

ПРИКЛАДНАЯ СТАТИСТИКА



Посиделка 6: сходимости случайных величин

Ульянкин Ппилиф

Повторять до сходимости — это как жарить до готовности.

(Неизвестный студент Вышки)

В этой посиделке мы поговорим про сходимости, которые часто встречаются в теории вероятностей и математической статистике. Не очень понятно, нафига придумывать какие-то новые сходимости, когда в матане мы уже изучили сходимость обычных последовательностей. Попробуем с этим разобраться.

1 Обычная сходимость

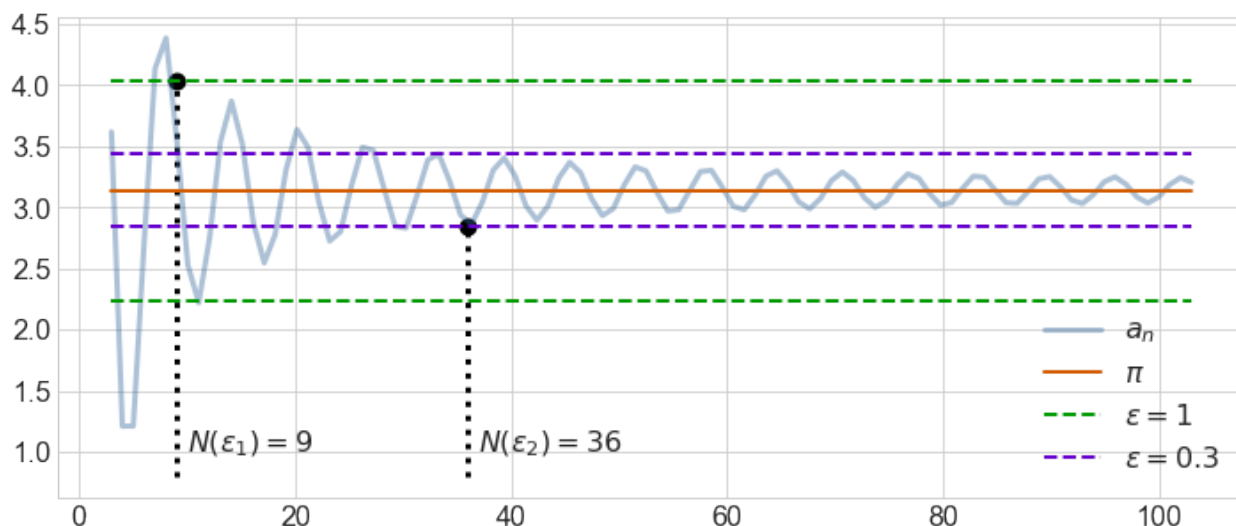
Говорят, что последовательность действительных чисел $(a_n)_{n=1}^{\infty}$ сходится к числу A т.е.

$$\lim_{n \rightarrow \infty} a_n = A,$$

если для любого сколь угодно малого положительного числа $\varepsilon > 0$ существует порядковый номер $N(\varepsilon) \in \mathbb{N}$ такой, что начиная с него a_n отличается от A меньше чем на ε т.е.

$$\forall \varepsilon > 0 \quad \exists N(\varepsilon) \in \mathbb{N} : \quad \forall n > N(\varepsilon) \quad |a_n - A| < \varepsilon.$$

Рассмотрим конкретный пример. Возьмём последовательность $a_n = 10 \cdot n^{0.99} \cdot \sin(n) + \pi$. Она сходится. Причём к числу π . Отложим по оси x номера членов последовательности, а по оси y значения.

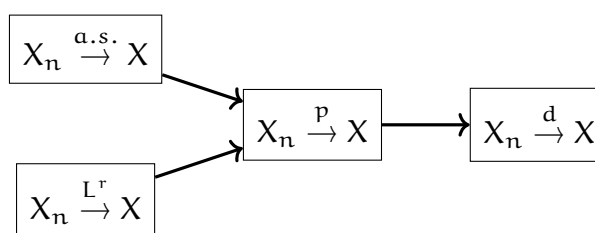


Зафиксируем произвольное число ε , например 1. Начиная с $N(\varepsilon) = 9$ все члены последовательности оказываются зажаты зелёной окрестностью.

Если взять $\varepsilon = 0.3$, то $N(\varepsilon)$ сдвинется вправо и окажется равно 36. Если мы для каждого ε можем найти место, начиная с которого все члены последовательности лежат внутри выделенного нами коридора, то последовательность сходится.

Так обстоят дела с детерминированными последовательностями. В случае, когда речь идёт о последовательностях из случайных величин, всё оказывается немного сложнее. Случайные величины могут с лёгкостью пробивать построенный нами коридор из эпсилонів в любом месте. Из-за этого приходится формулировать сходимость в терминах вероятностей.

Всего в теории вероятностей выделяют четыре вида сходимости. На картинке ниже стрелками показано, какая сходимость следует из какой.



- Самая слабая сходимость, **сходимость по распределению**. Чтобы сказать, что последовательность случайных величин X_n сходится по распределению к случайной величине X , обычно над стрелочкой пишут букву L или букву d . Иногда просто рисуют двойную стрелочку, \Rightarrow . Из-за того, что эта сходимость самая слабая её так иногда и называют — **слабой сходимостью**.
- Сходимость чуть сильнее — это **сходимость по вероятности**. Обычно её обозначают, подписывая над стрелкой букву p . Если последовательность сходится по вероятности, тогда она будет сходиться и по распределению.
- Сходимость по вероятности, в свою очередь следует из **сходимости почти наверное**. (almost surely). Чтобы обозначить эту сходимость, над стрелкой пишут $a.s.$ или п.н. Эту

сходимость часто называют **сильной сходимостью**.

- Также сходимость по вероятности следует из **сходимости в среднем порядка r** . Над стрелкой в случае такой сходимости либо подписывают порядок сходимости, либо пишут L^r . В случае, когда $r = 1$ говорят о **сходимости в среднем**. В случае $r = 2$ говорят о сходимости **в смысле среднего квадратического**.

