центра этого квадрата. Ясно, что $|F_n(x_n) - F(x_n)| \geqslant \varepsilon$ при всех натуральных n.

Так как функция $F \in \overline{\mathcal{P}}$, то существует такое число α, которое является точкой непрерывности $F(x)$, что $a = F(-\infty) \leqslant a + \varepsilon/3$ при всех $x \leqslant \alpha$, то есть при $x \leqslant \alpha$ график функции $F(x)$ лежит в горизонтальной полосе $\{(x, y) : x \leqslant \alpha, a \leqslant y \leqslant a + \varepsilon/3\}$ ширины $\varepsilon/3$.

Так как $F_n(-\infty) \to F(-\infty) = a$ и $F_n(\alpha) \to F(\alpha), n \to \infty$, поскольку $\alpha \in \mathbb{C}(F)$, то начиная с некоторого n

$$a - \varepsilon/6 \leqslant F_n(-\infty) \leqslant a + \varepsilon/3$$

и

$$a - \varepsilon/6 \leqslant F_n(x) \leqslant a + \varepsilon/3 + \varepsilon/6 \text{ при всех } x \leqslant \alpha.$$

Таким образом, начиная с некоторого n графики всех функций F_n, F будут лежать в горизонтальной полосе $\{(x,y): x \leqslant \alpha, a-\frac{\varepsilon}{6} \leqslant y \leqslant a+\frac{\varepsilon}{2}\}$ ширины $\frac{2\varepsilon}{3}$. В эту полосу нельзя вставить квадрат со стороной ε, поэтому для всех n, начиная с некоторого, $x_n \geqslant \alpha$. Таким образом, последовательность $\{x_n\}_{n\geqslant 1}$ ограничена снизу. Аналогично доказывается, что она ограничена сверху. Ограниченность последовательности $\{y_n\}_{n\geqslant 1}$ очевидна. Поскольку из ограниченной последовательности можно извлечь сходящуюся подпоследовательность, то мы можем считать, что наша последовательность $\{F_n\}_{n\geqslant 1}$ такова, что $F_n \Rightarrow F, n \to \infty$, между графиками функций F_n и F можно вставить квадрат $K_\varepsilon(x_n, y_n)$ со стороной ε и $x_n \to x_0, y_n \to y_0$ при $n \to \infty$. Заметим, наконец, что начиная с некоторого n во все квадраты $K_\varepsilon(x_n, y_n)$ со стороной ε можно вставить один и тот же квадрат $K_{\varepsilon/2}(x_0, y_0)$ со стороной $\varepsilon/2$ и центром (x_0, y_0).

Очевидно, что для всех $x \in \left(x_0 - \frac{\varepsilon}{4}, x_0 + \frac{\varepsilon}{4}\right)$ при достаточно больших n справедливо неравенство $|F_n(x) - F(x)| \geqslant \varepsilon/2$, что противоречит условию $F_n \Rightarrow F, n \to \infty$, поскольку на указанном интервале существуют точки непрерывности функции F.

Теорема доказана.

Стоит отметить, что в литературе вместо терминов "сходимость в основном" и "слабая сходимость" иногда используются термины "слабая сходимость" и "сходимость вполне" соответственно.

Связь между метрикой Леви и метриками ρ и κ .

Далее мы рассматриваем расстояния $L(F,G), \rho(F,G)$ и $\kappa(F,G)$ только между функциями распределения.

Определение 6. Пусть на множестве $\mathcal{V}$ заданы метрики μ и ν. Говорят, что метрика μ слабее или эквивалентна метрике ν (ν сильнее или эквивалентна метрике μ), если из $\nu(x_n, x) \to 0$ при $n \to \infty$, $\{x_n\}_{n\geqslant 1} \subset \mathcal{V}, x \in \mathcal{V}$, следует $\mu(x_n, x) \to 0$. Это соотношение между метриками обозначается $\mu \precsim \nu$ ($\nu \succsim \mu$). Говорят, что метрика μ строго слабее ν (ν строго сильнее μ), если $\mu \precsim \nu$ и в $\mathcal{V}$ существует последовательность $\{y_n\}_{n\geqslant 1}$ и точка $y \in \mathcal{V}$, такие, что $\mu(y_n, y) \to 0$, но $\nu(y_n, y) \not\to 0$ при $n \to \infty$. Это соотношение между метриками обозначается $\mu \prec \nu$ ($\nu \succ \mu$). Если $\mu \precsim \nu$ и $\nu \precsim \mu$, то метрики μ и ν называются эквивалентными (обозначение $\mu \sim \nu$).

Отметим, что не всякие метрики сравнимы.

Рассмотрим вначале связи между метриками L и ρ . Так как расстояние Леви $L(F,G) = \varepsilon$ между функциями распределения совпадает с максимальной стороной квадрата, который можно вставить между графиками этих функций, а для точек x оси абсцисс, находящихся под этим квадратом, $|F(x) - G(x)| \geqslant \varepsilon$, то очевидно, что $L(F,G) \leqslant \rho(F,G)$. Поэтому из сходимости $\rho(F_n, F) \to 0$, $n \to \infty$, следует сходимость $L(F_n, F) \to 0$. Простой пример функций распределения $F(x) = E_0(x)$ и $F_n(x) = E_{1/n}(x), n = 1, 2, \ldots$, показывает, что обратное неверно, поскольку для указанных функций распределения $\rho(F_n, F) \equiv 1$, а $L(F_n, F) = \frac{1}{n} \to 0$ при $n \to \infty$.

