Лекция 10. Сходимости в теорвере. Закон больших чисел.

22 апреля 2022 г.

1 Сходимость в вероятностных пространствах.

В этом разделе мы будем рассматривать бесконечные последовательности с.в. $\{X_n\}_{n\in\mathbb{N}}$ и определим, что такое сходимость такой последовательности, пытаясь создать аналогию с пределом числовой последовательности.

Сходимость по распределению.

Говорят, что $X_n \to X$ по распределению (все X_n и X – с.в. на одном и том же вероятностном пространстве), если функции распределения $F_{X_i}(x)$ поточечно сходятся к функции распределения $F_X(x)$, кроме, возможно, счетного числа точек разрыва $F_X(x)$. Последнее уточнение важно, так как возьмем, например, последовательность $X_n \sim U(0, \frac{1}{n})$. В пределе, казалось бы, они дадут с.в. $X = 0$. Но все $F_{X_n}(0) = 0$, в то время как $F_X(0) = 1$.

Также заметим, что сходимость по распределению не подразумевает сходимость плотностей вероятности. Например, можно рассмотреть $X_n \in [0, 1]$ с функцией распределения $f_{X_n} = 1 - \cos(2\pi nx)$. Они сходятся по распределению к $U(0, 1)$, но их плотности вообще не сходятся. Но обратное верно: если плотности сходятся почти всюду, то почти всюду сходятся и функции распределения.

Пример: запустили фабрику по выпуску костей $d6$. Но так как оборуждование новое и ненастроенное, кости получаются нечестными. Но со временем, они становятся все более и более честными, и их функция распределения сходится поточечно к функции распределения честной $d6$.

Сходимость по вероятности.

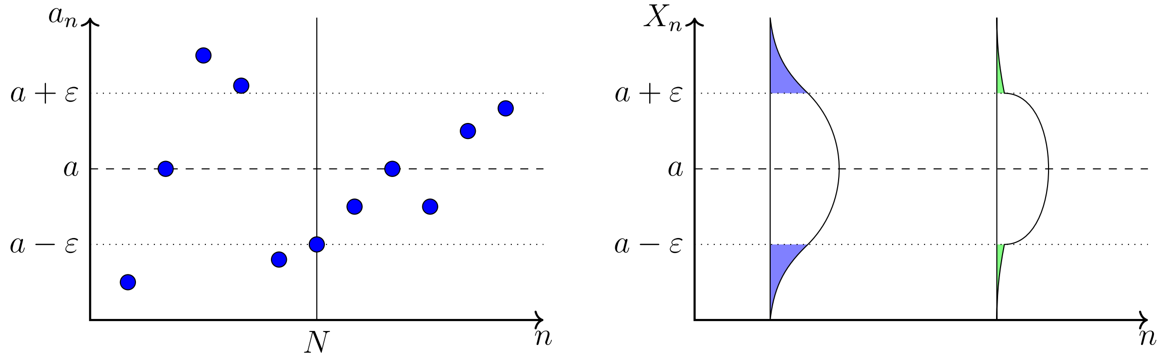
Говорят, что $X_n \to X$ по вероятности (все X_n и X – с.в. на одном и том же вероятностном пространстве), если для любого $\varepsilon > 0$

$$\boxed{\lim_{n\to+\infty} \Pr(|X - X_n| > \varepsilon) = 0}$$

То есть в данном случае это означает, что X_n с росто n становится все более и более зависимой от X, точнее даже почти равной. Особый случай, когда X — константа.

Тогда это означает, что последовательность X_n становится все более и более сконцентрированной вокруг этой константы.

Проведем аналогию с сходимостью последовательностей. Для последовательности $a_n \to a$ значит, что для любого чилсла $\varepsilon > 0$ мы можем нарисовать ε-трубку вокруг a, и в нее, начиная с какого-то N попадут все члены последовательности. В случае, когда с.в. $X_n \to a$ по вероятнсоти, это значит, что для любого $\delta > 0$ начиная с какого-то N почти вся вероятностная масса X_n (а именно хотя бы $1 - \delta$) будет сконцентрирована в ε-трубке.



Сходимость по вероятности подразумевает сходимость по распределению (доказать — упражнение).

Важные свойства:

- Если $g(x)$ — непрерывная функция, то $g(X_n)$ также сходятся к $g(X)$ по вероятности.
- Если $X_n \to X$ и $Y_n \to Y$, то $(X_n + Y_n) \to (X + Y)$ (везде сходимость по вероятности).