Последние два вида сходимости самые сильные. Между ними нет чёткой взаимосвязи. Давайте посмотрим по очереди на все четыре вида. Начнём со сходимости по вероятности.

2 Сходимость по вероятности

Определение. Последовательность случайных величин X_1, X_2, X_3, \dots **сходится по вероятности** к случайной величине X , если

$$\forall \varepsilon > 0 \quad \lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0.$$

Обычно это записывают, как $X_n \xrightarrow{p} X$, либо $\text{plim}_{n \rightarrow \infty} X_n = X$.

Довольно часто сходимость по вероятности можно доказать просто воспользовавшись определением.

Упражнение 1

Пусть $X_n \sim \text{Exp}(n)$. Покажите, что $X_n \xrightarrow{p} 0$.

Решение:

Пойдём чётко по определению

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \varepsilon) = \lim_{n \rightarrow \infty} \mathbb{P}(X_n > \varepsilon).$$

Воспользуемся определением функции распределения: $\mathbb{P}(X_n > \varepsilon) = 1 - \mathbb{P}(X_n \leq \varepsilon)$.

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n > \varepsilon) = \lim_{n \rightarrow \infty} 1 - (1 - e^{-n\varepsilon}) = \lim_{n \rightarrow \infty} e^{-n\varepsilon} = 0 \quad \forall \varepsilon > 0.$$

Получается, что

$$\text{plim}_{n \rightarrow \infty} X_n = 0.$$

Подумайте, что будет для $X_n \sim \text{Exp}\left(\frac{1}{n}\right)$.

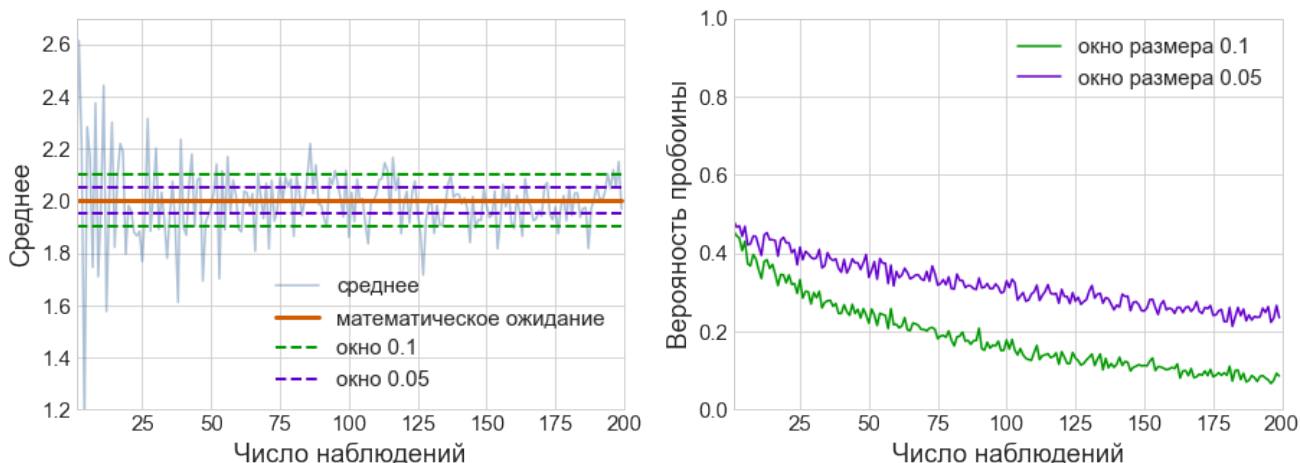
2.1 Сходимость по вероятности в картинках

Для обычной сходимости нам было нужно, чтобы начиная с какого-то номера $N(\varepsilon)$ вся последовательность оставалась в рамках коридора из эпсилонів. В случае сходимости по вероятности, когда мы оказываемся за некоторым номером $N(\varepsilon)$, последовательность может пробивать коридор. Но по мере нашего продвижения вправо, вероятность того, что она пробьёт коридор падает.

Давайте попробуем посмотреть на это на конкретном примере. Предположим, что в наших руках оказалась генеральная совокупность из нормального распределения. Мы сделали из неё выборку некоторого объёма и посчитали по ней выборочное среднее.

$$\bar{X} = \sum_{i=1}^n X_i$$

Выборочное среднее сходится по вероятности к математическому ожиданию при росте n . Этот факт обычно называют законом больших чисел. В следующей посиделке мы сформулируем его как полноценную теорему. Тут же давайте посмотрим как это выглядит в картинках



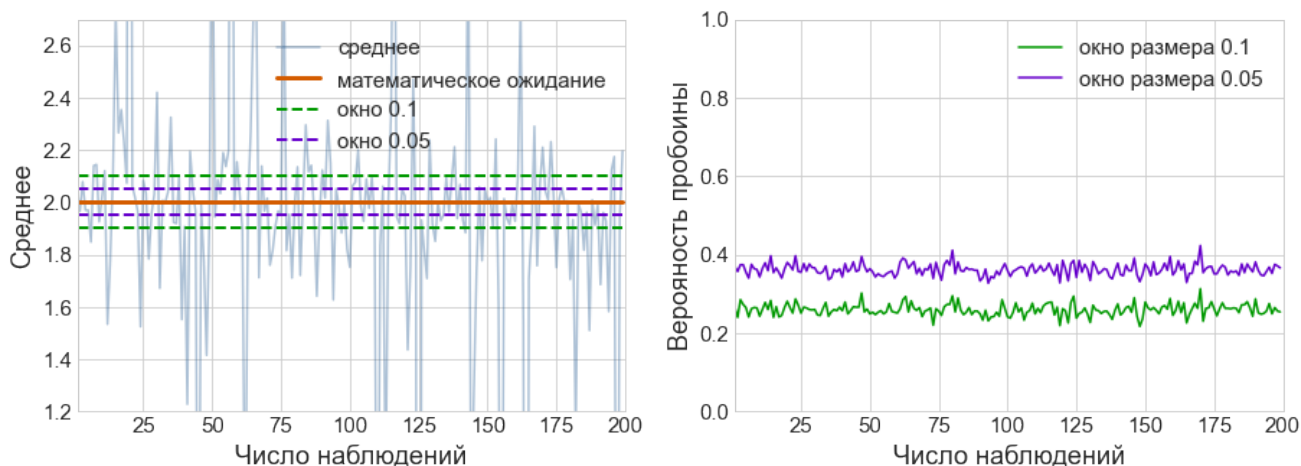
Зафиксируем коридор, ε , за которым мы будем в дальнейшем наблюдать. Как можно увидеть, поначалу, при маленьком числе наблюдений, последовательность довольно часто пробивает коридор. Нет такого, чтобы она один раз погрузилась в него и никогда больше из него не выходила. Со временем, количество пробоин в зафиксированном коридоре падает.