Однако, в некоторых случаях из сходимости в метрике Леви следует равномерная сходимость. Важнейший из них - случай, когда предельная функция непрерывна. Дело в том, что если хотя бы одна из функций распределения F и G , скажем F, непрерывна, величину $\rho(F,G)$ можно оценить сверху через $L(F,G)$. Это очевидно для дифференцируемых функций F таких, что $F'(x) \leqslant A$ при всех x.

Действительно, для любого $\varepsilon > L(F,G)$ при всех x выполнены неравенства

$$F(x-\varepsilon) - \varepsilon \leqslant G(x) \leqslant F(x+\varepsilon) + \varepsilon,$$

а если $F'(x) \leqslant A$, то $F(x+\varepsilon) \leqslant F(x) + A\varepsilon$ и $F(x) \leqslant F(x-\varepsilon) + A\varepsilon$. Из этих неравенств следует, что при всех x

$$F(x) - A\varepsilon - \varepsilon \leqslant G(x) \leqslant F(x) + A\varepsilon + \varepsilon,$$

а отсюда следует, что $\rho(F,G) \leqslant (1 + A)\varepsilon$. Переход к пределу при $\varepsilon \downarrow L(F,G)$ дает неравенство

$$\rho(F,G) \leqslant (1 + A)L(F,G).$$

В случае, когда на F налагается лишь условие непрерывности, доказательство аналогичного неравенства чуть сложнее и оно опирается на тот факт, что всякая непрерывная функция распределения равномерно непрерывна на всей действительной оси. Для того, чтобы доказать эту равномерную непрерывность, предположим противное и отрицание равномерной непрерывности функции F на всей действительной оси запишем в виде: существует число $\varepsilon > 0$ и последовательности $\{x'_n\}_{n\geqslant 1}$ и $\{x''_n\}_{n\geqslant 1}$ такие, что $|x'_n - x''_n| \leqslant 1/n$, но $|F(x'_n) - F(x''_n)| \geqslant \varepsilon$. Так как $F(x)$ является функцией распределения, то существуют число a такое, что $F(x) < \varepsilon$ при всех $x < a$ и число $b > a$, такое, что $F(x) > 1 - \varepsilon$ при всех $x > b$. Ясно, что оба числа x'_n и x''_n не могут лежать левее a или правее b. поэтому хотя бы одно из них принадлежит отрезку $[a, b]$, и поэтому все пары чисел x'_n, x''_n принадлежат отрезку $[a - 1, b + 1]$ Мы получили отрицание равномерной непрерывности непрерывной функции F на компакте $[a - 1, b + 1]$, что неверно.

Для всякой функции $F(x), -\infty < x < \infty$, можно ввести функцию

$$W_F(\delta) = \sup\{|F(x) - F(y)| : |x - y| \leqslant \delta\},\ \delta > 0,$$

которая называется ее модулем непрерывности. Для равномерно непрерывной функции F ее модуль непрерывности $W_F(\delta) \to 0$ при $\delta \to 0$.

Вернемся к оценке величины $\rho(F, G)$ через $L(F, G)$ в случае, когда функция распределения F непрерывна, а потому и равномерно непрерывна на всей действительной оси. Действуя точно так же, как и ранее, и заменяя лишь неравенства $F(x+\varepsilon) \leqslant F(x)+A\varepsilon$ и $F(x) \leqslant F(x-\varepsilon)+A\varepsilon$ на неравенства $F(x+\varepsilon) \leqslant F(x)+W_F(\varepsilon)$ и $F(x) \leqslant F(x-\varepsilon)+W_F(\varepsilon)$, мы убеждаемся в справедливости оценки

$$\rho(F, G) \leqslant L(F, G) + W_F(L(F, G)).$$

Поэтому, если $L(F_n, F) \to 0$ при $n \to \infty$, а предельная функция F непрерывна, то $\rho(F_n, F) \to 0$ то есть из сходимости $F_n \Rightarrow F$ следует сходимость $\rho(F_n, F) \to 0$ при $n \to \infty$, если предельная функция непрерывна.

Рассмотрим теперь связи между метриками L и κ.

Прежде всего заметим, что расстояние $\kappa(F, G) = \int\limits_{-\infty}^{\infty} |F(x) - G(x)| dx$ совпадает с площадью геометрической фигуры, состоящей из точек, лежащих между замкнутыми графиками функций распределения F и G. Расстояние Леви $L(F, G)$ совпадает со стороной максимального квадрата, который можно вставить между этими графиками, площадь этого квадрата есть $L^2(F, G)$, поэтому $L^2(F, G) \leqslant \kappa(F, G)$. Из этого неравенства следует, что из $\kappa(F_n, F) \to 0,\ n \to \infty$, вытекает, что $L(F_n, F) \to 0$. Обратное неверно, что показывает пример функций распределения $F(x) = E_0(x)$ и $F_n(x) = \left(1 - \frac{1}{n}\right) E_0(x) + \frac{1}{n} E_n(x)$. Для этих функций распределения расстояние $\kappa(F_n, F) \equiv 1$, но $L(F_n, F) = \frac{1}{n} \to 0,\ n \to \infty$. Таким образом, метрика Леви строго слабее метрики κ.

