

Лекции 7-8

Вероятностное моделирование с применением распределений вероятности

Распределение вероятностей

Распределение вероятностей — это закон, описывающий область значений случайной величины и вероятности их принятия.

Определение

Определение 1. Пусть задано вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$, и на нём определена случайная величина $X : \Omega \rightarrow \mathbb{R}$. В частности, по определению, X является измеримым отображением измеримого пространства (Ω, \mathcal{F}) в измеримое пространство $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, где $\mathcal{B}(\mathbb{R})$ обозначает борелевскую сигма-алгебру на \mathbb{R} . Тогда случайная величина X индуцирует вероятностную меру \mathbb{P}^X на \mathbb{R} следующим образом:

$$\mathbb{P}^X(B) = \mathbb{P}(X^{-1}(B)), \quad \forall B \in \mathcal{B}(\mathbb{R}).$$

Мера \mathbb{P}^X называется **распределением** случайной величины X . Иными словами, $\mathbb{P}^X(B) = \mathbb{P}(X \in B)$, таким образом $\mathbb{P}^X(B)$ задаёт вероятность того, что случайная величина X попадает во множество $B \in \mathcal{B}(\mathbb{R})$.

Способы задания распределений

Определение 2. Функция $F_X(x) = \mathbb{P}^X((-\infty, x]) = \mathbb{P}(X \leq x)$ называется (кумулятивной) функцией распределения случайной величины X . Из свойств вероятности вытекает

Теорема 1. Функция распределения $F_X(x)$ любой случайной величины удовлетворяет следующим трем свойствам:

1. F_X — функция неубывающая;
2. $\lim_{x \rightarrow -\infty} F_X(x) = 0, \lim_{x \rightarrow \infty} F_X(x) = 1$;
3. F_X непрерывна слева.

Из того факта, что борелевская сигма-алгебра на вещественной прямой порождается семейством интервалов вида $\{(-\infty, x]\}_{x \in \mathbb{R}}$, вытекает

Теорема 2. Любая функция $F(x)$, удовлетворяющая трём свойствам, перечисленным выше, является функцией распределения для какого-то распределения \mathbb{R}^X .

Для вероятностных распределений, обладающих определенными свойствами, существуют более удобные способы его задания.

Дискретные распределения

Определение 3. Случайная величина называется **простой** или **дискретной**, если она принимает не более, чем счётное число значений. То есть

$X(\omega) = a_i, \forall \omega \in A_i$, где $\{A_i\}_{i=1}^{\infty}$ — разбиение Ω .

Распределение простой случайной величины тогда по определению задаётся:

$$\mathbb{P}^X(B) = \sum_{i: a_i \in B} \mathbb{P}(A_i)$$

. Введя обозначение $p_i = \mathbb{P}(A_i)$, можно задать

$$\sum_{i=1}^{\infty} p_i = 1$$

функцию $p(a_i) = p_i$. Очевидно, что . Используя счётную аддитивность

\mathbb{P} , легко показать, что эта функция однозначно определяет распределение X .

$$\sum_{i=1}^{\infty} p_i = 1$$

Определение 4. Функция $p(a_i) = p_i$, где $i=1, 2, \dots$ часто называется **дискретным распределением**.

Пример 1. Пусть функция p задана таким образом, что $p(-1) = \frac{1}{2}$ и $p(1) = \frac{1}{2}$. Эта функция задаёт распределение случайной величины X такой, что $\mathbb{P}(X = \pm 1) = \frac{1}{2}$ ([распределение Бернулли](#)).

Теорема 3. Дискретное распределение обладает следующими свойствами:

1. $p_i \geq 0$;

2. $\sum_{i=1}^n p_i = 1$.

Непрерывные распределения

Непрерывное распределение — распределение вероятностей, не имеющее атомов.
Любое распределение вероятностей есть смесь дискретного и непрерывного.

