Числовые характеристики, сходимость

Высшая школа цифровой культуры
Университет ИТМО
dc@itmo.ru

Содержание

1 Начальные сведения о функциях от случайных величин

В этой, завершающей лекции по теории вероятностей, мы рассмотрим основные числовые характеристики случайных величин, имеющих дискретное или абсолютно непрерывное распределение. Кроме того, мы, наконец, увидим, а о чем же говорит в общем случае центральная предельная теорема. Начнем, однако, с важного пункта – с функций от случайных величин.

Пусть $(\Omega, \Sigma, \mathsf{P})$ – вероятностное пространство и ξ – случайная величина, заданная на этом вероятностном пространстве. Зачастую нам интересна не сама случайная величина, а функция от нее. В этом пункте мы рассмотрим функции $g : \mathbb{R} \to \mathbb{R}$ от случайной величины ξ, то есть функции $g(\xi)$. Как мы знаем, не любая функция, заданная на Ω, является случайной величиной. Нас интересуют такие функции, которые из случайной величины делают тоже случайную величину. Не останавливаясь на деталях отметим, что если g – непрерывная функция, то $g(\xi)$ будет случайной величиной. В принципе, можно выдвинуть даже такое «смелое» утверждение: все представляющие практический интерес функции g являются случайными величинами.

1.1 Монотонные преобразования

Ясно, что если ξ имеет дискретное распределение, то и $g(\xi)$ имеет дискретное распределение, которое вычислить достаточно легко. Даже если какие-то значения случайной величины «склеиваются», то и соответствующие им вероятности «складываются». Например, имея случайную величину, заданную рядом распределения

$$\begin{array}{c|c|c|c|} \xi & -1 & 1 & 2 \\ \hline \mathsf{P} & \frac{1}{4} & \frac{1}{3} & \frac{5}{12} \end{array},$$

можно вычислить распределение случайной величины ξ^2. Так как при возведении в квадрат значения -1 и 1 склеиваются, то ряд распределения для ξ^2 имеет вид

$$\begin{array}{c|c|c|} \xi^2 & 1 & 4 \\ \hline \mathsf{P} & \frac{1}{4}+\frac{1}{3} & \frac{5}{12} \end{array} \Leftrightarrow \begin{array}{c|c|c|} \xi^2 & 1 & 4 \\ \hline \mathsf{P} & \frac{7}{12} & \frac{5}{12} \end{array}$$

Простейшие примеры показывают, что в абсолютно непрерывном случае случайная величина $g(\xi)$ вовсе не всегда имеет абсолютно непрерывное распределение.

Пример 1.1.1 *Пусть ξ – случайная величина, имеющая абсолютно непрерывное распределение. Ясно, что если функция $g \equiv const$, то $g(\xi)$ имеет вырожденное распределение. Легко также понять, что если функция g кусочно-постоянна, то $g(\xi)$ имеет дискретное распределение*

В некоторых случаях, однако, мы можем гарантировать, что преобразование абсолютно непрерывной случайной величины снова даст абсолютно непрерывную случайную величину. Начнем с некоторых примеров, попутно показав еще один способ поиска плотности распределения функции от случайной величины.

Пример 1.1.2 *Пусть случайная величина ξ имеет функцию распределения $F_\xi(x)$ и плотность распределения $f_\xi(x)$. Тогда при $a \neq 0$ случайная величина $\eta = a\xi + b$ имеет плотность распределения*

$$f_\eta(x) = \frac{1}{|a|} f_\xi\left(\frac{x-b}{a}\right).$$

Пусть $a > 0$. Найдем функцию распределения случайной величины η.

$$F_\eta(x) = \mathsf{P}(\eta < x) = \mathsf{P}(a\xi + b < x) = \mathsf{P}\left(\xi < \frac{x-b}{a}\right) = F_\xi\left(\frac{x-b}{a}\right).$$

Тогда

$$F_\eta(x) = F_\xi\left(\frac{x-b}{a}\right) = \int\limits_{-\infty}^{\frac{x-b}{a}} f_\xi(t)dt.$$

В последнем интеграле сделаем замену $t = \frac{z-b}{a}$, тогда $dt = \frac{1}{a}dz$ и

$$F_\eta(x) = \int\limits_{-\infty}^{x} \frac{1}{a} \cdot f_\xi\left(\frac{z-b}{a}\right) dz.$$

Рассуждения при $a < 0$ практически аналогичны, проведите их самостоятельно. В итоге показано, что η имеет абсолютно непрерывное распределение, а ее плотность равна

$$f_\eta(x) = \frac{1}{|a|} f_\xi\left(\frac{x-b}{a}\right).$$

На самом деле в последнем примере не доказано, что интеграл от найденной «плотности» по всей вещественной оси равен единице (доказано лишь представление «функции распределения» в виде интеграла). Здесь нет обмана. Требуемый факт следует из того, что f_ξ – плотность случайной величины ξ. Осознайте этот момент самостоятельно.

Тем самым, мы получили еще один способ доказательства того, что какое-то распределение абсолютно-непрерывно. Напомним, что другой способ – построить функцию распределения, продифференцировать ее и убедиться, что производная является плотностью распределения, мы уже видели и применяли в предыдущей лекции.

Пользуясь разобранным примером с линейным преобразованием, можно получить важные следствия, которые мы будем использовать в дальнейшем.

Следствие 1.1.1 *Пусть* $\xi \sim \mathsf{N}_{0,1}$, *тогда* $\eta = \sigma\xi + a \sim \mathsf{N}_{a,\sigma^2}$.

Доказательство. Действительно,

$$f_\eta(x) = \frac{1}{\sigma} f_\xi\left(\frac{x-a}{\sigma}\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{\sigma^2}},$$

а последняя функция и есть не что иное, как плотность случайной величины, имеющей нормальное распределение N_{a,σ^2}. □

Следствие 1.1.2 *Пусть* $\xi \sim \mathsf{N}_{a,\sigma^2}$, *тогда* $\eta = \frac{\xi - a}{\sigma} \sim \mathsf{N}_{0,1}$.

Это следствие моментально следует из доказанной нами теоремы.

Следующее следствие часто используется даже в том же программировании при моделировании случайности. Зачастую генератор случайных чисел выдает некоторое число ξ в диапазоне $[0, 1]$ (моделирует равномерное распределение $\mathsf{U}_{0,1}$). А как получить случайную точку на отрезке $[a, b]$?

Следствие 1.1.3 *Пусть* $\xi \sim \mathsf{U}_{0,1}$. *Тогда* $\eta = a + \xi(b-a) \sim \mathsf{U}_{a,b}$.

Итак, случайная величина $a + \xi(b-a)$ имеет равномерное на $\mathsf{U}_{a,b}$ распределение. Ну и наоборот.

Следствие 1.1.4 *Пусть* $\xi \sim \mathsf{U}_{a,b}$. *Тогда* $\eta = \frac{\xi - a}{b - a} \sim \mathsf{U}_{0,1}$.

В то же время нельзя обойти вниманием важное обобщение рассмотренного примера.

Теорема 1.1.5 *Пусть случайная величина* ξ *имеет абсолютно непрерывное распределение с плотностью распределения* f_ξ, *а функция* $g : \mathbb{R} \to \mathbb{R}$ *монотонна. Если* g^{-1} *всюду непрерывно дифференцируема, то функция* $\eta = g(\xi)$ *имеет абсолютно непрерывное распределение с плотностью*

$$f_\eta(x) = f_\xi\left(g^{-1}(x)\right)\left(g^{-1}(x)\right)'.$$

Иными словами, приведенная теорема устанавливает, что монотонное преобразование «не портит» распределения, а также приводит формулу для вычисления такого важного объекта, как плотности полученной случайной величины.

1.2 Функции от нескольких случайных величин

Часто приходится иметь дело не с одной, а несколькими случайными величинами, и находить функции от них (например, находить распределение суммы или произведения случайных величин). В случае, когда система случайных величин имеет дискретное распределение, распределение функции пишутся зачастую намного проще, чем в случае, когда система имеет абсолютно непрерывное распределение (примеры мы видели во второй лекции). Мы не будем подробно останавливаться на этих вопросах и оставим их курсу теории вероятностей.