Оказывается, что в некоторых случаях из $L(F_n, F) \to 0,\ n \to \infty$, следует $\kappa(F_n, F) \to 0$. Один из них - случай, когда носители функций F_n принадлежат некоторому конечному отрезку $[a, b]$ (выше мы видели, что если $L(F_n, F) \to 0,\ n \to \infty$, и $\{F_n\}_{n \geqslant 1} \subset \mathcal{P}_{a,b}$, то $F \in \mathcal{P}_{a,b}$). Это вытекает из следующего утверждения.

Теорема 4. *Пусть $F, G \in \mathcal{P}_{a,b}$. Тогда*

$$\kappa(F, G) \leqslant (1 + b - a) L(F, G).$$

Для доказательства этого утверждения нам понадобится следующая

Лемма. *Пусть $F, G \in \mathcal{P}_{a,b}$ - непрерывные и строго монотонные на $[a, b]$ функции распределения. Тогда существует непрерывная и строго монотонная на $[a, b]$ функция $H \in \mathcal{P}_{a,b}$, такая, что*

$$|F(x) - H(x)| \leqslant \varepsilon,\ x \in (a, b),\ |H^{-1}(y) - G^{-1}(y)| \leqslant \varepsilon,\ y \in (0, 1),$$

где $\varepsilon = L(F, G)$, а H^{-1} и G^{-1} - функции, обратные H и G.

Доказательство. Так как функция $|F(x)-G(x)|$ непрерывна, то множество $\{x : |F(x) - H(x)| > 0\}$ является открытым подмножеством интервала (a, b), и поэтому оно состоит из не более, чем счетного множества интервалов (α, β) и нам достаточно построить функцию $H(x)$ на каждом из этих интервалов. Пусть $F(x) > G(x)$ на некотором интервале (α, β). Для каждого $x \in (\alpha, \beta)$ построим квадрат, левый верхний угол, абсцисса которого есть x, лежит на графике функции F, а правый нижний угол лежит на графике функции G. При увеличении x от α до β этот квадрат будет перемещаться, оставаясь между графиками F и G, его сторона будет изменяться, оставаясь не больше ε и очевидно, что его левый нижний угол прочертит график нужной нам функции $H(x)$ при $x \in (\alpha, \beta)$. На интервалах, где $F(x) < G(x)$, функция $H(x)$ определяется аналогично. Для тех x, где $F(x) = G(x)$, положим $H(x) = F(x)$.

Функция $H(x) \in \mathcal{P}_{a,b}$ построена, ее непрерывность и строгая монотонность на $[a, b]$ очевидны.

Доказательство теоремы 4. Пусть $F, G \subset \mathcal{P}_{a,b}$ - непрерывные строго монотонные на $[a, b]$ функции распределения и $\varepsilon = L(F, G)$. Воспользовавшись функцией H, существование которой гарантируется доказанной леммой, мы видим, что

$$\kappa(F, G) = \int_a^b |F(x)-G(x)|dx \leqslant \int_a^b |F(x)-H(x)|dx+\int_a^b |H(x)-G(x)|dx =$$

$$= \int_a^b |F(x) - H(x)|dx + \int_0^1 |H^{-1}(y) - G^{-1}(y)|dy \leqslant (b - a)\varepsilon + \varepsilon.$$

Последнее равенство в этой цепочке справедливо в силу того, что вторые слагаемые в его левой и правой части совпадают, поскольку они равны площади фигуры, состоящей из точек, лежащих между графиками функций H и G.

Для того, чтобы доказать неравенство теоремы 4 в общем случае, воспользуемся тем, что множество непрерывных строго монотонных функций из $\mathcal{P}_{a,b}$ плотно в средней метрике во всем множестве $\mathcal{P}_{a,b}$ (а в силу неравенства $L^2 \leqslant \kappa$ оно плотно и в метрике Леви), то есть для любой функции $F \in \mathcal{P}_{a,b}$ существует последовательность $\{F_n\}_{n\geqslant 1}$ непрерывных строго монотонных на $[a,b]$ функций, такая, что $\kappa(F_n, F) \to 0,\ n \to \infty$.

Действительно, пусть $F \in \mathcal{P}_{a,b}$ и множество ее точек разрыва счетно. Зафиксируем малое число $\varepsilon > 0$ и аппроксимируем $F(x)$ функцией $F^{(1)}(x)$ с конечным числом скачков. Для этого представим функцию $F(x)$ в виде суммы $F_c(x) + F_d(x)$, где функция

$$F_d(x) = \sum_{j:x_j<x} p_j,$$

$x_1, x_2, \dots$ - точки разрыва $F(x)$, а $p_1, p_2, \dots$ - величины скачков $F(x)$ в точках разрыва. Функция $F_d(x)$ изменяется только скачками в точках $x_1, x_2, \dots$, а функция $F_c(x) = F(x) - F_d(x)$ непрерывна. Так как сумма ряда $\sum\limits_{j\geqslant 1} p_j$ не превосходит единицы, то для любого $\varepsilon > 0$ существует число n_ε, такое, что $\sum\limits_{j>n_\varepsilon} p_j < \varepsilon$. Введем функцию $F_d^{(1)}(x)$, которая, как и $F_d(x)$, изменяется только скачками, но число этих скачков конечно (и не превосходит $n_\varepsilon + 1$). На интервале (a,b) мы положим