Можно попробовать оценить вероятность того, что последовательность из средних пробьёт на конкретном шаге установленный нами коридор. Для этого можно сгенерировать 1000 траекторий среднего, а дальше посмотреть, как часто на каждом шаге у нас возникает пробоина. Частота таких пробоин и будет оценкой вероятности пробить коридор. Именно её динамика для обоих наших коридоров нарисована на картинке справа.

Теперь мы знаем, как выглядит сходимость по вероятности. Интересно было бы посмотреть, как выглядит её отсутствие. Давайте вспомним распределение Коши. У него не существу-

ет математического ожидания и для него выборочному среднему некуда сходиться. Давайте построим для него пару таких же рисунков.

Мы видим, что среднее постоянно пробивает выделенный нами коридор. При этом, с ростом n , число пробоев не уменьшается. Вероятность пробоя не убывает и ведёт себя стабильно.



2.2 Друзья сходимости по вероятности

У сходимости по вероятности есть много друзей, которые помогают проверять её наличие.

- Первый и главный друг сходимости по вероятности — это закон больших чисел.

Теорема 2.1. Закон Больших Чисел (Пафнутий Львович Чебышёв)

Пусть X_1, \dots, X_n попарно независимые и одинаково распределённые случайные величины с конечным вторым моментом, $E(X_1^2) < \infty$, тогда имеет место сходимость:

$$\frac{X_1 + \dots + X_n}{n} \xrightarrow{p} E(X_1)$$

Про него мы подробно поговорим в следующей посиделке.

- Второй друг сходимости по вероятности — неравенства Чебышёва и Маркова. Они иногда помогают доказать наличие сходимости по определению.

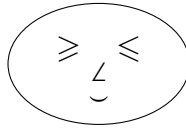
Теорема 2.2. Если у случайной величины X конечная дисперсия, тогда $\forall \varepsilon > 0$

$$\mathbb{P}(|X| \geq \varepsilon) \leq \frac{\mathbb{E}(X)}{\varepsilon}$$

А также

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{\text{Var}(X)}{\varepsilon^2}.$$

Чтобы запомнить, где какие неравенства ставить, можно представить себе что-то подозревающего китайца¹.



В упражнении, которое мы решали выше, неравенством Маркова можно было бы воспользоваться следующим образом

$$\mathbb{P}(X_n \geq \varepsilon) \leq \frac{\mathbb{E}(X)}{\varepsilon} = \frac{1}{n \cdot \varepsilon} \rightarrow 0, \text{ при } n \rightarrow \infty.$$

Получается искомая нами вероятность ограничена при $n \rightarrow \infty$ сверху нулём. Получается она будет равна нулю.

- Третий друг — это арифметические свойства предела. Из-за того, что сходимость по вероятности — наследник обычной сходимости, все свойства сохраняются [1].
- Четвёртый союзник — это теорема Манна-Вальда. В русской литературе её часто называют теоремой Слуцкого. На самом деле в сходимостях есть ещё одна теорема Слуцкого, которую мы будем обсуждать немного ниже. Чтобы не запутаться, будем называть эту теорему на западный манер².

Теорема 2.3. Манна-Вальда-Слуцкого

- Если $X_n \xrightarrow{p} a$ и функция $g(t)$ непрерывна в точке a , тогда $g(X_n) \xrightarrow{p} g(a)$.
- Если $X_n \xrightarrow{p} X$ и функция $g(t)$ непрерывна, тогда $g(X_n) \xrightarrow{p} g(X)$.

Доказательство этой теоремы и ещё ряда свойств можно найти, например, в книге Черновой [1].

2.3 Зачем всё это, зачем?

Нам это всё нужно для благих целей. Чаще всего сходимость по вероятности будет нам требоваться для доказательства состоятельности. Когда мы собираем выборку для того, чтобы оценить какой-нибудь параметр, мы надеемся, что из-за её больших размеров мы окажемся близко к истине. Состоятельность говорит нам именно про это.

Определение. Оценка $\hat{\theta}$ неизвестного параметра θ называется **состоятельной**, если при $n \rightarrow \infty$ она сходится по вероятности к истинному значению, то есть $\hat{\theta} \xrightarrow{p} \theta$.

В будущих посиделках мы будем доказывать состоятельность разных оценок, используя перечисленных выше союзников.

¹Эта рукопись может содержать устаревшие культурные стереотипы и показывается в оригинальном виде

²Запад развращает. Сначала ты начинаешь писать вместо M букву \mathbb{E} , а затем переименовываешь теоремы.

3 Сходимость по распределению

Определение. Последовательность случайных величин X_1, X_2, X_3, \dots **сходится по распределению** к случайной величине X , если при $n \rightarrow \infty$ последовательность $F_{X_n}(x) \rightarrow F_X(x)$ для всех x , в которых $F_X(x)$ непрерывна. Обычно это записывают, как $X_n \xrightarrow{d} X$ либо $X_n \Rightarrow X$. Ещё такую сходимость называют **слабой**, так как она следует из всех остальных сходимостей.

3.1 Сходимость по распределению в картинках

Если последовательность из функций сходится, она есть. Если не сходятся, её нет. Всё просто.