Остановимся на случае, когда случайные величины ξ_1, ξ_2 имеют абсолютно непрерывные распределения с плотностями f_{ξ_1} и f_{ξ_2} и независимы. Мы знаем из предыдущей лекции, что в этом случае плотность совместного распределения факторизуется и равна

$$f_{\vec{\xi}}(x_1, x_2) = f_{\xi_1}(x_1) f_{\xi_2}(x_2).$$

В статистике нам потребуется разбираться с распределением суммы нескольких независимых слагаемых. Оказывается, справедлива следующая торема.

Теорема 1.2.1 (Распределение суммы независимых с.в.) *Пусть случайные величины ξ_1 и ξ_2 независимы и имеют абсолютно непрерывные распределения с плотностями $f_{\xi_1}(x)$ и $f_{\xi_2}(x)$. Тогда сумма $\xi_1 + \xi_2$ имеет абсолютно непрерывное распределение, и ее плотность задается так называемой формулой свертки*

$$f_{\xi_1+\xi_2}(x) = \int\limits_{-\infty}^{+\infty} f_{\xi_1}(t) f_{\xi_2}(x-t)dt.$$

Итак, теорема устанавливает, что сумма независимых и абсолютно непрерывных случайных величин имеет абсолютно непрерывное распределение, а также дает формулу для вычисления плотности этого распределения.

Пример 1.2.1 *Рассмотрим две независимые случайные величины ξ_1 и ξ_2, имеющие стандартные нормальные распределения $\mathsf{N}_{0,1}$. Вычислим плотность суммы. Согласно формуле сверкти,*

$$f_{\xi_1+\xi_2}(x) = \int\limits_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-t)^2}{2}} dt = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} e^{-t^2+xt-\frac{x^2}{2}} dt.$$

В показателе степени экспоненты выделим полный квадрат по t, тогда

$$-t^2 + xt - \frac{x^2}{2} = -\left(t - \frac{x}{2}\right)^2 - \frac{x^2}{4},$$

а последний интеграл перепишется в виде

$$f_{\xi_1+\xi_2}(x) = \frac{1}{2\pi} e^{-\frac{x^2}{4}} \int\limits_{-\infty}^{+\infty} e^{-\left(t-\frac{x}{2}\right)^2} dt.$$

Сделав в нем замену $p = t - \frac{x}{2}$, придем к выражению

$$f_{\xi_1+\xi_2}(x) = \frac{1}{2\pi} e^{-\frac{x^2}{4}} \int\limits_{-\infty}^{+\infty} e^{-p^2} dp = \frac{1}{2\sqrt{\pi}} e^{-\frac{x^2}{4}} = \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-\frac{x^2}{2(\sqrt{2})^2}},$$

так как интеграл $\int\limits_{-\infty}^{+\infty} e^{-p^2} dp = \sqrt{\pi}$, будучи интегралом Эйлера-Пуассона. Итого, мы получили плотность случайной величины, имеющей распределение $\mathsf{N}_{0,2}$.

На самом деле справедливо более общее утверждение, а именно.

Пример 1.2.2 *Пусть случайные величины ξ_1 и ξ_2 независимы и $\xi_1 \sim \mathsf{N}_{a_1,\sigma_1^2}$, $\xi_2 \sim \mathsf{N}_{a_2,\sigma_2^2}$, тогда $\xi_1 + \xi_2 \sim \mathsf{N}_{a_1+a_2,\sigma_1^2+\sigma_2^2}$.*

Иными словами, сумма независимых нормально распределенных случайных величин с параметрами a_1, σ_1^2 и a_2, σ_2^2 тоже имеет нормальное распределение с параметрами $a_1 + a_2, \sigma_1^2 + \sigma_2^2$. Это утверждение при должном терпении может быть доказано аналогично тому, как это только что было сделано нами для двух независимых случайных величин, имеющих стандартное нормальное распределение. Мы опустим вычисления из-за громоздкости.

Используя аналог приведенной теоремы для дискретных распределений, можно доказать и следующие соотношения.

Пример 1.2.3 *Пусть случайные величины ξ_1 и ξ_2 независимы и $\xi_1 \sim \Pi_\lambda$, $\xi_2 \sim \Pi_\mu$, тогда $\xi_1 + \xi_2 \sim \mathsf{\Pi}_{\lambda+\mu}$.*

Пример 1.2.4 *Пусть случайные величины ξ_1 и ξ_2 независимы и $\xi_1 \sim \mathsf{Bin}_{n,p}$, $\xi_2 \sim \mathsf{Bin}_{m,p}$, тогда $\xi_1 + \xi_2 \sim \mathsf{Bin}_{n+m,p}$.*

В лекциях по статистике мы узнаем еще один важный пример распределения, так называемое гамма-распределение, к которому тоже применим формулу свертки.

2 Некоторые числовые характеристики случайных величин

Как было отмечено в простейшем случае, со случайными величинами связывают множество числовых характеристик.

2.1 Медиана

С понятием медианы мы уже встречались, напомним его.

Определение 2.1.1 *Число $a \in \mathbb{R}$ называется медианой случайной величины ξ, если*

$$\mathsf{P}(\xi \leq a) \geq \frac{1}{2} \quad \text{и} \quad \mathsf{P}(\xi \geq a) \geq \frac{1}{2}.$$

Полезно отметить, что в случае, когда функция распределения F_ξ случайной величины ξ непрерывна и монотонна, медиана a может быть найдена (единственным образом!), как решение уравнения $F_\xi(a) = \frac{1}{2}$. В общем же случае медиана, как было ранее показано в рассмотренном дискретном случае, может быть не единственна.

2.2 Математическое ожидание

Дадим отдельно определения математического ожидания для дискретной и абсолютно непрерывной случайных величин. Определение для дискретной случайной величины практически не отличается от введенного нами ранее

Определение 2.2.1 *Пусть случайная величина ξ имеет дискретное распределение. Математическим ожиданием случайной величины ξ называется число*

$$\mathsf{E}\xi = \sum_{\omega\in\Omega} \xi(\omega)\mathsf{P}(\omega) = \sum_{x\in\xi(\Omega)} x\mathsf{P}(\xi = x)$$

при условии, что написанный ряд сходится абсолютно (то есть если существует $\mathsf{E}|\xi|$). Иначе говорят, что математического ожидания не существует.

Математическое ожидание абсолютно непрерывной случайной величины определяется опять же согласно аналогии «сумма-интеграл».

Определение 2.2.2 *Пусть случайная величина ξ имеет абсолютно непрерывное распределение с плотностью $f_\xi(x)$. Математическим ожиданием случайной величины ξ называется число*

$$\mathsf{E}\xi = \int\limits_{-\infty}^{+\infty} x f_\xi(x)dx$$

при условии, что написанный интеграл сходится абсолютно (то есть если существует $\mathsf{E}|\xi|$). Иначе говорят, что математического ожидания не существует.

Как и раньше, смысл математического ожидания – среднее вероятностное значение случайной величины. Напомним и механическую интерпретацию. Пусть на оси в точках с координатами x_i сосредоточены массы, равные p_i, сумма которых равна 1. Тогда координата так называемого центра масс рассматриваемой системы определяется из соотношения

$$x_c = \frac{x_1p_1 + x_2p_2 + ... + x_np_n}{n},$$

что и является аналогом выражения для математического ожидания. Если единичная масса непрерывно распределена вдоль оси с линейной плотностью f_ξ, то получим формулу для центра масс, соответствующую математическому ожиданию в абсолютно непрерывном случае:

$$x_c = \int\limits_{-\infty}^{+\infty} x f_\xi(x)dx.$$

Мы уже сталкивались с вычислением математического ожидания у многих дискретных распределений. Перед тем как вычислить математическое ожидание у «новых» рассмотренных распределений, установим основные свойства.

2.3 Свойства математического ожидания

Отметим важные свойства математического ожидания. Их доказательство для дискретных случайных величин можно провести ровно также, как это было сделано в случае конечного пространства Ω, используя свойства абсолютно сходящихся рядов. Доказательства для непрерывных величин во многом опираются на такие свойства интеграла, как, например, линейность.