$$F_d^{(1)}(x) = \sum_{j:x_j<x,\ j\leqslant n_\varepsilon} p_j,$$

а сумму $\sum\limits_{j>n_\varepsilon} p_j$ оставшихся скачков $F_d(x)$ поместим в точку b. (Если функция $F_d(x)$ непрерывна в точке b, то у функции $F_d^{(1)}(x)$ будет скачок величины $\sum\limits_{j>n_\varepsilon} p_j$ в точке b, а если у функции $F_d(x)$ есть скачок величины c в точке b, то у функции $F_d^{(1)}(x)$ в точке b будет скачок величины $c + \sum\limits_{j>n_\varepsilon} p_j$.) Легко понять, что разность $F(x) - F^{(1)}(x)$ между функциями распределения $F(x) = F_c(x) + F_d(x)$ и $F^{(1)}(x) = F_c(x) + F_d^{(1)}(x)$ на интервале (a,b) меняется только скачками величины p_j в точках $x_j, j > n_\varepsilon$, а в точке b эта разность скачком убывает до нуля. Очевидно, что на (a,b) величина $0 \leqslant F(x) - F^{(1)}(x) \leqslant \varepsilon$ и

$$\kappa(F, F^{(1)}) \leqslant \int\limits_a^b |F_d(x) - F_d^{(1)}(x)| dx \leqslant \varepsilon(b-a).$$

Введем теперь функцию $F^{(2)}(x)$, положив $F^{(2)}(x) = 0$ для $x < a + \varepsilon$ и

$F^{(2)}(x) = F^{(1)}(x)$ при $x > a + \varepsilon$. Очевидно, что

$$\kappa(F^{(1)}, F^{(2)}) = \int\limits_a^{a+\varepsilon} F^{(1)}(x)dx \leqslant \varepsilon,$$

число скачков функции распределения $F^{(2)}(x)$ конечно и на интервале $(a, a+\varepsilon)$ скачков у $F^{(2)}(x)$ нет. Действуя так же, как и раньше, представим функцию $F^{(2)}(x)$ в виде суммы $F_c^{(2)}(x) + F_d^{(2)}(x)$ непрерывной функции $F_c^{(2)}$ и функции $F_d^{(2)}$, меняющейся только скачками, величины которых суть $q_1, \ldots, q_k$ и которые находятся в точках $y_1, \ldots, y_k$. Аппроксимируем функцию $F_d^{(2)}$ непрерывной функцией. Для этого выберем число $\varepsilon' > 0$ такое, что $\varepsilon' < \varepsilon$ и $\varepsilon' < \min\limits_{i \neq j} |y_i - y_j|$. Рассмотрим функцию $\mathcal{V}(x)$, которая совпадает с $F_d^{(2)}$ вне интервалов $(y_j - \varepsilon', y_j), j = 1, \ldots, k$, а на этих интервалах линейно возрастает от $F_d^{(2)}(y_j - 0)$ до $F_d^{(2)}(y_j + 0)$. Очевидно, что величина $\kappa(F^{(2)}, F^{(3)})$, где $F^{(3)}(x) = F_c^{(2)}(x) + \mathcal{V}(x)$ равна сумме площадей прямоугольных треугольников с основаниями $[y_j - \varepsilon', y_j]$ и высотами $F_d^{(2)}(y_j + 0) - F_d^{(2)}(y_j - 0) = q_j$. Так как $\sum\limits_{j=1}^{k} q_j \leqslant 1$, то $\kappa(F^{(2)}, F^{(3)}) \leqslant \varepsilon'/2 \leqslant \varepsilon/2$ и очевидно, что функция $F^{(3)}$ непрерывна.

Наконец, рассмотрим функцию распределения $F^{(4)}(x) = (1 - \varepsilon)F^{(3)}(x) + \varepsilon R_{a,b}(x)$, где $R_{a,b}$ - функция распределения, равномерного на (a, b). Очевидно, что функция распределения $F^{(4)} \in \mathcal{P}_{a,b}$ и она является непрерывной и строго возрастающей на $[a, b]$, при этом $\kappa(F^{(3)}, F^{(4)}) \leqslant \varepsilon(b - a)$.

Таким образом, для произвольной функции $F \in \mathcal{P}_{a,b}$ со счетным множеством скачков мы построили непрерывную и строго монотонную на $[a, b]$ функцию $F^{(4)}$ такую, что $\kappa(F, F^{(4)}) \leqslant 2\varepsilon(b - a) + 3\varepsilon/2$.

Из сказанного легко понять, как аппроксимируются функции $F \in \mathcal{P}_{a,b}$ с конечным числом скачков или непрерывные.

Тем самым для средней метрики плотность в множестве $\mathcal{P}_{a,b}$ его подмножества, состоящего из непрерывных и строго монотонных функций, доказана.

Пусть F и G произвольные функции из $\mathcal{P}_{a,b}$ и пусть $\{F_n\}_{n \geqslant 1}$ и $\{G_n\}_{n \geqslant 1}$ - последовательности, состоящие из непрерывных и строго монотонных на $[a, b]$ функций, такие, что $\kappa(F_n, F) \to 0$ и $\kappa(G_n, G) \to 0$, $n \to \infty$.

