Теория вероятностей. Лекция восемнадцатая
Слабая сходимость

Дмитрий Валерьевич Хлопин
glukanat@mail.ru

Институт математики и механики им. Н.Н.Красовского

27.02.2019

Что разобрали:

- Неравенства концентрации меры
- Сходимости случайных величин
- Усиленный закон больших чисел
- Слабая сходимость
- Характеристические функции
- Центральная предельная теорема

Различные виды сходимости отвечают за разные типы надежности: начиная с некоторого момента все будет хорошо, отклонения больше заданного достаточно редки, ущерб будет мал в среднем и т. п. Все эти сходимости имеют дело со случайной величиной, понимаемой с точностью почти всюду. При этом какие-то вещи могут быть узнаны с вероятностью 1 (в предположении, что мы не в худшем из миров).

Усиленный закон больших чисел (в форме Колмогорова)

Пусть $\{\xi_n\}$ – последовательность независимых одинаково распределенных случайных величин со средним m и конечной дисперсией. Тогда

$$\frac{\xi_1 + \ldots + \xi_n}{n} \xrightarrow{\text{п.в.}} m \text{ при } n \to \infty.$$

Закон повторного логарифма

Теорема 3. [без д-ва] Пусть дана последовательность независимых одинаково распределенных случайных величин ξ_n со средним 0 и дисперсией $\sigma^2 > 0$. Тогда

$$\sup_{n>k} \frac{\xi_1 + \ldots + \xi_n}{\sigma\sqrt{2n \ln\ln n}} \xrightarrow{\text{п.в.}} 0 \text{ при } k \uparrow \infty,$$

и для всех положительных ε выполнено

$$\mathbb{P}(\xi_1 + \ldots + \xi_n > (1+\varepsilon)\sigma\sqrt{2n \ln\ln n} \text{ бесконечно много раз}) = 0,$$
$$\mathbb{P}(\xi_1 + \ldots + \xi_n < (1-\varepsilon)\sigma\sqrt{2n \ln\ln n} \text{ бесконечно много раз}) = 1.$$

Закон нуля и единицы. Очень частный случай

Теорема 4. Пусть дана последовательность независимых случайных величин ξ_n. Пусть некоторое событие A принадлежит $\sigma(\xi_n, \xi_{n+1}, \dots)$ для всех $n \in \mathbb{N}$. Тогда $\mathbb{P}(A) = 0$ или $\mathbb{P}(A) = 1$.

Доказательство.

Отметим, что A не зависит от всех σ-алгебр $\sigma(\xi_1, \xi_2, \dots, \xi_{n-1})$, тогда не зависит от объединяющей их $\sigma(\xi_1, \xi_2, \dots, \xi_{n-1}, \xi_n \dots)$, но $A \in \sigma(\xi_1, \xi_2, \dots, \xi_{n-1}, \xi_n \dots)$, следовательно событие A не зависит от самого себя, то есть $\mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}^2(A)$. Решая квадратное уравнение, получаем требуемое.

Задача [1,5 балла] Пусть борелевская функция f от координат $x_1, \dots, x_n, \dots$ такова, что для всех $i \in \mathbb{N}$
$f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, x_{i+2}, \dots) = f(x_1, \dots, x_{i-1}, x_{i+1}, x_i, x_{i+2}, \dots)$.
Пусть $X_1, \dots, X_n, \dots$ — независимые в совокупности одинаково распределенные случайные величины. Докажите, что случайная величина $f(X_1, \dots, X_n, \dots)$ с вероятностью 1 – константа.

От сильной к слабой сходимости...

- Неравенства концентрации меры
- Сходимости случайных величин
- Усиленный закон больших чисел
- Слабая сходимость
- Характеристические функции
- Центральная предельная теорема

Подумать: рассмотрите последовательность случайных величин, распределенных по закону $U[1, 1+1/n]$ или, например, $N(1, 1/n)$. Соответствующие им плотности сходятся почти всюду к нулю. Но вроде понятно, что в пределе вся масса сосредоточена в единице...

Слабая сходимость

Пусть имеется вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$. Рассмотрим также различные вероятности на $(\mathbb{R}, \mathcal{B})$. Если μ – вероятность на $\mathbb{R}$, то $F_\mu(x) \stackrel{\triangle}{=} \mu((-\infty, x])$. Для $\mu = \xi \# \mathbb{P}$ сократим обозначение до $F_\mu = F_\xi$.