Лемма 2.3.1 *Во всех написанных свойствах предполагается, что математическое ожидание существует.*

1. *Пусть $g(\xi)$ – случайная величина, построенная по случайной величине ξ. Тогда если ξ имеет дискретное распределение, то*

$$\mathsf{E}(g(\xi)) = \sum_{\omega\in\Omega} g(\xi(\omega))\mathsf{P}(\omega) = \sum_i g(x_i)\mathsf{P}(\xi = x_i),$$

а если абсолютно непрерывное, то

$$\mathsf{E}(g(\xi)) = \int\limits_{-\infty}^{+\infty} g(x)f_\xi(x)dx.$$

2. *Пусть* $\xi \geq 0$, *тогда* $\mathsf{E}\xi \geq 0$. *Иными словами, если случайная величина неотрицательна, то и ее среднее вероятностное неотрицательно.*

3. $\mathsf{E}(a\xi + b\eta) = a\mathsf{E}\xi + b\mathsf{E}\eta$. *Иными словами, математическое ожидание линейно.*

4. *Пусть* $\xi \geq \eta$ *(это значит, что* $\forall\omega \in \Omega$ *выполняется* $\xi(\omega) \geq \eta(\omega)$*), тогда* $\mathsf{E}\xi \geq \mathsf{E}\eta$. *Иными словами, математическое ожидание монотонно.*

5. $|\mathsf{E}\xi| \leq \mathsf{E}|\xi|$.

6. $(\mathsf{E}|\xi\eta|)^2 \leq \mathsf{E}\xi^2\mathsf{E}\eta^2$.

7. *Если* ξ, η *независимы, то* $\mathsf{E}(\xi\eta) = \mathsf{E}\xi\mathsf{E}\eta$.

8. *Если* $\xi = \mathsf{const}$, *то* $\mathsf{E}\xi = \mathsf{const}$.

Как мы уже сказали, мы не будем останавливаться на доказательствах этих утверждений.

Как и ранее, обратное утверждение, к утверждению 7, неверно. Из того, что равенство выполнено не следует, что величины независимы. Так как это было тщательно изучено нами ранее, то на этом мы тоже сейчас подробно останавливаться не будем.

2.4 Вычисление математического ожидания у некоторых распределений

Давайте вычислим математические ожидания тех «новых» распределений, которые нам встретились. Начнем с дискретных распределений, и, конкретно, с геометрического распределения.

Пример 2.4.1 (Геометрическое распределение) *Пусть случайная величина* ξ *имеет геометрическое распределение, задаваемое таблицей*

ξ	1	2	$\ldots$	n	$\ldots$
P	p	$(1-p)p$	$\ldots$	$(1-p)^{n-1}p$	$\ldots$

$$\mathsf{E}\xi = \sum_{n=1}^{\infty} n(1-p)^{n-1}p = p\sum_{n=1}^{\infty} n(1-p)^{n-1}.$$

Для вычисления суммы последнего ряда, воспользуемся свойствами степенных рядов ведь, как легко заметить,

$$\sum_{n=1}^{\infty} n(1-p)^{n-1} = \sum_{n=1}^{\infty} (x^n)'\bigg|_{x=1-p} = \left(\sum_{n=1}^{\infty} x^n\right)'\bigg|_{x=1-p} = \left(\frac{x}{1-x}\right)'\bigg|_{x=1-p} =$$

$$= \frac{1}{(1-x)^2}\bigg|_{x=1-p} = \frac{1}{p^2},$$

так как сумма ряда

$$\sum_{n=1}^{\infty} x^n$$

при $|x| < 1$*, по сути, сумма бесконечно убывающей геометрической прогрессии, и равна* $\frac{x}{1-x}$*. Итого,*

$$\mathsf{E}\xi = \frac{p}{p^2} = \frac{1}{p}.$$

Итак, как оказалось, математическое ожидание случайной величины, имеющей геометрическое распределение, обратно пропорционально вероятности успеха.

Пример 2.4.2 *В предыдущей лекции мы рассматривали следующую постановку задачи. В среднем 7% людей страдают сахарным диабетом. Допустим, что проводится опрос населения города для выявления людей с такой патологией. Каково среднее число людей, которых требуется опросить, чтобы выявить человека с данной патологией?*

Случайная величина ξ *– количество опрошенных людей до первого выявления больного этим заболеванием имеет геометрическое распределение* G_p *с параметром* $p = 0.07$*. Тогда* $\mathsf{E}\xi \approx 14.3$*. Тем самым, в среднем следует опросить* $14 - 15$ *человек.*

Вычислим математическое ожидание распределения Пуассона.

Пример 2.4.3 (Распределение Пуассона) *Пусть* $\xi \sim \Pi_\lambda$*,* $\lambda > 0$*, то есть*

$$\mathsf{P}(\xi = k) = \frac{\lambda^k}{k!}e^{-\lambda}, \quad k \in \{0, 1, 2, 3, ..., n, ...\}.$$

Тогда

$$\mathsf{E}\xi = \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}e^{\lambda} = \lambda,$$

так как

$$\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = e^{\lambda}.$$

Итак, параметр λ у распределения Пуассона и есть математическое ожидание случайной величины, имеющей такое распределение. Рассмотрим пример, в котором используем предельную теорему для схемы Бернулли – теорему Пуассона.

Пример 2.4.4 *Для продвижения своей продукции на рынок фирма раскладывает по почтовым ящикам рекламные листки. Прежний опыт работы показывает, что примерно в одном случае из* 2000 *следует заказ. Допустим, что размещено* 10000 *рекламных листков. Каково среднее количество заказов?*

Мы помним, что в теореме Пуассона $\lambda \approx p_n \cdot n$. *В нашем случае* $p_n = \frac{1}{2000}$, *а* $n = 10000$, *тогда* $\lambda = 5$.

Кстати, мы знаем, что в схеме Бернулли математическое ожидание числа успехов S_n *равно пр. Так что используя схему Бернулли мы получим ровно такой же результат.*

Теперь рассмотрим равномерное распределение. Наверное, ответ можно угадать, подумайте.

Пример 2.4.5 (Равномерное распределение) *Пусть* $\xi \sim \mathsf{U}_{a,b}$, *то есть*

$$f_\xi(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b] \\ 0, & x \notin [a,b] \end{cases}.$$

Тогда

$$\mathsf{E}\xi = \int\limits_{-\infty}^{+\infty} x f_\xi(x)dx = \int\limits_a^b \frac{x}{b-a}dx = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$$

Итак, у равномерно распределенной на отрезке случайной величины среднее значение находится в середине отрезка, на котором плотность отлична от нуля. Неудивительно, не так ли?

Пример 2.4.6 (Показательное распределение) *Пусть случайная величина* ξ *имеет показательное распределение* $\xi \sim \mathsf{Exp}_\lambda$, $\lambda > 0$, *то есть ее плотность задается соотношением*

$$f_\xi(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \geq 0 \end{cases}.$$

Тогда

$$\mathsf{E}\xi = \int\limits_{-\infty}^{+\infty} x f_\xi(x)dx = \lambda \int\limits_0^{+\infty} x e^{-\lambda x}dx = \lambda\left(-\frac{x}{\lambda}e^{-\lambda x}\Big|_0^{+\infty} + \frac{1}{\lambda}\int_0^{+\infty} e^{-\lambda x}dx \right) = \frac{1}{\lambda}.$$

Итого, у показательного распределения математическое ожидание обратно пропорционально параметру.

Перейдем к вычислению математического ожидания стандартного нормального распределения. Опять же, его достаточно легко угадать, вспомнив график плотности такого распределения.

Пример 2.4.7 (Стандартное нормальное распределение) *Пусть* $\xi \sim \mathsf{N}_{0,1}$, *то есть*

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Тогда

$$\mathsf{E}\xi = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} x e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} d\left(\frac{x^2}{2}\right) = -\frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\Big|_{-\infty}^{+\infty} = 0.$$

Итак, у случайной величины со стандартным нормальным распределением математическое ожидание равно 0.