По неравенству треугольника

$$\kappa(F, G) \leqslant \kappa(F, F_n) + \kappa(F_n, G_n) + \kappa(G_n, G).$$

Второе слагаемое в правой части этого неравенства не превосходит $(1+b-a)L(F_n,G_n)$, а два оставшихся стремятся к нулю при $n\to\infty$.

Величина

$$L(F_n,G_n)\leqslant L(F_n,F)+L(F,G)+L(G,G_n).$$

Так как квадрат расстояния Леви всегда не превосходит среднего расстояния, то $L(F_n,F)\to 0$ и $L(G_n,G)\to 0$, $n\to\infty$. Отсюда следует утверждение теоремы.

Следствие теоремы 3. *Для любых функций распределения F и G и любых чисел $-\infty<a<b<\infty$*

$$\int\limits_a^b |F(x)-G(x)|dx\leqslant(1+b-a)L(F,G).$$

Действительно, функции

$$F_{a,b}(x)=\begin{cases}0, & x\leqslant a,\\ F(x), & a<x\leqslant b,\\ 1, & x>b,\end{cases}\quad \text{и}\quad G_{a,b}(x)=\begin{cases}0, & x\leqslant a,\\ G(x), & a<x\leqslant b,\\ 1, & x>b\end{cases}$$

принадлежат $\mathcal{P}_{a,b}$ и ясно, что

$$\int\limits_a^b |F(x)-G(x)|dx=\kappa(F_{a,b},G_{a,b})$$

В силу теоремы 4

$$\kappa(F_{a,b},G_{a,b})\leqslant(1+b-a)L(F_{a,b},G_{a,b})$$

и очевидно, что $L(F_{a,b},G_{a,b})\leqslant L(F,G)$, поскольку расстояние Леви совпадает со стороной максимального квадрата, который можно вставить между графиками функций распределения, а каждый из квадратов, которые можно вставить между графиками $F_{a,b}$ и $G_{a,b}$ можно вставить и между графиками F и G.

Сходимость $\xrightarrow{w}$. Здесь мы рассмотрим еще один вид сходимости распределений, который также называется слабой сходимостью. Этот вид сходимости мы будем рассматривать только на множестве $\mathcal{P}$ функций распределения, хотя его можно рассматривать и на множестве $\overline{\mathcal{P}}$.

Определение 6. Говорят, что последовательность $\{F_n\}_{n\geqslant 1} \subset \mathcal{P}$ *слабо сходится* к функции распределения $F \in \mathcal{P}$, если для любой непрерывной ограниченной функции f

$$\int\limits_{-\infty}^{\infty} f(x)dF_n(x) \to \int\limits_{-\infty}^{\infty} f(x)dF(x)$$

при $n \to \infty$. Этот вид сходимости мы будем обозначать $F_n(x) \xrightarrow{w} F(x),\ n \to \infty$.

Теорема 5. *Сходимости* $\Rightarrow$ *и* $\xrightarrow{w}$ *эквивалентны.*

Д о к а з а т е л ь с т в о. Прежде всего заметим, что если $F_n(x) \xrightarrow{w} F(x),\ n \to \infty$, то семейство $\{F_n\}_{n\geqslant 1}$ слабо относительно компактно. Пусть $a < 0$ - произвольное число, которое является точкой непрерывности всех функций $F, \{F_n\}_{n\geqslant 1}$ и пусть функция $f(x) = 1$ при $x < a$, $f(x)$ линейно убывает от 1 до 0, когда x возрастает от a до $a+1$ и $f(x) = 0$ при $x > a+1$. Ясно, что для любого n

$$F_n(x) = \int\limits_{-\infty}^{a} dF_n(x) \leqslant \int\limits_{-\infty}^{\infty} f(x)dF_n(x) =$$

$$= \int\limits_{-\infty}^{\infty} f(x)dF_n(x) - \int\limits_{-\infty}^{\infty} f(x)dF(x) + \int\limits_{-\infty}^{\infty} f(x)dF(x) =$$

$$= \int\limits_{-\infty}^{\infty} f(x)dF_n(x) - \int\limits_{-\infty}^{\infty} f(x)dF(x) + \int\limits_{-\infty}^{a+1} dF(x) \leqslant$$

$$\leqslant \int\limits_{-\infty}^{\infty} f(x)dF_n(x) - \int\limits_{-\infty}^{\infty} f(x)dF(x) + F(a+1).$$

В силу того, что F является функцией распределения, для любого $\varepsilon > 0$ можно найти такое a с указанным выше свойством, что $F(a+1) < \varepsilon/2$. Разность двух первых слагаемых в правой части последней цепочки равенств и неравенств станет меньше $\varepsilon/2$ начиная с некоторого n в силу того, что $F_n(x) \xrightarrow{w} F(x),\ n \to \infty$, и f - непрерывная ограниченная функция. Таким образом, $F_n(a) < \varepsilon$ начиная с некоторого n и понятно, что число a можно уменьшить так, чтобы неравенство $F_n(a) < \varepsilon$ стало справедливо для всех $n \geqslant 1$. Отсюда следует, что $h^-(x; \{F_n\}_{n\geqslant 1}) \to 0,\ n \to \infty$, аналогично доказывается, что

$h^+(x;\{F_n\}_{n\geqslant 1}) \to 0,\ n \to \infty$, а отсюда вытекает слабая относительная компактность семейства $\{F_n\}_{n\geqslant 1}$.