Упражнение 1

Пусть последовательность X_1, X_2, \dots это последовательность случайных величин с функциями распределения

$$F_{X_n}(x) = \begin{cases} 1 - \left(1 - \frac{1}{n}\right)^{nx} & x \geq 0 \\ 0 & \text{иначе} \end{cases}$$

Найдите предел по распределению для X_n .

Решение:

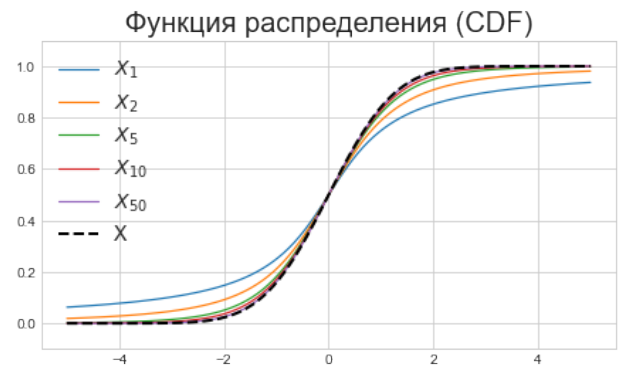
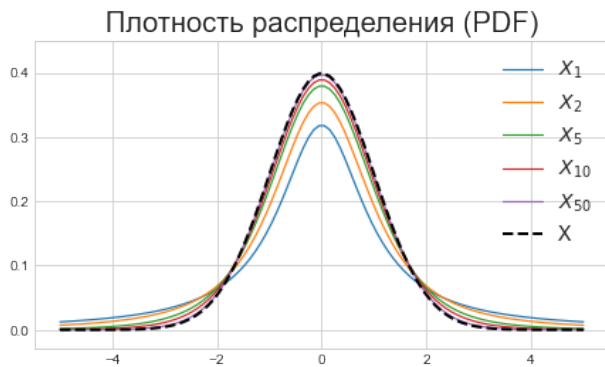
Надо просто найти предел для $F_{X_n}(x)$:

$$F_X(x) = \lim_{n \rightarrow \infty} F_{X_n}(x) = 1 - e^{-x}.$$

Получаем в пределе экспоненциальное распределение, $\text{Exp}(1)$.

Другой пример, который часто встречается на практике — распределение Стьюдента с n степенями свободы, $t(n)$. Оно славится тем, что обладает более тяжёлыми хвостами, чем у нормального распределения. При этом, если $n \rightarrow \infty$, распределение Стьюдента сходится по распределению к стандартному нормальному, $t(n) \xrightarrow{d} N(0, 1)$.

Можно проиллюстрировать этот факт на картинке. Чёрная пунктирная линия — стандартное нормальное распределение. Остальные линии это распределённые по стьюденту случайные величины. Видно, что и плотности и функции распределения понемногу приближаются к предельной.



Можно выписать функцию плотности для распределения Стьюдента и осознать как именно так получается, но мы оставим это читателю в качестве домашнего задания. В определении не очень понятной остаётся формулировка «для всех x , в которых $F_X(x)$ непрерывна.» Чтобы лучше понять её, давайте решим ещё одну задачку.

Упражнение 2

Пусть $X_n \sim U[0; 1]$, а $Y_n = \min(X_1, \dots, X_n)$. Найдите $F_{Y_n}(y)$, а затем докажите, что $Y_n \xrightarrow{d} 0$ и что $Y_n \xrightarrow{p} 0$.

Решение:

Для начала найдём функцию распределения минимума из n независимых случайных величин

$$F_{Y_n}(y) = \mathbb{P}(Y_n \leq y) = \mathbb{P}(\min(X_1, X_2, \dots, X_n) \leq y) = 1 - \mathbb{P}(\min(X_1, X_2, \dots, X_n) > y).$$

Если минимум больше y , то и все случайные величины должны быть больше него

$$\begin{aligned} F_{Y_n}(y) &= 1 - \mathbb{P}(X_1 > y, X_2 > y, \dots, X_n > y) = \\ &= 1 - \mathbb{P}(X_1 > y) \cdot \mathbb{P}(X_2 > y) \cdot \dots \cdot \mathbb{P}(X_n > y) = \\ &= 1 - [1 - \mathbb{P}(X_1 \leq y)] \cdot [1 - \mathbb{P}(X_2 \leq y)] \cdot \dots \cdot [1 - \mathbb{P}(X_n \leq y)] = \\ &= 1 - [1 - F_{X_n}(y)]^n = 1 - [1 - y]^n. \end{aligned}$$

Получается, что

$$F_{Y_n}(y) = \begin{cases} 0, & \text{если } y < 0 \\ 1, & \text{если } y \geq 1 \\ 1 - [1 - y]^n, & \text{если } y \in [0; 1]. \end{cases}$$

Докажем, что $Y_n \xrightarrow{d} 0$. Функция распределения вырожденной случайной величины, принимающей только значение 1 выглядит как

$$F_Y(y) = \begin{cases} 1, & \text{если } y \geq 0 \\ 0, & \text{если } y < 0 \end{cases}$$

Найдём предел $F_n(y)$ во всех точках y и сравним его с $F(y)$

$$y \geq 1 \quad F_{Y_n}(y) = 1 \rightarrow 1 = F_Y(y)$$

$$y < 0 \quad F_{Y_n}(y) = 0 \rightarrow 0 = F_Y(y)$$

$$0 < y < 1 \quad F_{Y_n}(y) = 1 - [1 - y]^n \rightarrow 1 = F_Y(y)$$

$$y = 0 \quad F_{Y_n}(y) = 0 \rightarrow 0 \neq 1 = F_Y(y)$$

В точке $y = 0$ сходимости нет. Определение сходимости по распределению разрешает нам проигнорировать эту точку, потому что у предельной функции $F_Y(y)$ в ней есть разрыв.

3.2 Друзья сходимости по распределению

Доказывать сходимость по распределению проще, чем сходимость по вероятности. Достаточно просто понять, как ведут себя функции распределения. Тем не менее, у сходимости по распределению тоже есть друзья, которыми мы дальше будем пользоваться.