Пример 2.4.8 (Нормальное распределение) *Пусть* $\xi \sim \mathsf{N}_{a,\sigma^2}$, *то есть*

$$f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Как мы знаем, случайная величина ξ *может быть представлена, как* $\xi = \sigma\eta + a$, *где* $\eta \sim \mathsf{N}_{0,1}$. *Согласно свойствам математического ожидания,*

$$\mathsf{E}\xi = \mathsf{E}(\sigma\eta + a) = \sigma\mathsf{E}\eta + a = a.$$

Значит, математическое ожидание случайной величины ξ *равно* a.

Математическое ожидание определено не всегда. Давайте приведем примеры таких случайных величин, у которых нет математического ожидания.

Пример 2.4.9 (Распределение Коши) *Говорят, что случайная величина* ξ *имеет распределение Коши, если ее плотность имеет вид*

$$f_\xi(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}.$$

Легко проверить, что написанная функция действительно задает плотность вероятности. Попробуем вычислить математическое ожидание случайной величины ξ.

$$\mathsf{E}\xi = \frac{1}{\pi} \int\limits_{-\infty}^{+\infty} \frac{x}{1+x^2} dx = \frac{1}{2\pi} \ln(1+x^2)\Big|_{-\infty}^{+\infty},$$

но последнего предела не существует. Значит, нет и математического ожидания.

Расходимость интеграла легко установить и из того, что подынтегральная функция эквивалентна функции $\frac{1}{x}$ *при* $x \to +\infty$.

Конечно, можно привести и пример случайной величины с дискретным распределением, не имеющей математического ожидания.

Пример 2.4.10 *Предположим, что случайная величина ξ принимает натуральные значения с вероятностями*

$$\mathsf{P}(\xi = k) = \frac{1}{k(k+1)}, \quad k \in \{1, 2, ...\}.$$

Проверим вначале, что такие вероятности задают распределение случайной величины. Для этого требуется проверить, что их сумма равна 1. Действительно,

$$\sum_{k=1}^{+\infty} \frac{1}{k(k+1)} = \lim_{n\to+\infty} \sum_{k=1}^{n} \frac{1}{k(k+1)} = \lim_{n\to+\infty} \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right) =$$

$$= \lim_{n\to+\infty} \left(1 - \frac{1}{n+1}\right) = 1.$$

В свою очередь, математическое ожидание равно следующему выражению

$$\mathsf{E}\xi = \sum_{k=1}^{+\infty} \frac{k}{k(k+1)} = \sum_{k=1}^{+\infty} \frac{1}{k+1},$$

но последний ряд, очевидно, расходится, как гармонический.

2.5 Дисперсия и моменты старших порядков

Мы уже знакомы с понятием дисперсии из простейшего случая. Посмотрим, как определяются моменты (и центральные моменты) старших порядков. Введем следующие определения в предположении, что существует $\mathsf{E}|\xi|^k$.

Определение 2.5.1 *Число $\mathsf{E}\xi^k$ называется моментом k-ого порядка, или k-ым моментом случайной величины ξ.*

Определение 2.5.2 *Число $\mathsf{E}|\xi|^k$ называется абсолютным моментом k-ого порядка, или абсолютным k-ым моментом случайной величины ξ.*

Определение 2.5.3 *Число $\mathsf{E}(\xi - \mathsf{E}\xi)^k$ называется центральным моментом k-ого порядка, или центральным k-ым моментом случайной величины ξ.*

Определение 2.5.4 *Число $\mathsf{E}|\xi - \mathsf{E}\xi|^k$ называется абсолютным центральным моментом k-ого порядка, или абсолютным центральным k-ым моментом случайной величины ξ.*

За что отвечают эти моменты? Рассмотрим пример, который нам давно известен, и на котором мы уже обсуждали понятия как математического ожидания, так и дисперсии. Пусть ξ задана таблицей распределения

ξ	0	100
P	$\frac{99}{100}$	$\frac{1}{100}$

Тогда, конечно,

$$\mathsf{E}\xi = 0 \cdot \frac{99}{100} + 100 \cdot \frac{1}{100} = 1,$$
$$\mathsf{E}\xi^2 = 0^2 \cdot \frac{99}{100} + 100^2 \cdot \frac{1}{100} = 100,$$
$$\mathsf{E}\xi^3 = 0^3 \cdot \frac{99}{100} + 100^3 \cdot \frac{1}{100} = 100^2,$$

да и вообще, $\mathsf{E}\xi^k = 100^{k-1}$. То есть чем старше момент, тем сильнее он реагирует на большие, но маловероятные значения.

Отметим важный технический факт, а именно.

Теорема 2.5.1 *Пусть существует момент k-ого порядка случайной величины ξ. Тогда при $0 < r < k$ существует и момент r-ого порядка*

Доказательство. Доказательство этого факта базируется на очевидном неравенстве

$$|x|^r \le \max\{|x|^k, 1\} \le 1 + |x|^k.$$

Используя свойство монотонности математического ожидания, подставив вместо x случайную величину ξ, получим

$$\mathsf{E}|\xi|^r \le \mathsf{E}|\xi|^k + 1,$$

что и доказывает утверждение. □

Определение 2.5.5 *Центральный момент второго порядка называется дисперсией случайной величины ξ. Как и раньше,*

$$\mathsf{D}\xi = \mathsf{E}(\xi - \mathsf{E}\xi)^2.$$

Как уже отмечалось, дисперсия показывает степень разброса случайной величины вокруг ее математического ожидания. Из предыдущего утверждения следует, что из существования второго момента следует существование как математического ожидания, так и дисперсии. Параллельно с дисперсией, как мы знаем, рассматривают среднее квадратическое отклонение.

Определение 2.5.6 *Пусть существует дисперсия* $\mathsf{D}\xi$ *случайной величины* ξ. *Тогда величина*

$$\sigma_\xi = \sqrt{\mathsf{D}\xi}$$

называется среднеквадратическим отклонением.

Повторим уже отмеченные ранее свойства дисперсии. Они доказываются совершенно аналогично с использованием свойств математического ожидания.

Лемма 2.5.1 *Во всех свойствах предполагается, что вторые моменты случайных величин существуют.*

1. $\mathsf{D}\xi \geq 0$.
2. $\mathsf{D}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2$
3. $\mathsf{D}(\xi + a) = \mathsf{D}\xi$.
4. $\mathsf{D}(c\xi) = c^2\mathsf{D}\xi$.
5. *Если* ξ *и* η *– независимые случайные величины, то* $\mathsf{D}(\xi+\eta) = \mathsf{D}\xi + \mathsf{D}\eta$.
6. *Для произвольных случайных величин* ξ *и* η *с конечными дисперсиями справедливо равенство*

$$\mathsf{D}(\xi + \eta) = \mathsf{D}\xi + \mathsf{D}\eta + 2\left(\mathsf{E}(\xi\eta) - \mathsf{E}\xi\mathsf{E}\eta\right).$$

2.6 Вычисление дисперсий некоторых распределений

Вслед за математическим ожиданием, вычислим дисперсии «новых» рассмотренных распределений.

Пример 2.6.1 (Геометрическое распределение) *Пусть случайная величина* ξ *имеет геометрическое распределение, задаваемое таблицей*

ξ	1	2	$\ldots$	n	$\ldots$
P	p	$(1-p)p$	$\ldots$	$(1-p)^{n-1}p$	$\ldots$

Напомним, что $\mathsf{E}\xi = \frac{1}{p}$. *Вычислим дисперсию. Для вычисления величины* $\mathsf{E}\xi^2$ *бывает удобно вычислить* $\mathsf{E}\left(\xi(\xi-1)\right)$ *и воспользоваться тождеством*

$$\mathsf{E}\xi^2 = \mathsf{E}\left(\xi(\xi-1)\right) + \mathsf{E}\xi.$$

Итак, снова используя свойства степенных рядов,

$$\mathsf{E}\left(\xi(\xi-1)\right)=\sum_{k=1}^{\infty}k(k-1)(1-p)^{k-1}p=p(1-p)\sum_{k=0}^{\infty}k(k-1)(1-p)^{k-2}=$$

$$=p(1-p)\sum_{k=0}^{\infty}\left(x^k\right)''\Big|_{x=1-p}=p(1-p)\left(\sum_{k=0}^{\infty}x^k\right)''\Bigg|_{x=1-p}=$$

$$=p(1-p)\left(\frac{1}{1-x}\right)''\Bigg|_{x=1-p}=\frac{2(1-p)}{p^2}.$$

Значит,

$$\mathsf{E}\xi^2=\frac{2(1-p)}{p^2}+\frac{1}{p}=\frac{2-p}{p^2}.$$

Тогда

$$\mathsf{D}\xi=\mathsf{E}\xi^2-(\mathsf{E}\xi)^2=\frac{2-p}{p^2}-\frac{1}{p^2}=\frac{1-p}{p^2}.$$