Будем говорить, что μ_n слабо сходится к вероятности μ, обозначая её $\mu_n \xrightarrow{w} \mu$, если для каждой непрерывной ограниченной $\phi : \mathbb{R} \to \mathbb{R}$

$$\int_{\mathbb{R}} \phi(x)\mu_n(dx) \to \int_{\mathbb{R}} \phi(x)\mu(dx).$$

Подумать: докажите, что если каждая μ_n сосредоточена в некоторой x_n, то есть $\mu_n(A) \stackrel{\triangle}{=} |A \cap \{x_n\}|$, то μ_n сходится тогда и только тогда, когда сходится x_n.

Подумать: стоит ли требовать сходимость интегралов для всех измеримых функций?

Слабая сходимость, куда еще слабее...

Функция $\phi : \mathbb{R} \to \mathbb{R}$ называется финитной, если она равна нулю вне некоторого компакта $[a, b]$. Множество непрерывных финитных функций обозначим через $C_0(\mathbb{R})$, множество бесконечное число раз дифференцируемых финитных функций — через $C_0^\infty(\mathbb{R})$.

Предложение 7. μ_n сходится слабо к μ тогда и только тогда, когда для каждой бесконечное число раз дифференцируемой финитной функции $\phi : \mathbb{R} \to \mathbb{R}$ выполнено

$$\int_{\mathbb{R}} \phi(x)\mu_n(dx) \to \int_{\mathbb{R}} \phi(x)\mu(dx).$$

Эквивалентность определения слабой сходимости

Доказательство. Нужна лишь достаточность.

Каждому $\varepsilon > 0$ выберем $[a,b]$ такой, что $\mu[a,b] \geq 1-\varepsilon$. Пусть ϕ – произвольная непрерывная функция. Найдется $\phi^\varepsilon \in C_0^\infty(\mathbb{R})$ такая, что $|\phi^\varepsilon(x) - \phi(x)| \leq \varepsilon$ на $x \in [a,b]$. Можно считать, что $|\phi| \leq c$, $|\phi^\varepsilon| \leq 2c$.

Теперь, для всех n

$$\left|\int_{-\infty}^{+\infty} \phi(x)\mu_n(dx) - \int_{-\infty}^{+\infty} \phi^\varepsilon(x)\mu_n(dx)\right| \leq \int_a^b |\phi(x) - \phi^\varepsilon(x)|\mu_n(dx)$$
$$+ \left|\int_{\mathbb{R}\setminus[a,b]} \phi(x)\mu_n(dx)\right| + \left|\int_{\mathbb{R}\setminus[a,b]} \phi^\varepsilon(x)\mu_n(dx)\right| \leq \varepsilon + 3c\varepsilon.$$

Аналогично,

$$\left|\int_{-\infty}^{+\infty} \phi(x)\mu(dx) - \int_{-\infty}^{+\infty} \phi^\varepsilon(x)\mu(dx)\right| \leq \varepsilon + 3c\varepsilon.$$

Теперь сходимость следует из $\displaystyle\int_{-\infty}^{+\infty} \phi^\varepsilon \mu_n(dx) \to \int_{-\infty}^{+\infty} \phi^\varepsilon \mu(dx)$ при $n \to \infty$.

Интегрирование по частям

Для произвольной $\phi \in C_0^\infty(\mathbb{R})$

$$\int_{-\infty}^{+\infty} \phi(x)\mu(dx) = -\int_{-\infty}^{+\infty} \phi'(x)F_\mu(x)dx.$$

Для этого запишем

$$\phi(x) = \int_{-\infty}^{x} \phi'(t)dt = \int_{-\infty}^{+\infty} \mathbf{1}_{(-\infty,x]}(t)\phi'(t)dt.$$

Тогда, переставляя интегралы и используя равенства $\mathbf{1}_{(-\infty,x]}(t) = \mathbf{1}_{[t,+\infty)}(x)$, $\int_t^{+\infty} \mu(dx) = 1 - F_\mu(t)$, мы получаем, что

$$\begin{aligned}\int_{-\infty}^{+\infty} \phi(x)\mu(dx) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathbf{1}_{(-\infty,x]}(t)\phi'(t)dt\mu(dx)\\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathbf{1}_{[t,+\infty)}(x)\mu(dx)\phi'(t)dt = \int_{-\infty}^{+\infty}(1 - F_\mu(t))\phi'(t)dt\\ &= -\int_{-\infty}^{+\infty} F_\mu(t)\phi'(t)dt.\end{aligned}$$

Сколько-то фактов без доказательства

Слабая сходимость задает топологию на множестве всевозможных вероятностных мер. Эта топология может быть описана метрикой, например введенной ранее метрикой Канторовича. В функане, откуда и пошло название, рассматривают такую сходимость как сходимость на линейных отображениях вида

$$\phi \mapsto \int_{\mathbb{R}} \phi(x)\mu(dx).$$

Для полных сепарабельных метрических пространств Ω имеет место:
Теорема Александрова [без д-ва]. Следующие условия эквивалентны:

- $\mu_n \xrightarrow{w} \mu$;
- $\mu(A) \leq \liminf_{n\to\infty} \mu_n(A)$ для всякого открытого множества A;
- $\mu(A) \geq \limsup_{n\to\infty} \mu_n(A)$ для всякого замкнутого множества A;
- $\mu(A) = \lim_{n\to\infty} \mu_n(A)$ для всякого борелевского множества A, у границы которого мера μ равна нулю.