- Во-первых, это ЦПТ.

Теорема 3.1. Центральная Предельная Теорема (Прокопий Петрович Ляпунов)

Пусть X_1, \dots, X_n попарно независимые и одинаково распределённые случайные величины с конечным вторым моментом, $E(X_i^2) < \infty$, тогда при $n \rightarrow \infty$ имеет место сходимость по распределению:

$$\bar{x} \xrightarrow{d} N\left(E(X_i), \frac{\text{Var}(X_i)}{n}\right)$$

Этот же факт можно переписать немного иначе:

$$\sqrt{n} \cdot [\bar{x} - E(X_i)] \xrightarrow{d} N(0, \text{Var}(X_i))$$

или даже как

$$\sqrt{n} \cdot \frac{\bar{x} - E(X_i)}{\sqrt{\text{Var}(X_i)}} \xrightarrow{d} N(0, 1).$$

Если говорить простым языком, то при определённых условиях сумма достаточно большого числа случайных величин имеет распределение близкое к нормальному. **Главное,**

чтобы случайные величины были похожи и не было такого, что одна резко выделяется на фоне остальных. Разговору о ЦПТ мы посвятим следующую посиделку.

- Во-вторых, это обобщение ЦПТ — Дельта-метод. Его мы введём, когда будем строить статистику средних.
- В-третьих, это ещё одна теорема Слуцкого. Она так называется и на Западе.

Теорема 3.2. Слуцкого Если $X_n \xrightarrow{p} a \neq 0$ и $Y_n \xrightarrow{d} Y$, тогда

- $Y_n + X_n \xrightarrow{d} Y + a$
- $Y_n \cdot X_n \xrightarrow{d} Y \cdot a$
- $\frac{Y_n}{X_n} \xrightarrow{d} \frac{Y}{a}$.

Ещё раз подчеркнём, что сходимость по распределению самая слабая. Если есть сходимость по вероятности, то будет и сходимость по распределению. В обратную сторону это верно только в случае сходимости к константе. На доказательство этих замечательных фактов можно, например, в лекциях Дэвида Хантера. [2] Вам нужна будет вторая глава.

3.3 Зачем всё это, зачем?

Помните, в самом начале мы говорили про схему матстата? Мы сказали, что когда у нас есть точечная оценка, нам хочется понять, насколько она точная. Для этого надо пытаться выяснить как оценка распределена. Нужен союзник, который подскажет. Сходимость по распределению, в частности ЦПТ, помогает находить асимптотических союзников для этих целей.

4 Сходимость в среднем

Когда мы придумываем определение для сходимости X_n к X , мы обычно придумываем какое-то расстояние между X и X_n . Если оно становится всё меньше и меньше, значит последовательность сходится. Если мы определим это расстояние как $\mathbb{P}(|X_n - X| > \varepsilon)$, тогда мы получим сходимость по вероятности. Другой способ определить расстояние между случайными величинами — посмотреть на математическое ожидание.

Определение. Последовательность случайных величин X_1, X_2, X_3, \dots **сходится в среднем порядка r** к случайной величине X , если

$$\lim_{n \rightarrow \infty} \mathbb{E}(|X_n - X|^r) = 0.$$

Обычно это записывают, как $X_n \xrightarrow{L^r} X$ либо $X_n \xrightarrow{\text{m.s.}} X$ в случае, когда $r = 2$.

Из этой сходимости следует сходимость по вероятности и сходимость по распределению. При этом её довольно легко проверить. Чаще всего она нужна, как техническое условие в раз-

ных теоремах, поэтому мы с ней будем сталкиваться довольно редко. Тем не менее, для полноты картины с этим видом сходимости тоже нужно немного поработать.

Упражнение 1

Пусть $X_n \sim \mathcal{U} \left[0; \frac{1}{n}\right]$. Покажите, что $X_n \xrightarrow{L^r} 0$.

Решение:

Воспользуемся определением. Плотность распределения для X_n выглядит как

$$f_{X_n}(x) = \begin{cases} n, & 0 \leq x \leq \frac{1}{n} \\ 0, & \text{иначе} \end{cases}.$$

Найдём

$$\mathbb{E}(|X_n - 0|^r) = \int_0^{\frac{1}{n}} x^r \cdot n \, dx = \frac{1}{(r+1)n^r} \rightarrow 0 \quad \forall r \geq 1.$$

Выходит, что $X_n \xrightarrow{L^r} 0$ при $r \geq 1$. В подобранном примере все x положительные, поэтому проблем с модулем не возникло.

Если у нас есть сходимость в среднем, то остальные две сходимости тоже будут присутствовать. Сходимость по вероятности можно углядеть, если понять, что дисперсия случайной величины X_n с каждым новым n всё меньше. Сходимость по распределению к вырожденному распределению можно углядеть, если заметить, что в плотности распределения фигурирует $1/n$. Точно такое же $1/n$ будет фигурировать в функции распределения и последовательность из них будет сходиться к вырожденному распределению.

Сходимость по вероятности слабее. Иногда бывает так, что сходимость по вероятности есть, но сходимости в среднем нет. Пусть последовательность дискретных случайных величин задаётся следующим образом:

$$X_n = \begin{cases} n^2, & p = \frac{1}{n} \\ 0, & p = 1 - \frac{1}{n} \end{cases}$$

Эта последовательность будет сходиться к нулю по вероятности

$$\lim_{n \rightarrow \infty} P(|X_n - 0| \geq \varepsilon) = \lim_{n \rightarrow \infty} P(X_n = n^2) = \lim_{n \rightarrow \infty} \frac{1}{n} = 0.$$

Число n растёт во времени, а $\varepsilon > 0$ произвольное маленькое число. Чтобы разница $X_n - 0$ оказалась больше этого произвольного маленького числа, нужно, чтобы случайная величина

приняла значение n^2 , поэтому $P(|X_n - 0| \geq \varepsilon) = P(X_n = n^2)$.