Пример 2.6.2 *Напомним условие задачи, которую мы уже рассматривали. В среднем 7% людей страдают сахарным диабетом. Допустим, что проводится опрос населения города для выявления людей с такой патологией. Случайная величина ξ – количество опрошенных людей до первого выявления больного этим заболеванием. Найдем дисперсию ξ. Так как ξ имеет геометрическое распределение G_p с параметром $p=0.07$, то, согласно только что сказанному,*

$$\mathsf{D}\xi=\frac{1-0.07}{0.07^2}\approx 189.8,$$

а тогда среднее квадратическое отклонение равно $\sigma_\xi=\sqrt{\mathsf{D}\xi}\approx 13.8$. Напомним, что среднее число опрошенных $\mathsf{E}\xi$ равно 14.3, тогда, грубо говоря, можно утверждать, что в среднем значения случайной величины ξ находятся в интервале

$$(\mathsf{E}\xi-\sigma_\xi,\mathsf{E}\xi+\sigma_\xi)\approx(14.3-13.8,14.3+13.8)=(0.5,28.1)\,.$$

Вычислим дисперсию случайной величины, имеющей распределения Пуассона.

Пример 2.6.3 (Распределение Пуассона) *Пусть $\xi\sim\Pi_\lambda$, $\lambda>0$, то есть*

$$\mathsf{P}(\xi=k)=\frac{\lambda^k}{k!}e^{-\lambda},\quad k\in\{0,1,2,3,...,n,...\}.$$

Как мы знаем, $\mathsf{E}\xi = \lambda$. *Вычислим* $\mathsf{E}\xi^2$ *аналогичным образом, сначала вычислив*

$$\mathsf{E}\left(\xi(\xi-1)\right) = \sum_{k=0}^{\infty} k(k-1)\frac{\lambda^k}{k!}e^{-\lambda} = \lambda^2 e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!} = \lambda^2 e^{-\lambda}e^{\lambda} = \lambda^2.$$

Значит,

$$\mathsf{D}\xi = \mathsf{E}\left(\xi(\xi-1)\right) + \mathsf{E}\xi - (\mathsf{E}\xi)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Итак, у распределения Пуассона совпадают математическое ожидание и дисперсия.

Пример 2.6.4 (Равномерное распределение) *Пусть* $\xi \sim \mathsf{U}_{a,b}$, *то есть*

$$f_\xi(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b] \\ 0, & x \notin [a,b] \end{cases}.$$

мы знаем, что $\mathsf{E}\xi = \frac{a+b}{2}$. *Вычислим* $\mathsf{E}\xi^2$, *тогда*

$$\mathsf{E}\xi^2 = \int\limits_{-\infty}^{+\infty} x^2 f_\xi(x)dx = \int\limits_a^b \frac{x^2}{b-a}dx = \frac{b^2+ab+a^2}{3}.$$

Отсюда следует, что

$$\mathsf{D}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2 = \frac{b^2+ab+a^2}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}.$$

На рисунке **??** красной точкой обозначено математическое ожидание случайной величины ξ, имеющей равномерное распределение $\mathsf{U}_{a,b}$. Зелеными точками показаны концы интервала $(\mathsf{E}\xi - \sigma_\xi,\ \mathsf{E}\xi + \sigma_\xi)$.

Пример 2.6.5 (Показательное распределение) *Пусть случайная величина* ξ *имеет показательное распределение* $\xi \sim \mathsf{Exp}_\lambda$, $\lambda > 0$, *то есть ее плотность задается соотношением*

$$f_\xi(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \geq 0 \end{cases}.$$

Мы знаем, что $\mathsf{E}\xi = \frac{1}{\lambda}$. *Значит,*

$$\mathsf{E}\xi^2 = \int\limits_{-\infty}^{+\infty} x^2 f_\xi(x)dx = \lambda\int\limits_0^{+\infty} x^2 e^{-\lambda x}dx = -x^2e^{-\lambda x}\Big|_0^{+\infty} + 2\int\limits_0^{+\infty} xe^{-\lambda x}dx = \frac{2}{\lambda^2},$$

так как последний интеграл равен $\frac{\mathsf{E}\xi}{\lambda} = \frac{1}{\lambda^2}$. *Тогда*

$$\mathsf{D}\xi = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

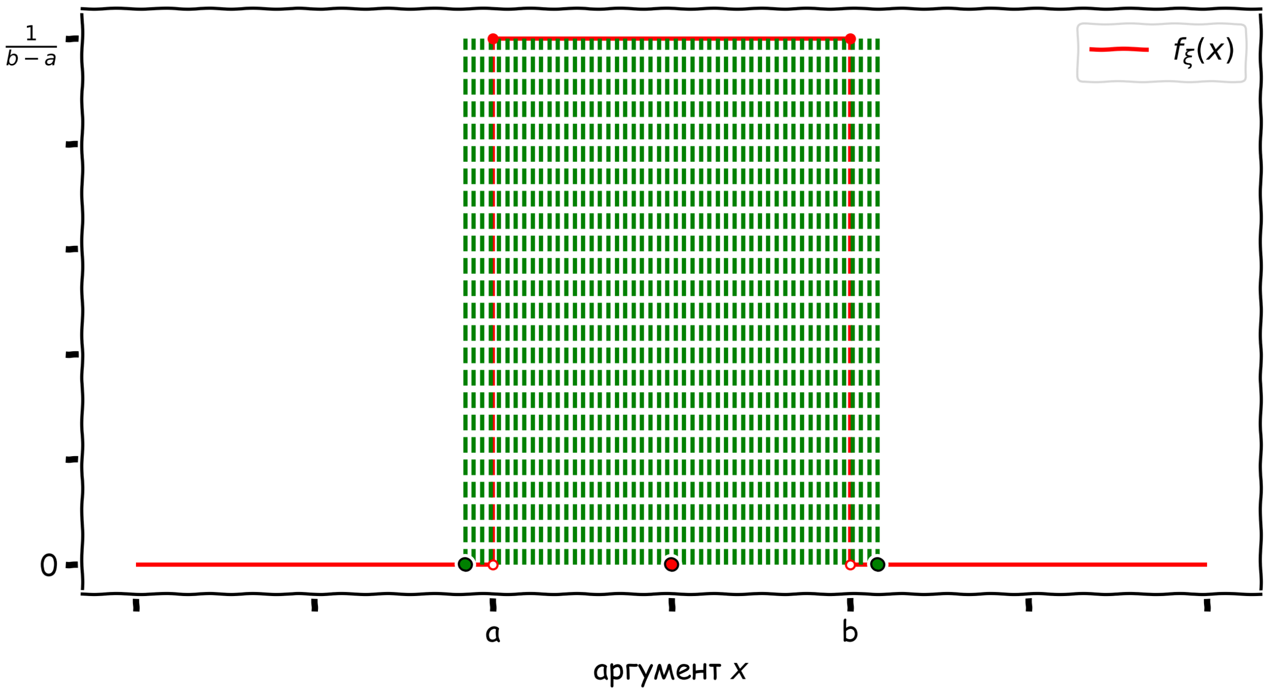

Рис. 1: Математическое ожидание и среднеквадратическое отклонение

На рисунке ?? красной точкой обозначено математическое ожидание случайной величины ξ, имеющей показательное распределение Exp_λ с параметром λ. Зелеными точками показаны концы интервала $(\mathsf{E}\xi - \sigma_\xi,\ \mathsf{E}\xi + \sigma_\xi)$.

Пример 2.6.6 (Стандартное нормальное распределение) *Пусть* $\xi \sim \mathsf{N}_{0,1}$, *то есть*

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Мы уже знаем, что $\mathsf{E}\xi = 0$. *Тогда*

$$\mathsf{E}\xi^2 = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{2}} dx = -\frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} x d\left(e^{-\frac{x^2}{2}}\right) =$$

$$-\frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Big|_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = 1.$$

Значит,

$$\mathsf{D}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2 = 1.$$

Итак, дисперсия стандартного нормального распределения равна 1.