Абсолютно непрерывные распределения

Определение 5. Распределение случайной величины X называется абсолютно непрерывным, если существует неотрицательная функция $f_X : \mathbb{R} \rightarrow \mathbb{R}_+$, такая

$$\mathbb{P}^X(B) \equiv \mathbb{P}(X \in B) = \int_B f_X(x) dx$$

что . Функция f_X тогда называется плотностью распределения случайной величины X .

Пример 2. Пусть $f(x) = 1$, когда $0 \leq x \leq 1$, и 0 — в противном случае. Тогда

$$\mathbb{P}(a < X < b) = \int_a^b 1 \, dx = b - a, \text{ если } (a, b) \subset [0, 1].$$

Очевидно, что для любой плотности распределения f_X верно равенство

$$\int_{-\infty}^{\infty} f_X(x) \, dx = 1.$$

. Верна и обратная

Теорема 4. Если функция $f : \mathbb{R} \rightarrow \mathbb{R}$ такая, что:

1. $f(x) \geq 0, \forall x \in \mathbb{R};$

2. $\int_{-\infty}^{\infty} f(x) dx = 1,$

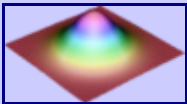
то существует распределение \mathbb{P}^X такое, что $f(x)$ является его плотностью.

Просто применение формулы Ньютона-Лейбница приводит к простому соотношению между кумулятивной функцией и плотностью абсолютно непрерывного распределения.

Теорема 5. Если $f(x)$ — непрерывная плотность распределения, а $F(x)$ — его кумулятивная функция, то

1. $F'(x) = f(x), \forall x \in \mathbb{R},$

2. $F(x) = \int_{-\infty}^x f(t) dt.$



Вероятностные распределения

Одномерные

Многомерные

Дискретные:

[Бернулли](#) | [биномиальное](#) | [геометрическое](#) |
[гипергеометрическое](#) | [логарифмическое](#) |
[отрицательное биномиальное](#) | [Пуассона](#) |
[дискретное равномерное](#)

[мультиномиальное](#)

**Абсолютно
непрерывные:**

[Бета](#) | [Вейбулла](#) | [Гамма](#) |
[гиперэкспоненциальное](#) | [Колмогорова](#) |
[Коши](#) | [Лапласа](#) | [логнормальное](#) |
[нормальное \(Гаусса\)](#) | [логистическое](#) |
[Парето](#) | [полукруговое](#) | [непрерывное](#)
[равномерное](#) | [Райса](#) | [Рэлея](#) | [Стьюдента](#) |
[Фишера](#) | [хи-квадрат](#) | [экспоненциальное](#) |
[variance-gamma](#)

[многомерное](#)
[нормальное](#)

Примечания

При построении распределения по эмпирическим (опытным) данным следует избегать ошибок округление

Функция вероятности

Функция вероятности в [теории вероятностей](#) — наиболее часто используемый способ охарактеризовать [дискретное распределение](#).

Определения

Функция произвольной вероятности

Пусть \mathbb{P} является вероятностной мерой на \mathbb{R}^n , то есть определено вероятностное пространство $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mathbb{P})$, где $\mathcal{B}(\mathbb{R}^n)$ обозначает борелевскую σ -алгебру на \mathbb{R}^n .

Определение 1. Вероятностная мера называется дискретной, если её носитель \mathbb{P} не более, чем счётен, то есть существует не более, чем счётное подмножество $X \subset \mathbb{R}^n$ такое, что $\mathbb{P}(X) = 1$.

Определение 2. Функция $p : \mathbb{R}^n \rightarrow [0, 1]$, определённая следующим образом:

$$p(x) = \begin{cases} \mathbb{P}(\{x\}), & x \in X \\ 0, & x \in \mathbb{R}^n \setminus X \end{cases}$$

называется функцией вероятности \mathbb{P} .