При этом, наша последовательность не будет сходиться к нулю в среднем

$$\lim_{n \rightarrow \infty} E(|X_n|^r) = \lim_{n \rightarrow \infty} (n^{2r} \cdot \frac{1}{n} + 0 \cdot (1 - \frac{1}{n})) = \lim_{n \rightarrow \infty} n^{2r-1} = \infty \quad \forall r \geq 1.$$

У сходимости в среднем есть довольно интересное свойство, связанное с моментами.

Утверждение 4.1. Пусть $1 \leq r \leq s$. Если $X_n \xrightarrow{L^s} X$, тогда $X_n \xrightarrow{L^r} X$.

Для сходимости в среднем также работает теорема о непрерывном отображении (Манна-Вальда-Слуцкого), то есть если $X_n \xrightarrow{L^r} X$ и функция $g(t)$ непрерывна, тогда $g(X_n) \xrightarrow{L^r} g(X)$.

5 Сходимость почти наверное

Осталось разобраться с последним видом сходимости, сходимостью почти наверное. До этого мы с вами никогда не вдавались в то, как именно устроена случайная величина. Для сходимости почти наверное — это жутко важно. Поэтому немного вспомним основные определения.

5.1 Сигма-алгебра

Упражнение 1

Рассмотрим простой случайный эксперимент: игральный кубик подбрасывают два раза. Множество Ω исходов данного эксперимента содержит 36 элементов. Вася знает результат подбрасываний, и сообщает Пете значение случайной величины Z — произведения очков на выпавших гранях.

Всегда ли сможет Петя, владея своей информацией, определить произошли ли события: A — оба раза выпала единица, B — хотя бы раз выпала пятерка, C — хотя бы раз выпала шестерка?

Решение:

Про события A и B Петя всегда сможет сказать произошли ли они: в первом случае достаточно сравнить Z с единицей, во втором — определить, делится ли Z на 5. Однако может сложиться такая ситуация, что Петя не будет уверен, произошло ли C : например, если Z окажется равным шести, то, может быть, это была пара $(6, 1)$ и тогда C произошло, а может это была пара $(2, 3)$ и тогда C не произошло.

Можно составить список событий, про которые Петя, зная Z , всегда сможет сказать, произошли ли они. Обозначим этот список событий буквой \mathcal{F} . В отличие от Пети Вася может уверенно сказать произошло ли любое событие и аналогичный список событий для него — все подмножества множества Ω .

В список \mathcal{F} входят довольно много событий, а сам список обладает рядом важных свойств, а именно:

- а. Множество $\Omega \in \mathcal{F}$ — даже ничего не зная, можно быть уверенным, что событие Ω произошло.
- б. Множество $\emptyset \in \mathcal{F}$ — даже ничего не зная, можно быть уверенным, что событие \emptyset не произошло.
- в. Если множества A и B входят в список \mathcal{F} , то $A \cup B$ и $A \cap B$ входят в список \mathcal{F} . Если Петя способен определить, произошли ли A и B по отдельности, то он путем простых логических заключений способен определить, произошли ли $A \cup B$ и $A \cap B$.

Подобные списки событий являются очень важными для нас и называются сигма-алгебрами (σ -алгебрами).

Определение. Набор подмножеств множества Ω называется **σ -алгеброй**, если обладает следующими тремя свойствами:

SA1 Множество $\Omega \in \mathcal{F}$

SA2 Из того, что множество $A \in \mathcal{F}$ следует, что $A^c \in \mathcal{F}$

SA3 Если $A_1 \in \mathcal{F}$, $A_2 \in \mathcal{F}$, $A_3 \in \mathcal{F}$, и т.д., то $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Сигма-алгебры нужны для того, чтобы моделировать наделенность информацией рационального индивида. Также они нужны для некоторых технических целей в различных теоремах. Сигма-алгебры помогают моделировать ситуации, когда в сознании наблюдателя несколько событий по каким-то причинам «слипаются», и наблюдатель не может отличить их друг от друга.

Определение. **Случайная величина** — это функция, которая отображает сигма-алгебру на множество действительных чисел, $X(\omega) : \mathcal{F} \rightarrow \mathbb{R}$.

До текущего момента мы не особо вдавались в подробности, связанные с тем как устроена случайная величина $X(\omega)$ при разных значениях ω . Всем сходимость выше было на это наплевать. Со сходимостью почти наверное другая история.

5.2 Сходимость почти наверное

Чтобы стало немного легче понять определение, давайте решим упражнение. Пусть у нас есть монетка, и мы подкидываем её один раз. Пространство элементарных исходов в таком случае выглядит как: $\Omega = \{O, P\}$. Определим последовательность случайных величин следующим образом.

$$X_n(\omega) = \begin{cases} \frac{n}{n+1}, & \omega = O \\ (-1)^n, & \omega = P \end{cases}$$

Пусть мы один раз подбрасываем монетку, а затем выписываем последовательность. Для каких ω последовательность случайных величин сходится?

Ответ прост. Ежели выпал орёл, то последовательность вошла в траекторию $\frac{n}{n+1}$ и сошлась к единице. Если выпала решка, то всё пропало. Найдём вероятность

$$\mathbb{P}(\lim_{n \rightarrow \infty} X_n(\omega) = 1).$$

Нам нужно найти вероятность события, которое заключается в том, что наша последовательность сойдётся. Очевидно, что такая вероятность равна 0.5, так как последовательность будет сходиться только при выпадении орла. Обычно эту вероятность записывают вот так:

$$P(\{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = 1\}).$$

Читаем по символам: вероятность множества омег, на которых последовательность сходится. То есть вероятность события, состоящего в том, что последовательность сходится. Ежели эта вероятность равна единице, то сходимост наступат почти наверное. Если среди омег можно найти какую-то подпоследовательность из случайных величин, где сходимости нет, то всё пропало, сходимости нет. В нашем случае таким набором омег является решка. Ей соответствует расходящаяся последовательность.