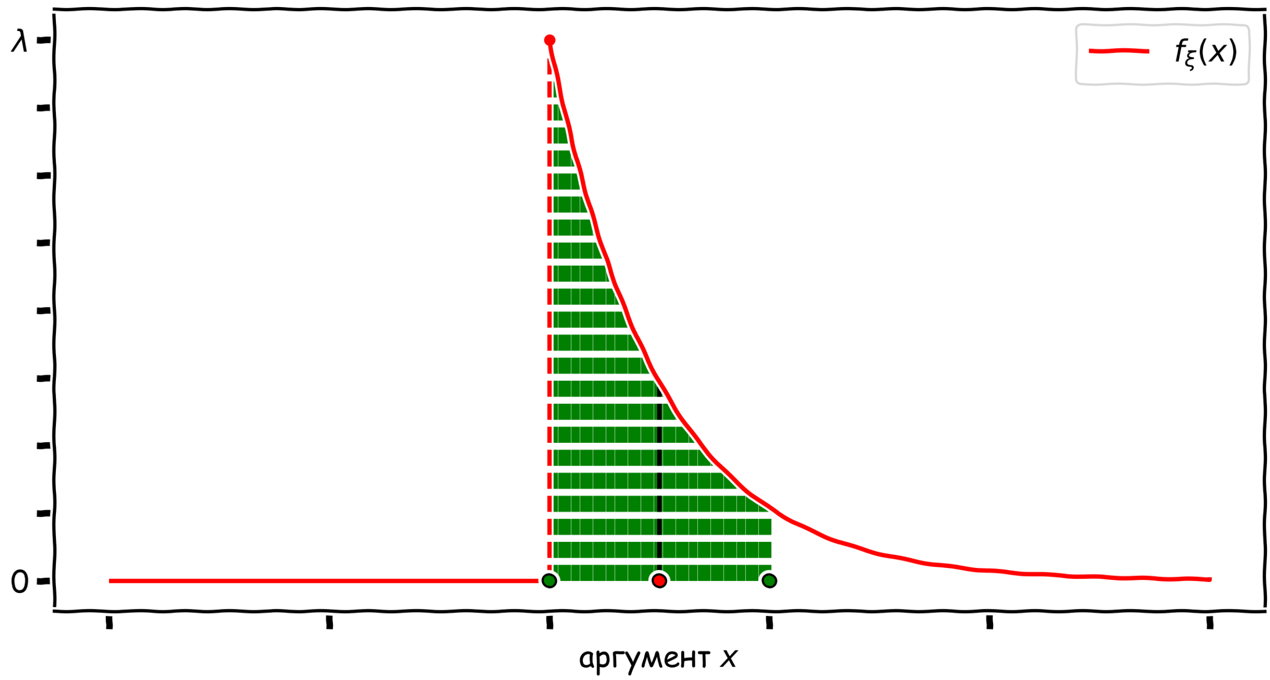

Рис. 2: Математическое ожидание и среднеквадратическое отклонение

Пример 2.6.7 (Нормальное распределение) *Пусть $\xi \sim \mathsf{N}_{a,\sigma^2}$, то есть*

$$f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Как мы знаем, случайная величина ξ может быть представлена, как $\xi = \sigma\eta + a$, где $\eta \sim \mathsf{N}_{0,1}$. Согласно свойствам дисперсии,

$$\mathsf{D}\xi = \mathsf{D}(\sigma\eta + a) = \sigma^2 \mathsf{D}\eta = \sigma^2.$$

Значит, второй параметр нормального распределения – это дисперсия.

На рисунке **??** красной точкой обозначено математическое ожидание случайной величины ξ, имеющей нормальное распределение N_{a,σ^2}. Зелеными точками показаны концы интервала $(\mathsf{E}\xi - \sigma_\xi,\ \mathsf{E}\xi + \sigma_\xi)$.

2.7 Ковариация и корреляция

Понятия ковариации и корреляции переносятся с простейшего случая абсолютно аналогичным образом, так что все дальнейшее нам прекрасно известно. Давайте конспективно пройдемся по основным моментам

Определение 2.7.1 *Величина $\mathsf{E}(\xi\eta) - \mathsf{E}\xi\mathsf{E}\eta$ называется ковариацией случайных величин ξ и η и обозначается $\mathsf{cov}(\xi, \eta)$.*

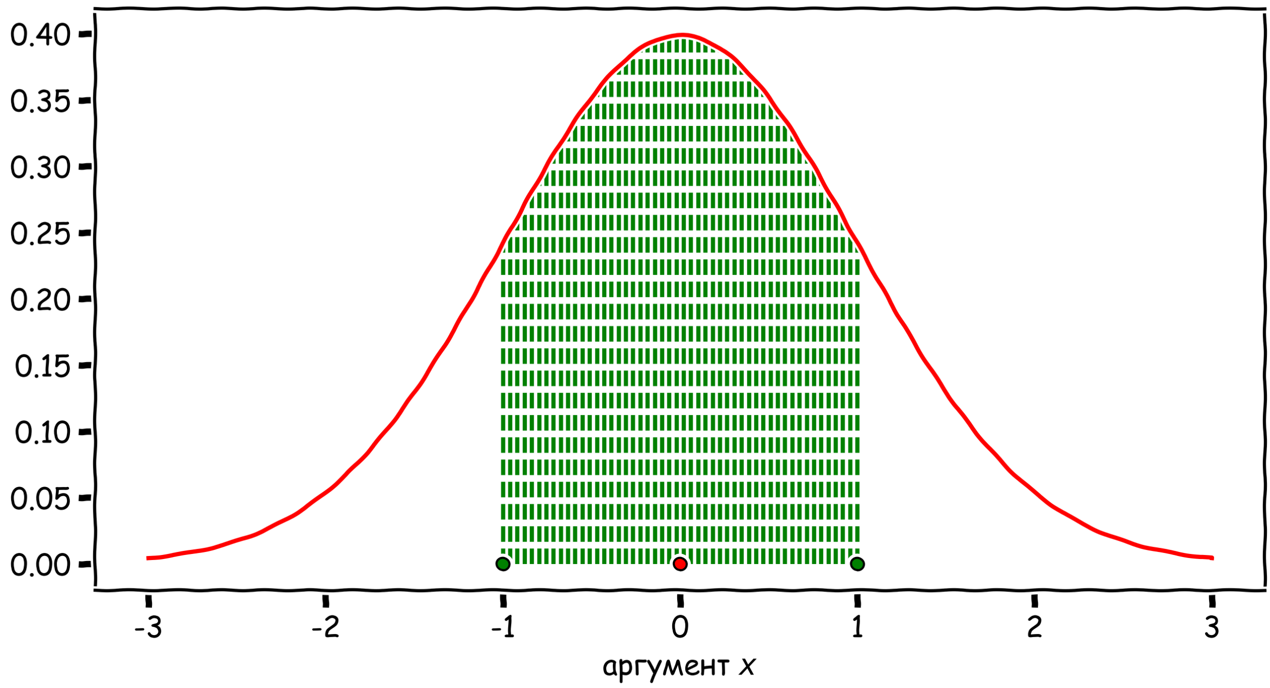

Рис. 3: Математическое ожидание и среднеквадратическое отклонение

Замечание 2.7.1 *Прямые выкладки показывают, что ковариация может быть определена и следующим способом:*

$$\mathsf{cov}(\xi, \eta) = \mathsf{E}\left((\xi - \mathsf{E}\xi)(\eta - \mathsf{E}\eta)\right).$$

Отметим свойства ковариации, моментально следующие из ее определения и свойств математического ожидания.

Лемма 2.7.1 *Ковариация обладает следующими свойствами.*

1. $\mathsf{cov}(\xi, \eta) = \mathsf{cov}(\eta, \xi)$;
2. $\mathsf{cov}(a\xi, b\eta) = ab \cdot \mathsf{cov}(\xi, \eta)$;
3. $\mathsf{cov}(\xi + c, \eta + d) = \mathsf{cov}(\xi, \eta)$;
4. $\mathsf{cov}(\xi, \xi) = \mathsf{D}\xi$.