Сходимость случайных величин по распределению

Будем говорить, что ξ_n сходится к ξ по распределению, обозначая её $\xi_n \xrightarrow{d} \xi$, если $\mu_{\xi_n} = \xi_n \# \mathbb{P}$ сходится к $\mu_\xi = \xi \# \mathbb{P}$ слабо, то есть для всякой непрерывной ограниченной функции ϕ

$$\int_\Omega \phi(\xi_n(\omega))\mathbb{P}(d\omega) \to \int_\Omega \phi(\xi(\omega))\mathbb{P}(d\omega).$$

Подумать: приведите пример $\xi : \Omega \to \{-1, 1\}$ такой, что $\xi_n = \xi \xrightarrow{d} \xi$, $\xi_n = \xi \xrightarrow{d} -\xi$. Можно ли сделать вывод, что $\xi = -\xi$? Еще раз прочитайте предыдущий абзац.

Подумать: сравните определения, вдруг там что напутано...

Из $F_{\xi_n}(t) \to F_\xi(t)$ п.в. следует $\mu_n \xrightarrow{w} \mu$

Предложение 8. Последовательность случайных величин ξ_n сходится по распределению к случайной величине ξ тогда и только тогда, когда для всех точек непрерывности функции распределения F_ξ

$$F_{\xi_n}(t) \to F_\xi(t) \text{ при } n \to \infty.$$

Докажем больше: $\mu_n \xrightarrow{w} \mu$ тогда и только тогда, когда

$$F_{\mu_n}(t) \to F_\mu(t)$$

во всех точках непрерывности F_μ.

Доказательство. Докажем сначала обратное. Пусть ϕ — бесконечное число раз дифференцируемая финитная функция. Тогда для не более чем счетного числа точек t (лишь в точках разрыва) возможно не имеет место сходимость $\phi'(t)F_{\mu_n}(t) \to \phi'(t)F_\mu(t)$. По теореме Лебега о мажорируемой сходимости, используя интегрирование по частям, заключаем, что $\int_{-\infty}^{+\infty} \phi(x)\mu_n(dx)$ сходятся к $\int_{-\infty}^{+\infty} \phi(x)\mu(dx)$.

Из $\mu_n \xrightarrow{w} \mu$ следует $F_{\xi_n}(t) \to F_\xi(t)$ п.в.

Пусть $\mu_n \xrightarrow{w} \mu$. Пусть t – точка непрерывности F_μ, тогда для любого $\varepsilon > 0$ можно найти $\delta > 0$ такое, что $F_\mu(t-\delta) \geq F_\mu(t) - \varepsilon$. Выберем ϕ равным 1 на $(-\infty, t-\delta]$ и 0 на $[t, +\infty)$. На $[t-\delta, t]$ продлим ϕ аффинно. При достаточно больших n имеем

$$F_{\mu_n}(t) = \int_{-\infty}^{t} \mu_n(dx) \geq \int_{-\infty}^{+\infty} \phi(x)\mu_n(dx)$$

$$> \int_{-\infty}^{+\infty} \phi(x)\mu(dx) - \varepsilon > \int_{-\infty}^{t-\delta} \phi(x)\mu(dx) - \varepsilon = F_\mu(t-\delta) - \varepsilon \geq F_\mu(t) - 2\varepsilon.$$

Аналогичным образом мы получаем, что при достаточно больших n $F_{\mu_n}(t) \leq F_\mu(t) + 2\varepsilon$.

$$\xi_n \xrightarrow{P} \xi \Rightarrow \xi_n \xrightarrow{d} \xi,\ \xi_n \xrightarrow{\text{п.в.}} \xi \Rightarrow \xi_n \xrightarrow{d} \xi,\ \xi_n \xrightarrow{L^p} \xi \Rightarrow \xi_n \xrightarrow{d} \xi$$

Предложение 9. Если $\xi_n \xrightarrow{P} \xi$, то $\xi_n \xrightarrow{d} \xi$.