Покажем, что из $F_n(x) \xrightarrow{w} F(x),\ n \to \infty$, следует $L(F_n, F) \to 0$. Проведем это доказательство от противного, то есть предположим, что существует последовательность $\{F_n\}_{n\geqslant 1} \subset \mathcal{P}$ такая, что $F_n \xrightarrow{w} F \in \mathcal{P},\ n \to \infty$, но $L(F_n, F) \not\to 0$. Отсюда следует, что существует число $\varepsilon > 0$ и подпоследовательность $\{F_{n_k}\}_{k\geqslant 1}$ такая, что $L(F_{n_k}, F) \geqslant \varepsilon$, и мы можем считать что вся последовательность $\{F_n\}_{n\geqslant 1}$ обладает этим свойством. Поэтому между графиками функций $F_n,\ n \geqslant 1$, и F можно вставить квадрат $K_\varepsilon(x_n, y_n)$ со стороной ε где x_n и y_n координаты центра этого квадрата.

Так как множество $\mathcal{H} = \{F\} \bigcup \{F_n\}_{n\geqslant 1}$ слабо относительно компактно, то существуют числа $a < 0,\ b > 0$ такие, что графики всех функций семейства $\mathcal{H}$ лежат в горизонтальной полосе $\{(x, y) : x \leqslant a, 0 \leqslant y \leqslant \varepsilon/2\}$ слева от a и в полосе $\{(x, y) : x \geqslant b, 1 - \varepsilon/2 \leqslant y \leqslant 1\}$ справа от b. Квадраты $K_\varepsilon(x_n, y_n)$ со стороной ε в эти полосы шириной $\varepsilon/2$ не помещаются, поэтому $a < x_n < b$ для всех $n \geqslant 1$. Очевидно, что $0 < y_n < 1,\ n \geqslant 1$. Так как последовательности $\{x_n\}_{n\geqslant 1}$ и $\{y_n\}_{n\geqslant 1}$ ограничены, из них можно выделить сходящиеся подпоследовательности и мы можем считать, что сама последовательность $\{F_n\}_{n\geqslant 1}$ такова, что $F_n(x) \xrightarrow{w} F(x),\ n \to \infty$, и между графиками функций $F_n(x), n \geqslant 1$, и F можно вставить квадраты $K_\varepsilon(x_n, y_n)$ со стороной ε такие, что $x_n \to x_0,\ y_n \to y_0,\ n \to \infty$. Во все эти квадраты для всех n, больших некоторого n_0, можно вставить один и тот же квадрат $K_{\varepsilon/2}(x_0, y_0)$. Ясно, что для каждого $n > n_0$ на интервале $(x_0 - \varepsilon/4, x_0 + \varepsilon/4)$ разность $F_n(x) - F(x)$ знакопостоянна и по абсолютной величине не меньше $\varepsilon/2$.

Рассмотрим функцию $f(x)$, равную нулю при $x < x_0 - \varepsilon/4$, равную $\varepsilon/2$ при $x > x_0 + \varepsilon/4$ и линейно возрастающую от нуля до $\varepsilon/2$ при возрастании x от $x_0 - \varepsilon/4$ до $x_0 + \varepsilon/4$. Ее производная $f'(x) = 0$ при $x < x_0 - \varepsilon/4$ и при $x > x_0 + \varepsilon/4$ и $f'(x) = 1$ при $x \in (x_0 - \varepsilon/4, x_0 + \varepsilon/4)$.

Рассмотрим при $n > n_0$

$$\int\limits_{-\infty}^{\infty} f(x)d(F_n(x) - F(x)) = f(x)(F_n(x) - F(x))|_{-\infty}^{+\infty} - \int\limits_{-\infty}^{\infty} f'(x)(F_n(x) - F(x))dx.$$

Первое слагаемое в правой части этого равенства равно нулю в силу того, что $F_n(\pm\infty) = F(\pm\infty)$, а абсолютная величина второго слагаемого равна

$$\left|\int\limits_{x_0-\varepsilon/4}^{x_0+\varepsilon/4}(F_n(x)-F(x))dx\right| = \int\limits_{x_0-\varepsilon/4}^{x_0+\varepsilon/4}|F_n(x)-F(x)|\,dx \geqslant \frac{\varepsilon}{2}\int\limits_{x_0-\varepsilon/4}^{x_0+\varepsilon/4}dx =$$

$$= \left(\frac{\varepsilon}{2}\right)^2 > 0.$$

То есть мы нашли непрерывную ограниченную функцию $f(x)$, для которой при $n > n_0$

$$\left|\int\limits_{-\infty}^{\infty} f(x)dF_n(x) - \int\limits_{-\infty}^{\infty} f(x)dF(x)\right| \geqslant \left(\frac{\varepsilon}{2}\right)^2 > 0,$$