Абсолютно аналогично мотивируется введение корреляции – она оказывается безразмерной, в отличие от ковариации.

Определение 2.7.2 *Коэффициентом корреляции двух величин ξ и η с отличными от нуля дисперсиями, называется величина*

$$\rho(\xi, \eta) = \frac{\mathsf{cov}(\xi, \eta)}{\sqrt{\mathsf{D}\xi}\sqrt{\mathsf{D}\eta}} = \frac{\mathsf{cov}(\xi, \eta)}{\sigma_\xi \sigma_\eta}$$

Как и раньше, коэффициент корреляции показывает степень линейной зависимости. Чем ближе его значение к единице, тем «более линейна» зависимость.

Лемма 2.7.2 *Коэффициент корреляции обладает следующими свойствами:*

1. $|\rho(\xi, \eta)| \leq 1$;
2. *Если* ξ, η *– независимы, то* $\rho(\xi, \eta) = 0$;
3. $|\rho(\xi, \eta)| = 1$ *тогда и только тогда, когда* $\xi = a\eta + b$, *причем* $a \cdot \rho(\xi, \eta) > 0$.

Опять же, мы не останавливаемся на детальных пояснениях, они даны во второй лекции.

3 Сходимость последовательностей случайных величин

Все предельные теоремы основаны, конечно же, на предельном переходе. В то же время, мы теперь работаем со случайными величинами – функциями на пространстве элементарных исходов. И что понимать под сходимостью функций? Давайте разбираться.

3.1 Различные типы сходимостей и их связи

Для того чтобы сформулировать закон больших чисел и центральную предельную теорему, нам понадобится (в отличие от рассмотренного ранее простейшего случая) ввести понятия различных сходимостей последовательностей случайных величин. Первый тип сходимости нам уже знаком из закона больших чисел.

Определение 3.1.1 (Сходимость по вероятности) *Говорят, что последовательность случайных величин* ξ_n, *заданных на вероятностном пространстве* $(\Omega, \Sigma, \mathsf{P})$, *сходится к случайной величине* ξ, *заданной на том же вероятностном пространстве, по вероятности, если*

$$\forall \varepsilon > 0 \quad \lim_{n \to \infty} \mathsf{P}\left(|\xi_n - \xi| > \varepsilon\right) = 0.$$

Иными словами, вероятность события $\{\omega \in \Omega : \ |\xi_n(\omega) - \xi(\omega)| > \varepsilon\}$ стремится к нулю с ростом n для каждого наперед заданного $\varepsilon > 0$. Сходимость по вероятности часто обозначают следующим образом

$$\xi_n \xrightarrow[n \to +\infty]{\mathsf{P}} \xi.$$

Отметим свойства сходимости по вероятности, который мы будем часто использовать в дальнейшем.

Лемма 3.1.1 *Пусть* $\xi_n \xrightarrow[n\to+\infty]{\mathsf{P}} \xi$, $\eta_n \xrightarrow[n\to+\infty]{\mathsf{P}} \eta$, *тогда*

1. $\xi_n + \eta_n \xrightarrow[n\to+\infty]{\mathsf{P}} \xi + \eta$.

2. $\xi_n \eta_n \xrightarrow[n\to+\infty]{\mathsf{P}} \xi\eta$.

3. *Пусть функция* $g(x)$ *непрерывна, тогда* $g(\xi_n) \xrightarrow[n\to+\infty]{\mathsf{P}} g(\xi)$

Эти свойства нам вполне привычны. Если говорить кратко, то они утверждают, что предел суммы равен сумме пределов, а предел произведения – произведению пределов. Кроме того, предел «проносится» через непрерывную функцию.

Из курса анализа нам знакомо понятие поточечной сходимости, а именно: $\xi_n \xrightarrow[n\to+\infty]{} \xi$ поточечно на множестве Ω, если

$$\forall \omega \in \Omega \;\Rightarrow\; \xi_n(\omega) \xrightarrow[n\to+\infty]{} \xi(\omega).$$

Поточечная сходимость означает сходимость соответствующей числовой последовательности в каждой точке множества Ω. В теории вероятностей логичнее рассматривать ослабленную поточечную сходимость.

Определение 3.1.2 (Сходимость почти наверное) *Говорят, что последовательность случайных величин* ξ_n*, заданных на вероятностном пространстве* $(\Omega, \Sigma, \mathsf{P})$*, сходится к случайной величине* ξ*, заданной на том же вероятностном пространстве, почти наверное (или с вероятностью* 1*), если*

$$\mathsf{P}\left(\omega \in \Omega : \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\right) = 1.$$

Последнее определение, как уже говорилось, расширяет понятие поточечной сходимости функциональной последовательности. Ведь при каждом $\omega \in \Omega$ последовательность $\xi_n(\omega)$ является числовой и может как сходиться, так и нет, как числовая последовательность, к $\xi(\omega)$. Поточечная сходимость требует наличие сходимости для каждого $\omega \in \Omega$. Так как в теории вероятностей событиями, имеющими нулевую вероятность, часто можно пренебречь, то резонно несколько ослабить понятие поточечной сходимости. Тут и возникает сходимость почти наверное. Для обозначения сходимости почти наверное используют обозначение

$$\xi_n \xrightarrow[n\to+\infty]{\text{п. н.}} \xi.$$

Как связаны сходимости почти наверное и по вероятности? Оказывается, сходимость почти наверное сильнее и влечет сходимость по вероятности, а именно справедлива следующая теорема.

Теорема 3.1.1 *Из сходимости почти наверное следует сходимость по вероятности, то есть*

$$\xi_n \xrightarrow[n\to+\infty]{\text{п. н.}} \xi \;\Rightarrow \xi_n \xrightarrow[n\to+\infty]{\mathsf{P}} \xi$$

Можно показать, что обратное утверждение не имеет места. Мы не будем доказывать эту теорему, так как она требует знаний из теории меры.

Следующий тип сходимости мы тоже встречали, хотя так его и не называли, когда говорили об интегральной теореме Муавра-Лапласа.

Определение 3.1.3 (Сходимость по распределению) *Говорят, что последовательность случайных величин ξ_n, заданных на вероятностном пространстве $(\Omega, \Sigma, \mathsf{P})$, сходится к случайной величине ξ, заданной на том же вероятностном пространстве, по распределению (слабо), если*

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

во всех точках x, в которых предельная функция $F_\xi(x)$ непрерывна.

Сходимость по распределению часто обозначают следующим образом:

$$\xi_n \xrightarrow[n\to+\infty]{\mathsf{d}} \xi.$$

Оказывается, сходимость по вероятности влечет сходимость по распределению.

Теорема 3.1.2 *Из сходиомости по вероятности следует сходимость по распределению, то есть*

$$\xi_n \xrightarrow[n\to+\infty]{\mathsf{P}} \xi \;\Rightarrow\; \xi_n \xrightarrow[n\to+\infty]{\mathsf{d}} \xi.$$

В этом случае тоже можно показать, что обратное утверждение в общем случае не имеет места. Однако, если $\xi_n \xrightarrow[n\to+\infty]{\mathsf{d}} \xi = \mathsf{const}$, то $\xi_n \xrightarrow[n\to+\infty]{\mathsf{P}} \xi = \mathsf{const}$. Итак, слабая сходимость к постоянной влечет сходимость по вероятности.

В итоге можно сказать, самой «сильной» в рассмотренной цепочке оказывается сходимость почти наверное, затем идет сходимость по вероятности, а потом – сходимость по распределению.

Напоследок отметим важные свойства слабой сходимости.

Лемма 3.1.2 *Пусть* $\eta_n \xrightarrow[n\to+\infty]{\mathsf{d}} \eta$ *и* $\xi_n \xrightarrow[n\to+\infty]{\mathsf{P}} \xi = \mathsf{const}$, *тогда*

$$\xi_n \cdot \eta_n \xrightarrow[n\to+\infty]{\mathsf{d}} \mathsf{const} \cdot \eta.$$

и

$$\xi_n + \eta_n \xrightarrow[n\to+\infty]{\mathsf{d}} \mathsf{const} + \eta.$$

Описанными типами сходимостей, а также их свойствами, мы будем постоянно пользоваться при изложении математической статистики, так что советуем крепко их усвоить.