Функция вероятности случайной величины

Определение 3. Пусть $X : \Omega \rightarrow \mathbb{R}^n$ — случайная величина (случайный вектор).

Тогда она индуцирует вероятностную меру \mathbb{P}^X на \mathbb{R}^n , называемую распределением. Случайная величина называется дискретной, если её распределение дискретно. Функция вероятности p_X случайной величины X имеет вид:

$$p_X(x) = \mathbb{P}^X(\{x\}) \equiv \mathbb{P}(X = x).$$

или короче

$$p_X(x_i) = \mathbb{P}(X = x_i) = p_i, \quad i \in \mathbb{N},$$

где $X = \{x_1, x_2, x_3, \dots\} \subset \mathbb{R}^n$.

Свойства функции вероятности

Из свойств [вероятности](#) очевидно следует:

- $p_X(x_i) \geq 0, \forall i \in \mathbb{N}$.
- $\sum_{i=1}^{\infty} p_X(x_i) = 1$.

Функция распределения случайной величины может быть выражена через её функцию вероятности:

$$F_X(x) = \sum_{x' \leq x} p_X(x')$$

- Если $X = (X_1, X_2)$, то

$$\sum_{x_2} p_{X_1, X_2}(x_1, x_2) = p_{X_1}(x_1)$$
$$\sum_{x_1} p_{X_1, X_2}(x_1, x_2) = p_{X_2}(x_2),$$

где p_{X_1, X_2} — функция вероятности вектора (X_1, X_2) , а p_{X_i} — функция вероятности величины X_i , $i = 1, 2$. Это свойство очевидно обобщается для случайных векторов размерности $n > 2$.

- Математическое ожидание функции от дискретной величины, когда оно существует, имеет вид:

$$\mathbb{E}[g(X)] = \sum_{i=1}^n g(x_i) p_i,$$

при условии что ряд в правой части абсолютно сходится.

Функция распределения

Функция распределения в [теории вероятностей](#) — функция, характеризующая [распределение случайной величины](#) или случайного вектора. При соблюдении известных условий (см. ниже) полностью определяет случайную величину.

Определение

Пусть дано вероятностное пространство $(\mathbb{R}, \mathcal{F}, \mathbb{P})$, и на нём определена случайная величина X с распределением \mathbb{P}^X . Тогда функцией распределения случайной величины X называется функция $F_X: \mathbb{R} \rightarrow [0, 1]$, задаваемая формулой:

$$F_X(x) = \mathbb{P}(X \leq x) \equiv \mathbb{P}^X((-\infty, x]).$$

Т.е. функцией распределения (вероятностей) случайной величины X называют функцию $F(x)$, значение которой в точке x равно вероятности события $\{X \leq x\}$, т.е. события, состоящего только из тех элементарных исходов, для которых $X(\omega) \leq x$.

Свойства

- F_X непрерывна слева: [1]

$$\lim_{\varepsilon \rightarrow 0-} F_X(x + \varepsilon) = F_X(x)$$

- F_X не убывает на всей числовой прямой.

- $\lim_{x \rightarrow -\infty} F_X(x) = 0$.

- $\lim_{x \rightarrow +\infty} F_X(x) = 1$.

- Распределение случайной величины \mathbb{P}^X однозначно определяет функцию распределения.
 - Верно и обратное: если функция $F(x)$ удовлетворяет четырём перечисленным выше свойствам, то существует вероятностное пространство и определённая на нём случайная величина, такая что $F(x)$ является её функцией распределения.
-

По определению непрерывности справа, функция F_X имеет правый предел $F_X(x+)$ в любой точке $x \in \mathbb{R}$, и он совпадает со значением функции $F_X(x)$ в этой точке.

- В силу неубывания, функция F_X также имеет и левый предел $F_X(x-)$ в любой точке $x \in \mathbb{R}$, который может не совпадать со значением функции. Таким образом, функция F_X либо непрерывна в точке, либо имеет в ней разрыв первого рода.