Пример: вы смотрите на то, какие числа выдает генератор псевдослучайных чисел. Пусть X_n — ваше предположение о том, что он выдаст следующим числом после того, как вы пронаблюдали уже n числе, выданных им. Чем больше n, тем больше вероятность правильно угадать паттерн, по которому действует генератор. Поэтому вероятность, что $X_n = X$ (где X — число, выданное генератором после вашего предположения), стремится к 1.

Еще пример: Пусть X_i – независимые с.в., все $\sim U(0, 1)$. Пусть $Y_n = \min_{i \in [1..n]}(X_i)$. Тогда $\Pr(|Y_n - 0| \geq \varepsilon) = \varepsilon^n \to 0$.

NB: сходимость по вероятности не подразумевает сходимость матожиданий. Пусть X_n такие, что

$$\Pr(X_n = 0) = 1 - \frac{1}{n}$$
$$\Pr(X_n = n^2) = \frac{1}{n}$$

Тогда $X_n \to 0$ по вероятности, но $E[X_n] = n^2 \cdot \frac{1}{n} = n$ — расходится.

Сходимость почти наверное. Говорят, что $X_n \to X$ по вероятности (все X_n и X – с.в. на одном и том же вероятностном пространстве), если

$$\Pr(\lim_{n\to+\infty} X_n = X) = 1$$

Это означает, что при всех возможных элементарных исходах из Ω (кроме, возможно, какого-то множество исходов с суммарной вероятностью ноль) X_n образуют такую последовательность, которая сходится к значению X на этом же элементарном исходе:

$$\Pr\left(\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}\right) = 1.$$

Это наиболее сильный вид сходимости, который влечет за собой сходимость по вероятности (и следовательно, сходимость по распределению). Доказать это — тоже упражнение.

Пример: Пусть продолжительность жизни какого-то животного — случайная величина. Пусть X_i — количество еды, которое съело это животное в i-й день своей жизни. При любом исходе с какого-то момента X_i станет нулем, то есть последовательность сходится к нулю почти наверное.

2 Законы больших чисел

Слабый закон больших чисел.

Пусть X_i — независимые одинаково распределенные с.в. с матожиданием μ и дисперсией σ^2. Пусть $M_n = \frac{1}{n}\sum_{i=1}^n X_i$ — среднее первых n случайных величин. Тогда

- $E[M_n] = E[\frac{1}{n}\sum_{i=1}^n X_i] = \frac{1}{n} \cdot n\mu = \mu$
- $\operatorname{Var}(M_n) = \operatorname{Var}(\frac{1}{n}\sum_{i=1}^n X_i) = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n}$

То есть с ростом n среднее значение M_n сохраняется, а вот дисперсия уменьшается (то есть увеличивается концентрация вокруг матожидания). Поэтому давайте применим неравенство Чебышева:

$$\Pr(|M_n - \mu| \geq \varepsilon) \leq \frac{\operatorname{Var}(M_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \to 0.$$

То есть M_n сходится по вероятности к μ, что и называется слабым законом больших чисел.

$$\forall\varepsilon > 0 \text{ выполняется } \lim_{n\to+\infty} \Pr(|M_n - \mu| > \varepsilon) = 0$$

Возможные применения.

- Для точного измерения при наличии шумов. Пусть хотим узнать какую-то величину μ, но при ее измерении у нас всегда есть какая-то случайная погрешность W (причем $E(W_i) = 0$). Тогда чем больше измерений мы проведем, тем меньше будет вероятность, что среднее наших измерений сильно отклоняется от реального значения.

- Среднее по экспериментам. Пусть есть нечестная монета с вероятностью p, что выпадет орел. Тогда чем больше экспериментов мы проведем, тем лучше сможем определить этот p.

Сильный закон больших чисел (закон Колмогорова).
В тех же условиях, что и для слабого закона, верно и более сильное утверждение:

$$\boxed{\Pr(\lim_{n \to +\infty} M_n = \mu) = 1}$$

То есть M_n сходится к μ почти наверное. Также существует версия и для случая, когда X_i имеют разные распределения (но с конечными матожиданиями и дисперсиями такими, что $\sum_{i=1}^{+\infty} \frac{1}{i^2} \operatorname{Var}(X_i)$ — сходится). В таком случае $(M_n - E[M_n])$ сходится к нулю почти наверное.