что невозможно, если $F_n(x) \xrightarrow{w} F(x),\ n \to \infty$. Предположим теперь, что $L(F_n, F) \to 0,\ n \to \infty$. Зафиксируем непрерывную ограниченную функцию $f(x)$ и пусть $M = \sup\limits_{-\infty<x<\infty}|f(x)|$. Ясно, что для любых $-\infty < a < b < \infty$

$$\int\limits_{-\infty}^{\infty} f(x)dF_n(x) - \int\limits_{-\infty}^{\infty} f(x)dF(x) =$$

$$= \int\limits_{-\infty}^{a} f(x)dF_n(x) - \int\limits_{-\infty}^{a} f(x)dF(x) + \int\limits_{a}^{b} f(x)d(F_n(x)-F(x)) -$$

$$- \int\limits_{b}^{\infty} f(x)d(1-F_n(x)) - \int\limits_{b}^{\infty} f(x)d(1-F(x)).$$

Зафиксируем число $\varepsilon > 0$. Так как $L(F_n, F) \to 0,\ n \to \infty$, то семейство $\{F_n\}_{n\geqslant 1}$ является слабо относительно компактным, таким же будет семейство $\mathcal{H} = \{F\}\bigcup\{F_n\}_{n\geqslant 1}$, и мы можем выбрать точки $a < 0$ и $b > 0$ так, чтобы они были точками непрерывности всех функций семейства $\mathcal{H}$ и так, чтобы $h^-(-a;\mathcal{H}) < \frac{\varepsilon}{8M}$ и $h^+(b;\mathcal{H}) < \frac{\varepsilon}{8M}$. При таком выборе чисел a и b сумма абсолютных величин двух первых и двух последних слагаемых в правой части последнего равенства не превосходит $\varepsilon/2$.

Рассмотрим величину $\int\limits_a^b f(x)d(F_n(x)-F(x))$. Для данной непрерывной функции f и данного ε существует непрерывно дифференцируемая функция g такая, что

$$|f(x)-g(x)| \leqslant \varepsilon/8,\ x \in [a,b],$$

(в качестве g можно взять многочлен). Пусть $\max\limits_{a \leqslant x \leqslant b} |g'(x)| = N$. Отметим, что величина N зависит от функции f, чисел a и b и от ε. Очевидно, что

$$\left|\int\limits_a^b f(x)d(F_n(x) - F(x))\right| \leqslant \int\limits_a^b |f(x) - g(x)|d(F_n(x) + F(x)) +$$

$$+\left|\int\limits_a^b g(x)d(F_n(x) - F(x))\right|.$$

Первое слагаемое в правой части этого неравенства не превосходит $\varepsilon/4$, а второе слагаемое равно

$$\left|g(x)(F_n(x) - F(x))|_a^b - \int\limits_a^b g'(x)(F_n(x) - F(x))dx\right| \leqslant$$

$$\leqslant M|F_n(b) - F(b)| + M|F_n(a) - F(a)| + N\int\limits_a^b |F_n(x) - F(x)|dx.$$

Два первых слагаемых в правой части этого неравенства стремятся к нулю при росте n в силу того, что точки a и b являются точками непрерывности функции $F(x)$, а последнее слагаемое стремится к нулю в силу того, что по следствию к теореме 3

$$\int\limits_a^b |F_n(x) - F(x)|dx \leqslant (1 + b - a)L(F_n, F),$$

поэтому сумма этих трех слагаемых будет меньше $\varepsilon/4$ при всех n, начиная с некоторого. Собирая вместе полученные утверждения, мы видим, что

$$\left|\int\limits_{-\infty}^{\infty} f(x)d(F_n(x) - F(x))\right| < \varepsilon$$

при всех n, начиная некоторого.

Теорема доказана.

Связь между сходимостью по вероятностью случайных величин и слабой сходимостью их функций распределения. Пусть на некотором вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P})$ задана последовательность $\{\xi_n\}_{n\geqslant 1}$ случайных величин, сходящаяся по вероятности к некоторой случайной величине ξ. Оказывается, что последовательность F_{ξ_n} функций распределения этих случайных величин слабо сходится к функции распределения F_ξ. Обратное, вообще говоря, неверно, что показывает пример последовательности $\{\eta_n\}_{n\geqslant 1}$ случайных величин таких, что $\eta_{2k+1} = \zeta,\ k = 0, 1, 2, \ldots,\ \eta_{2k} = \chi,\ k = 1, 2, \ldots,$ где случайные величины ζ и χ имеют одну и ту же функцию распределения F и $\mathsf{P}(\zeta = \chi) \neq 1$. Эта последовательность $\{\eta_n\}_{n\geqslant 1}$ имеет две предельные точки ζ и χ (в смысле сходимости по вероятности, почти наверное и т.д.; в смысле любой сходимости на множестве с л у ч а й н ы х в е л и ч и н) и поэтому не может быть сходящейся. При этом функции распределения F_{η_n} при всех n совпадают с F и эта последовательность сходится к F в любом смысле. Но в одном важном частном случае (он связан с законом больших чисел) из слабой сходимости распределений следует сходимость соответствующих случайных величин по вероятности. Это - случай, когда предельная функция распределения вырождена.