Доказательство. Положим

$$\mu \triangleq \mu_\xi, \quad \mu_n \triangleq \mu_{\xi_n}.$$

Пусть $\phi \in C_0^\infty(\mathbb{R})$, тогда $|\phi(x)| \le c$, $|\phi'(x)| \le c$. Выберем $\varepsilon > 0$ и введем

$$G_n^\varepsilon \triangleq \{\omega : |\xi_n - \xi| < \varepsilon\}.$$

Имеем, что

$$\begin{aligned}
\left| \int_{-\infty}^{+\infty} \phi(x)\mu(dx) - \int_{-\infty}^{+\infty} \phi(x)\mu_n(dx) \right| &= |\mathbb{E}\phi(\xi_n) - \mathbb{E}\phi(\xi)| \\
&\le |\mathbb{E}\phi(\xi_n) - \mathbb{E}\phi(\xi)|\mathbf{1}_{G_n^\varepsilon} + |\mathbb{E}\phi(\xi_n)|\mathbf{1}_{\Omega \smallsetminus G_n^\varepsilon} + \mathbb{E}|\phi(\xi)|\mathbf{1}_{\Omega \smallsetminus G_n^\varepsilon} \\
&\le c\varepsilon + 2c\mathbb{P}(\Omega \smallsetminus G_n^\varepsilon).
\end{aligned}$$

Поскольку $\mathbb{P}(\Omega \smallsetminus G_n^\varepsilon) = \mathbb{P}(|\xi_n - \xi| \ge \varepsilon) \to 0$ при $n \to \infty$, получаем, что

$$\lim_{n\to\infty} \left| \int_{-\infty}^{+\infty} \phi(x)\mu(dx) - \int_{-\infty}^{+\infty} \phi(x)\mu_n(dx) \right| \le c\varepsilon.$$

Устремляя ε к нулю, получаем требуемое.

$\xi_n \xrightarrow{d} \xi \Rightarrow \xi_n \xrightarrow{P} \xi$, если ξ — константа

Предложение 10 Если $\xi_n \xrightarrow{d} m$ для некоторой константы m, то $\xi_n \xrightarrow{P} m$.

Доказательство.

Пусть $\xi = m$. Отметим, что F_ξ непрерывна всюду, кроме точки m, более того $F_\xi(m-\varepsilon) = 0, F_\xi(m+\varepsilon) = 1$ для всех положительных ε. Остальное - голый счет: для всех положительных ε при $n \uparrow \infty$ имеем

$$\mathbb{P}(|\xi_n - m| \leq \varepsilon) \geq \mathbb{P}(m - \varepsilon < \xi_n \leq m + \varepsilon)$$
$$= F_{\xi_n}(m+\varepsilon) - F_{\xi_n}(m-\varepsilon) \to F_\xi(m+\varepsilon) - F_\xi(m-\varepsilon) = 1.$$

Что разобрали:

- Неравенства концентрации меры
- Сходимости случайных величин
- Усиленный закон больших чисел
- Слабая сходимость. Банальности
- Характеристические функции
- Центральная предельная теорема

Слабая сходимость предполагает, что изучаемое явление — черный ящик, не пытаясь залезть внутрь, она пытается угадать реакции на раздражители (значения у непрерывных функций). Она не требует знания ω, наоборот, она скорее предполагает, что Вы его не знаете и никогда не узнаете. И дает все остальное.

Впрочем, слабая сходимость еще долго не закончится: характеристические функции тоже о слабой сходимости...

На пять минут...

1. Эту задачу или ее вариацию я подкину на одну из следующих пятиминуток. Пусть $X_n (n > 3)$ — последовательность независимых случайных величин, $\mathbb{P}(X_n = n) = \mathbb{P}(X_n = -n) = \frac{1}{n \ln n}$, $\mathbb{P}(X_n = 0) = 1 - \frac{2}{n \ln n}$. Опровергнуть или доказать: сходится ли эта последовательность по вероятности, в среднем $(p = 2)$, почти всюду?

2. Опровергнуть или доказать: для любой случайной величины ξ последовательность ξ/n сходится по вероятности, в среднем $(p = 1)$, почти всюду?

Полное решение. Поскольку $\xi(\omega)/n \to 0$ для всех ω, значит сходимость к нулю почти всюду доказана. Следовательно, доказана и по вероятности.

Сходимость в среднем при $p = 2$ эквивалентна $\mathbb{E}|\xi/n - 0|^2 = \mathbb{E}|\xi/n|^2 = \mathbb{E}|\xi|^2/n^2 \to 0$, что так, если матожидание $\mathbb{E}|\xi|^2$ существует. Для случайной величины, принимающей значение 5^n с вероятностью $1/2^n$ для всех $n \in \mathbb{N}$, такой сходимости нет.

Ответ: Да, для почти всюду и по вероятности; нет, для в среднем.