Определение. Говорят, что последовательность случайных величин X_1, X_2, \dots **сходится к случайной величине X почти наверное**, если

$$P(\{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}) = 1.$$

Обычно это записывают как $X_n \xrightarrow{a.s.} X$.

Обратите внимание, что в сходимости по вероятности речь шла о пределе вероятности, а тут речь идёт про вероятность предела.

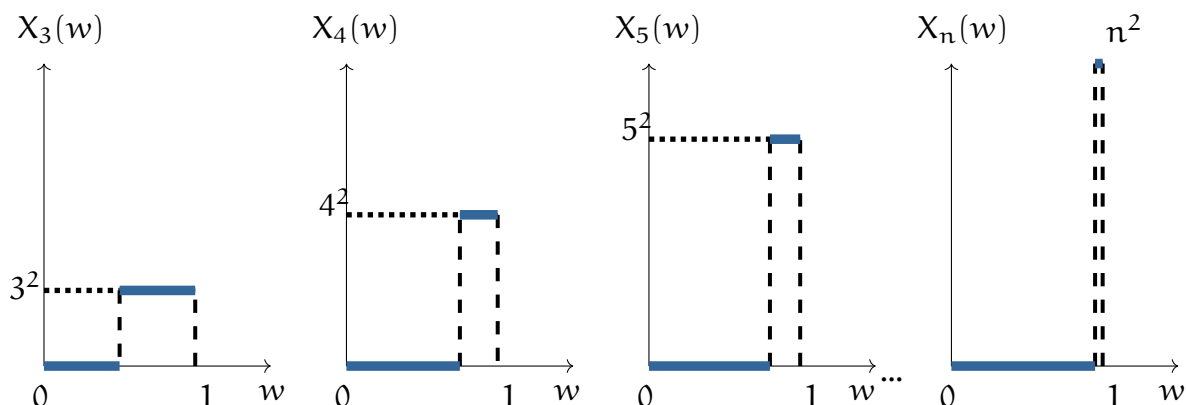
Давайте попробуем проиллюстрировать всё вот это вот на примере, который мы уже видели выше. Пусть $\Omega = [0; 1]$. На нём задана дискретная случайная величина

$$X_n = \begin{cases} n^2, & p = \frac{1}{n} \\ 0, & p = 1 - \frac{1}{n} \end{cases}$$

Сходимость почти наверное для этой случайной величины зависит от того, как именно она устроена на пространстве элементарных исходов. Например, структура может быть такой

$$X_n = \begin{cases} n^2, & \omega \in (1 - \frac{1}{n}, 1] \\ 0, & \omega \in [0, 1 - \frac{1}{n}]. \end{cases}$$

Получается, что у нас отрезочек из элементарных исходов, на котором случайная величина принимает большие значения, становится всё меньше и меньше.

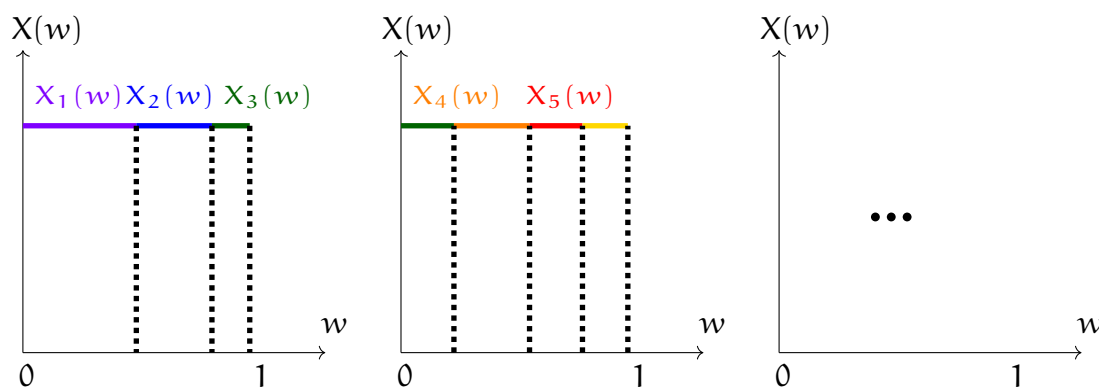


Какое бы $\omega \in [0; 1]$ мы не взяли, рано или поздно оно выпадет из той части отрезка, где случайная величина принимает значение n^2 . Например, для $\omega = 0.85$ последовательность будет $1, 2^2, 3^2, 4^2, 5^2, 0, 0, 0, 0, 0 \dots$

То есть, начиная с какого-то момента ω попал в ту часть элементарных исходов, которые соответствуют нулю. Так будет происходить для любого ω . Это и означает сходимость почти наверное.

Если строить случайную величину из $[0; 1]$ иначе, сходимости может не быть. Заставим отрезок длины $1/n$, на котором выпадает n^2 , бегать по отрезку $[0; 1]$ так, чтобы любое ω попадало в него бесконечное число раз и тем самым существовала бы подпоследовательность $X_{n_k} \rightarrow \infty$.

Будем откладывать отрезочки длины $\frac{1}{n}$ друг за другом. Когда мы будем пробивать суммой отрезочков единицу, будем начинать сначала. Мы сможем бегать значениями $\frac{1}{n}$ по нашему отрезку вечно, так как ряд $\sum \frac{1}{n}$ расходится. Например, для второй картинке $\frac{1}{2} + \frac{1}{3} + \frac{1}{4} > 1$, и часть $X_4(\omega)$ покрывает начало отрезочка на третьей картинке.



Для $\omega = 0.85$ мы тогда получим последовательность $0, 0, 9, 0, 0, 0, 0, 0, 100, 0, 0, \dots$ Внутри постоянно вылезает n^2 . Он вылезает всё реже и реже. Этого достаточно для сходимости по вероятности, но мало для сходимости почти наверное. Вероятность предела не будет равна единице и сходимости не будет.

Ещё раз обратите внимание на то, что сходимость почти наверное не имеет прямых связей со сходимостью в среднем. Для рассмотренного выше примера сходимость в среднем не выполняется. Сходимость почти наверное может как выполняться, так и не выполняться в зависимости от устройства пространства элементарных исходов.