3.2 Неравенства Маркова и Чебышёва

Аналогично тому, что было сделано в простейшем случае, сформулируем неравенства Маркова и Чебышева.

Теорема 3.2.1 (Неравенство Маркова) *Пусть случайная величина* $\xi \geq 0$ *и существует* $\mathsf{E}\xi$. *Тогда для каждого* $a > 0$

$$\mathsf{P}\left(\xi \geq a\right) \leq \frac{\mathsf{E}\xi}{a}.$$

Неравенство Маркова в общем случае удобно доказывать с помощью так называемых индикаторов, доказательство можно найти в дополнительных материалах. Ясно, что если допустить, что математическое ожидание случайной величины может равняться $+\infty$, то последнее неравенство все равно останется справедливым. Из неравенства Маркова вытекает хорошо знакомое нам неравенство Чебышёва.

Теорема 3.2.2 (Неравенство Чебышёва) *Пусть существует момент второго порядка* $\mathsf{E}\xi^2$. *Тогда для каждого* $a > 0$

$$\mathsf{P}\left(|\xi - \mathsf{E}\xi| \geq a\right) \leq \frac{\mathsf{D}\xi}{a^2}.$$

Неравенство Чебышёва нам тоже хорошо знакомо по дискретному случаю.

3.3 Закон больших чисел

В предыдущих лекциях мы уже встречались с законом больших чисел и мотивировали его важность. Напомним, что теоремы типа ЗБЧ устанавливают сходимость среднего арифметического независимых и одинаково распределенных случайных величин к их среднему (математическому ожиданию). Именно эти теоремы математически обосновывают следующее умозаключение: если мы в одинаковых условиях проводим много раз один и тот же эксперимент по измерению какого-то значения, в каждом эксперименте есть случайная ошибка, а ошибки независимы и одинаково распределены, то при большом числе измерений среднее арифметическое наших измерений будет сколько угодно близко к истинному «среднему» значению.

Отметим несколько формулировок законов больших чисел, которые будем использовать в дальнейшем в математической статистике. Закон больших чисел в форме Чебышёва мы уже встречали ранее, сформулируем его. В дальнейшем мы будем классически использовать обозначение

$$S_n = \xi_1 + \xi_2 + ... + \xi_n.$$

Теорема 3.3.1 (ЗБЧ в форме Чебышёва) *Пусть $\xi_1, ..., \xi_n, ...$ – последовательность попарно независимых и одинаково распределенных случайных величин, причем существует второй момент $\mathsf{E}\xi_1^2$. Тогда*

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} \mathsf{P}\left(\left|\frac{S_n}{n} - \mathsf{E}\xi_1\right| > \varepsilon\right) = 0.$$

Иными словами, среднее арифметическое попарно независимых и одинаково распределенных случайных величин сходится к математическому ожиданию величины ξ_1 по вероятности. Доказательство этой теоремы опирается на неравенство Чебышёва и ничем не отличается от проведенного ранее. Кстати, можно заметить, что требование попарной независимости можно заменить на требование попарной некоррелированности. Доказательство при этом не изменится.

Доказательство приводимых далее законов больших чисел достаточно трудно и, в принципе, не проливает много света на саму идею этой группы теорем, поэтому доказательства опустим, но формулировки обязательно приведем.

Теорема 3.3.2 (ЗБЧ в форме Хинчина) *Пусть $\xi_1, ..., \xi_n, ...$ – последовательность независимых в совокупности и одинаково распределенных случайных величин, причем существует первый момент $\mathsf{E}\xi_1$. Тогда*

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} \mathsf{P}\left(\left|\frac{S_n}{n} - \mathsf{E}\xi_1\right| > \varepsilon\right) = 0.$$

Обратите внимание, мы ослабили условие на момент, но усилили требование независимости: теперь она в совокупности, а не попарная.

Следующую формулировку закона больших чисел часто называют усиленным законом больших чисел.

Теорема 3.3.3 (ЗБЧ в форме Колмогорова) *Пусть* $\xi_1, ..., \xi_n, ...$ *– последовательность независимых в совокупности и одинаково распределенных случайных величин, причем существует первый момент* $\mathsf{E}\xi_1$. *Тогда*

$$\mathsf{P}\left(\omega \in \Omega : \lim_{n\to\infty} \frac{S_n}{n} = \mathsf{E}\xi_1\right) = 1.$$

Тем самым, ЗБЧ в форме Колмогорова устанавливает, что среднее арифметическое сходится к математическому ожиданию не только по вероятности, но и почти наверное. Как мы знаем, сходимость почти наверное сильнее, чем сходимость по вероятности. Именно поэтому данная теорема часто носит название «усиленного» закона больших чисел.

3.4 Центральная предельная теорема

Как мы знаем, среднее арифметическое попарно независимых случайных величин $\xi_1, ..., \xi_n$ с конечным вторым моментом сходится по вероятности к $\mathsf{E}\xi_1$. Иными словами,

$$\frac{S_n}{n} - \mathsf{E}\xi_1 = \frac{S_n - n\mathsf{E}\xi_1}{n} \xrightarrow[n\to+\infty]{\mathsf{P}} 0.$$

Можно ли найти такую функцию $f(n)$, что при умножении на эту функцию сходимость будет не к нулю, а к какому-то числу, отличному от нуля, то есть

$$f(n) \cdot \frac{S_n - n\mathsf{E}\xi_1}{n} \to a \neq 0, \quad a \in \mathbb{R}?$$

Найдя такую функцию $f(n)$, мы сможем говорить о так называемой скорости сходимости к нулю. Центральная предельная теорема говорит о том, что в качестве $f(n)$ можно взять $f(n) = \sqrt{n}$.

Теорема 3.4.1 (ЦПТ в форме Ляпунова) *Пусть* $\xi_1, ..., \xi_n, ...$ *– независимые в совокупности, одинаково распределенные случайные величины, у которых существует второй момент, а дисперсия отлична от нуля. Тогда имеет место слабая сходимость*

$$\frac{S_n - n\mathsf{E}\xi_1}{\sqrt{n\mathsf{D}\xi_1}} \xrightarrow[n\to+\infty]{\mathsf{d}} \mathsf{N}_{0,1}$$

Итак, согласно центральной предельной теореме, «центрированная» (так как вычитается математическое ожидание S_n) и «нормированная» (так как делится на дисперсию S_n) случайная величина S_n сходится слабо к стандартному нормальному распределению.

Как мы уже говорили, теорема Муавра-Лапласа является частным случаем ЦПТ. И правда, напомним, что если S_n – количество успехов в схеме Бернулли из n испытаний, то справедливо представление

$$S_n = \xi_1 + \xi_2 + ... + \xi_n,$$

где случайные величины ξ_i независимы, одинаково распределены и имеют распределение Бернулли B_p, то есть принимают значение 1 с вероятностью p и ноль с вероятностью $1-p$, а смысл соответствующей величины ξ_i – это успех или неудача в i-ом испытании. Заметим, что $\mathsf{E}\xi_1 = p$, $\mathsf{D}\xi_1 = pq$, $q = 1 - p$, а тогда, согласно ЦПТ,

$$\frac{S_n - n\mathsf{E}\xi_1}{\sqrt{n\mathsf{D}\xi_1}} \xrightarrow[n\to+\infty]{\mathsf{d}} \mathsf{N}_{0,1},$$

откуда

$$\mathsf{P}\left(a \le \frac{S_n - n\mathsf{E}\xi_1}{\sqrt{n\mathsf{D}\xi_1}} \le b\right) = \mathsf{P}\left(a \le \frac{S_n - np}{\sqrt{npq}} \le b\right) \xrightarrow[n\to+\infty]{} \Phi_{0,1}(b) - \Phi_{0,1}(a) =$$

$$= \frac{1}{\sqrt{2\pi}} \int\limits_a^b e^{-\frac{x^2}{2}} dx.$$

Многочисленные другие приложения ЦПТ мы увидим практически во всех лекциях по математической статистике. Ну что, теперь весь аппарат у нас готов, и мы готовы к ней приступить к статистике.