Теорема 5. *Из сходимости $\xi_n \xrightarrow{\mathsf{P}} \xi,\ n \to \infty$, случайных величин следует сходимость $F_{\xi_n} \Rightarrow F_\xi$ их функций распределения. Если последовательность функций распределения $F_{\xi_n} \Rightarrow E_c,\ n \to \infty$, то $\xi_n \xrightarrow{\mathsf{P}} c$.*

Д о к а з а т е л ь с т в о. Пусть $\xi_n \xrightarrow{\mathsf{P}} \xi,\ n \to \infty$, то есть для любого $\varepsilon > 0$

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

Определим событие $A_{n,\varepsilon} = \{\omega : |\xi_n(\omega) - \xi(\omega)| < \varepsilon\}$ и обозначим $\overline{A_{n,\varepsilon}}$ его дополнение. Легко видеть, что для любого действительного x

$$F_{\xi_n}(x) = \mathsf{P}(\xi_n < x) = \mathsf{P}(\xi + \xi_n - \xi < x) =$$

$$+\mathsf{P}(\{\xi + \xi_n - \xi < x\} \cap A_{n,\varepsilon}) + \mathsf{P}(\{\xi + \xi_n - \xi < x\} \cap \overline{A_{n,\varepsilon}}).$$

Нетрудно убедиться в справедливости включений

$$\{\xi + \varepsilon < x\} \cap A_{n,\varepsilon} \subset \{\xi + \xi_n - \xi < x\} \cap A_{n,\varepsilon} \subset \{\xi - \varepsilon < x\} \cap A_{n,\varepsilon}$$

(так как на событии $A_{n,\varepsilon}$ величина $\xi_n - \xi$ меньше ε, то если на этом событии $\xi + \varepsilon < x$, то и подавно $\xi + \xi_n - \xi < x$), откуда следует, что

$$\mathsf{P}(\{\xi + \varepsilon < x\} \cap A_{n,\varepsilon}) \leqslant \mathsf{P}(\{\xi + \xi_n - \xi < x\} \cap A_{n,\varepsilon}) \leqslant \mathsf{P}(\{\xi - \varepsilon < x\} \cap A_{n,\varepsilon}).$$

Так как $A_{n,\varepsilon} = \Omega \setminus \overline{A_{n,\varepsilon}}$, то

$$\mathsf{P}(\{\xi + \varepsilon < x\} \cap A_{n,\varepsilon}) = \mathsf{P}(\xi + \varepsilon < x) - \mathsf{P}(\{\xi + \varepsilon < x\} \cap \overline{A_{n,\varepsilon}})$$

и

$$\mathsf{P}(\{\xi-\varepsilon<x\}\cap A_{n,\varepsilon})=\mathsf{P}(\xi-\varepsilon<x)-\mathsf{P}(\{\xi-\varepsilon<x\}\cap\overline{A_{n,\varepsilon}})$$

Собирая вместе полученные неравенства, мы видим, что

$$\mathsf{P}(\xi+\varepsilon<x)-\mathsf{P}(\{\xi+\varepsilon<x\}\cap\overline{A_{n,\varepsilon}})+\mathsf{P}(\{\xi+\xi_n-\xi<x\}\cap\overline{A_{n,\varepsilon}})\leqslant F_{\xi_n(x)}\leqslant$$

$$\leqslant\mathsf{P}(\xi-\varepsilon<x)-\mathsf{P}(\{\xi-\varepsilon<x\}\cap\overline{A_{n,\varepsilon}})+\mathsf{P}(\{\xi+\xi_n-\xi<x\}\cap\overline{A_{n,\varepsilon}}).$$

Переходя в правом из этих неравенств к верхнему пределу по $n\to\infty$, а в левом - к нижнему, и учитывая, что вероятности всех событий, в которых участвует событие $\overline{A_{n,\varepsilon}}$, не превосходят $\mathsf{P}(\overline{A_{n,\varepsilon}})\to 0$ при $n\to\infty$ для любого $\varepsilon>0$, мы получим

$$F_\xi(x-\varepsilon)\leqslant\varliminf_{n\to\infty}F_{\xi_n}(x)\leqslant\varlimsup_{n\to\infty}F_{\xi_n}(x)\leqslant F_\xi(x+\varepsilon).$$

Отсюда следует, что для любого $x\in\mathbb{C}(F_\xi)$

$$\varliminf_{n\to\infty}F_{\xi_n}(x)=\varlimsup_{n\to\infty}F_{\xi_n}(x)=\lim_{n\to\infty}F_{\xi_n}(x)=F_\xi(x).$$

Так как на множестве $\mathcal{P}$ понятия слабой сходимости и сходимости в основном совпадают, то утверждение $F_{\xi_n}\Rightarrow F_\xi,\ n\to\infty$, становится очевидным.

Пусть $F_{\xi_n}\Rightarrow E_c,\ n\to\infty$. Так как точками непрерывности предельной функции E_c являются все точки $x\neq c$, то для любого $\varepsilon>0$

$$F_{\xi_n}(c-\varepsilon)\to 0,\ F_{\xi_n}(c+\varepsilon)\to 1,\ n\to\infty.$$

Поэтому

$$\mathsf{P}(c-\varepsilon<\xi_n<c+\varepsilon)\geqslant\mathsf{P}(c-\frac{\varepsilon}{2}\leqslant\xi_n<c+\varepsilon)=F_{\xi_n}(c+\varepsilon)-F_{\xi_n}(c-\frac{\varepsilon}{2})\to 1,\ n\to\infty.$$

Отсюда сразу следует, что для любого $\varepsilon>0$

$$\mathsf{P}(|\xi_n-c|\geqslant\varepsilon)\to 0,\ n\to\infty.$$

Теорема доказана.