5.3 Почему почти наверное

Важно понимать тонкую разницу между событием, которое не произойдёт никогда и тем, вероятность которого равна нулю. Нулевая вероятность события вовсе не означает того, что это событие никогда не наступит. Чисто теоретически наступление такого события возможно, но вероятность этого события равна нулю.

Например, вероятность того, что непрерывная случайная величина попадёт в какую-то конкретную точку равна нулю. Тем не менее, сгенерировав такую случайную величину на компьютере, мы увидим, что она приняла конкретное значение, то есть попала в некоторую точку. Удивительно, но при каждой реализации случайной величины мы наблюдаем невозможное событие. Точно также важно понимать разницу между утверждением верным всегда и почти наверное верным.

Ещё один пример. Представим себе, как игрок в дартс бросает в мишень дротик. Пусть, как бы игрок не метнул дротик, он всегда попадает в мишень. Вероятность того, что дротик попадёт в какой-то конкретный регион мишени равна отношению площади этого региона к общей площади мишени. Например, вероятность того, что дротик попадет в правую половину мишени равна 0.5. Рассмотрим событие "дротик попадает в конкретную хорду мишени". Площадь хорды равна нулю. То есть дротик не попадет в хорду почти наверное.

Тем не менее множество точек на хорде не пусто и попадание в неё является чисто теоретически возможным. То же самое можно сказать и о любой другой точке на мишени. Так как любая точка будет иметь нулевую площадь, дротик не попадет в неё почти наверное. Однако дротик точно должен попасть в какую-то точку мишени! Таким образом при каждом бросании дротика происходит событие, имеющее нулевую вероятность. Вот такой вот странный объект — вероятность.

5.4 Признаки сходимости почти наверное

Доказывать сходимость почти наверное через определение — сложно. Есть признаки, которые в этом помогают. Подробнее про них можно почитать в ?? и ??. Давайте перечислим несколько из них и попробуем использовать для решения какой-нибудь задачи.

Теорема 5.1. Пусть у нас есть последовательность случайных величин X_1, X_2, X_3, \dots . Если $\forall \varepsilon > 0$ выполняется

$$\sum_{n=1}^{\infty} \mathbb{P}(|X_n - X| > \varepsilon) < \infty,$$

тогда $X_n \xrightarrow{\text{a.s.}} X$

Если оказалось, что сумма бесконечная, это не означает отсутствия сходимости. Сформулируем более мощное условие

Теорема 5.2. Пусть у нас есть последовательность случайных величин X_1, X_2, X_3, \dots Определим для $\varepsilon > 0$ набор событий

$$A_m = \{|X_n - X| < \varepsilon \quad \forall n \geq m\}.$$

Последовательность $X_n \xrightarrow{a.s.} X$ тогда и только тогда, когда $\forall \varepsilon > 0$ выполняется

$$\lim_{m \rightarrow \infty} \mathbb{P}(A_m) = 1.$$

Упражнение 2

Пусть $X_n \sim \text{Bern}(\frac{1}{n})$. Покажите, что $X_n \xrightarrow{a.s.} 0$.

Решение:

Попробуем воспользоваться достаточным условием. Зафиксируем маленькое $\varepsilon > 0$. Для него

$$\sum_{n=1}^{\infty} \mathbb{P}(|X_n| > \varepsilon) = \sum_{n=1}^{\infty} \mathbb{P}(X_n = 1) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

Никаких выводов сделать не можем. Попробуем воспользоваться вторым признаком и выпишем соответствующую последовательность из событий

$$A_m = \{|X_n| < \varepsilon, \quad \forall n \geq m\} = \{X_n = 0, \quad \forall n \geq m\}.$$

Найдём вероятность одного из таких событий

$$\begin{aligned} \mathbb{P}(A_m) &= \mathbb{P}(\{X_n = 0, \quad \forall n \geq m\}) = \mathbb{P}(\{X_n = 0, \quad n = m, m+1, \dots, N \quad \forall N > m\}) = \\ &= \mathbb{P}(X_m = 0) \cdot \mathbb{P}(X_{m+1} = 0) \cdot \dots \cdot \mathbb{P}(X_N = 0) = \frac{m-1}{m} \cdot \frac{m}{m+1} \cdot \dots \cdot \frac{N-1}{N} = \frac{m-1}{N}. \end{aligned}$$

Если выбрать N достаточно большим, можно получить, что $\mathbb{P}(A_m)$ меньше любого наперед положительного числа. Получается, что $\mathbb{P}(A_m) = 0 \quad \forall m \geq 2$. Отсюда мы делаем вывод, что $\lim_{m \rightarrow \infty} \mathbb{P}(A_m) = 0$ и у нас нет сходимости почти наверное к нулю.

Последнее, что стоит здесь упомянуть — для сходимости почти наверное также работает теорема о непрерывном отображении, то есть если $X_n \xrightarrow{a.s.} X$ и функция $g(t)$ непрерывна, тогда $g(X_n) \xrightarrow{a.s.} g(X)$.

5.5 Зачем всё это, зачем?

Самый важный пример сходимости почти наверное — **усиленный закон больших чисел**. Про него мы также более подробно поговорим в следующей посиделке.

Теорема 5.3. Усиленный Закон Больших Чисел

Пусть X_1, \dots, X_n попарно независимые и одинаково распределённые случайные величины с конечным математическим ожиданием, $E(X_i) < \infty$, тогда имеет место сходимость:

$$\frac{X_1 + \dots + X_n}{n} \xrightarrow{\text{a.s.}} E(X_i)$$

Список литературы

- [1] *Н.И.Чернова* (2007). Теория вероятностей. // <https://tvims.nsu.ru/chernova/tv/lec/node53.html>.
- [2] *David Hunter* (2006). Asymptotic Tools. // <http://personal.psu.edu/drh20/asymp/fall2006/lectures/>.
- [3] *H. Pishro-Nik* (2014). Introduction to probability, statistics, and random processes. // <https://www.probabilitycourse.com>.