Тождества

Из свойств вероятности следует, что $\forall x \in \mathbb{R}, \forall a, b \in \mathbb{R}$, таких что $a < b$:

- $\mathbb{P}(X > x) = 1 - F_X(x)$;
- $\mathbb{P}(X < x) = F_X(x-)$;
- $\mathbb{P}(X \geq x) = 1 - F_X(x-)$;
- $\mathbb{P}(X = x) = F_X(x) - F_X(x-)$;
- $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$;
- $\mathbb{P}(a \leq X \leq b) = F_X(b) - F_X(a-)$;
- $\mathbb{P}(a < X < b) = F_X(b-) - F_X(a)$;
- $\mathbb{P}(a \leq X < b) = F_X(b-) - F_X(a-)$.

Дискретные распределения

Если случайная величина X дискретна, то есть её распределение однозначно задаётся функцией вероятности

$$\mathbb{P}(X = x_i) = p_i, \quad i = 1, 2, \dots,$$

то функция распределения F_X этой случайной величины кусочно-постоянна и может быть записана как:

$$F_X(x) = \sum_{i: x_i \leq x} p_i.$$

Эта функция непрерывна во всех точках $x \in \mathbb{R}$, таких что $x \neq x_i, \forall i$, и имеет разрыв первого рода в точках $x = x_i, \forall i$.

Непрерывные распределения

Распределение \mathbb{P}^X называется непрерывным, если такова его функция распределения F_X . В этом случае:

$$\mathbb{P}(X = x) = 0, \quad \forall x \in \mathbb{R},$$

и

$$F_X(x - 0) = F_X(x), \quad \forall x \in \mathbb{R},$$

а следовательно формулы имеют вид:

$$\mathbb{P}(X \in |a, b|) = F_X(b) - F_X(a),$$

где $|a, b|$ означает любой интервал, открытый или закрытый, конечный или бесконечный.

Абсолютно непрерывные распределения

Распределение \mathbb{P}^X называется абсолютно непрерывным, если существует неотрицательная почти всюду (относительно меры Лебега) функция $f_X(x)$, такая что:

$$F_X(x) = \int_{-\infty}^x f_X(t) dt .$$

Функция f_X называется плотностью распределения. Известно, что функция абсолютно непрерывного распределения непрерывна, и, более того, если $f_X \in C(\mathbb{R})$, то $F_X \in \mathcal{D}(\mathbb{R})$, и

$$\frac{d}{dx} F_X(x) = f_X(x), \quad \forall x \in \mathbb{R} .$$

Вариации и обобщения

Иногда в российской литературе берётся такое определение функции распределения:

$$F_X(x) = \mathbb{P}(X < x) \equiv \mathbb{P}^X((-\infty, x)).$$

Определённая так функция распределения будет непрерывна слева, а не справа.

Многомерные функции распределения

Пусть $(\Omega, \mathcal{F}, \mathbb{P})$ фиксированное вероятностное пространство, и $X = (X_1, \dots, X_n): \Omega \rightarrow \mathbb{R}^n$ — случайный вектор. Тогда распределение \mathbb{P}^X , называемое *распределением случайного вектора X* или *совместным распределением случайных величин X_1, \dots, X_n* , является вероятностной мерой на \mathbb{R}^n . Функция этого распределения $F_X: \mathbb{R}^n \rightarrow [0, 1]$ задаётся по определению следующим образом:

$$F_X(x_1, \dots, x_n) = \mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n) \equiv \mathbb{P}^X \left(\prod_{i=1}^n (-\infty, x_i] \right),$$

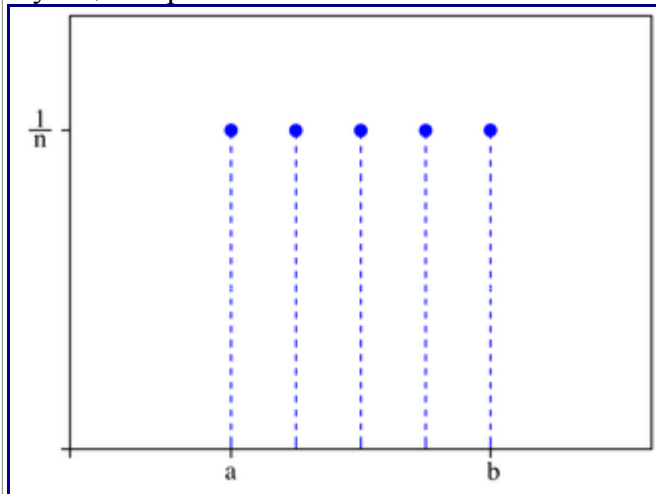
где \prod в данном случае обозначает [декартово произведение множеств](#).

Свойства многомерных функций распределения аналогичны одномерному случаю. Также сохраняется взаимно-однозначное соответствие между распределениями на \mathbb{R}^n и многомерными функциями распределения. Однако, формулы для вычисления вероятностей существенно усложняются, и потому функции распределения редко используются для $n > 1$.

Дискретное равномерное распределение

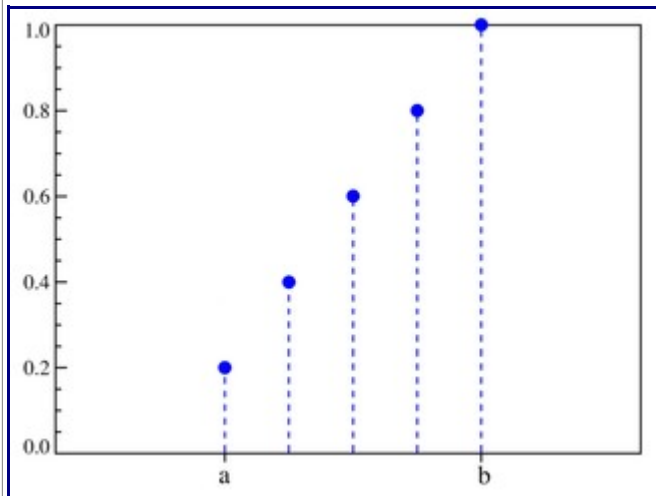
Дискретное равномерное распределение

Функция вероятности



$n=5$, где $n=b-a+1$

Функция распределения



$n=5$, где $n=b-a+1$.

Обозначение	
Параметры	$a \in (\dots, -2, -1, 0, 1, 2, \dots)$ $b \in (\dots, -2, -1, 0, 1, 2, \dots)$ $n = b - a + 1$
<u>Носитель</u>	$k \in \{a, a + 1, \dots, b - 1, b\}$
<u>Функция вероятности</u>	$\frac{1}{n}, \quad a \leq k \leq b$ $0, \quad \text{else}$
<u>Функция распределения</u>	$0, \quad k < a$ $\frac{k-a+1}{n}, \quad a \leq k \leq b$ $1, \quad k > b$
<u>Математическое ожидание</u>	$\frac{a + b}{2}$
<u>Медиана</u>	$\frac{a + b}{2}$

<u>Мода</u>	нет
<u>Дисперсия</u>	$\frac{n^2 - 1}{12}$
<u>Коэффициент асимметрии</u>	0
<u>Коэффициент эксцесса</u>	$-\frac{6(n^2 + 1)}{5(n^2 - 1)}$
<u>Информационная энтропия</u>	$\ln n$
<u>Производящая функция моментов</u>	$\frac{e^{at} - e^{(b+1)t}}{n(1 - e^t)}$
<u>Характеристическая функция</u>	$\frac{e^{iat} - e^{i(b+1)t}}{n(1 - e^{it})}$

В теории вероятностей случайная величина имеет **дискретное равномерное распределение**, если она принимает конечное число значений с равными вероятностями.

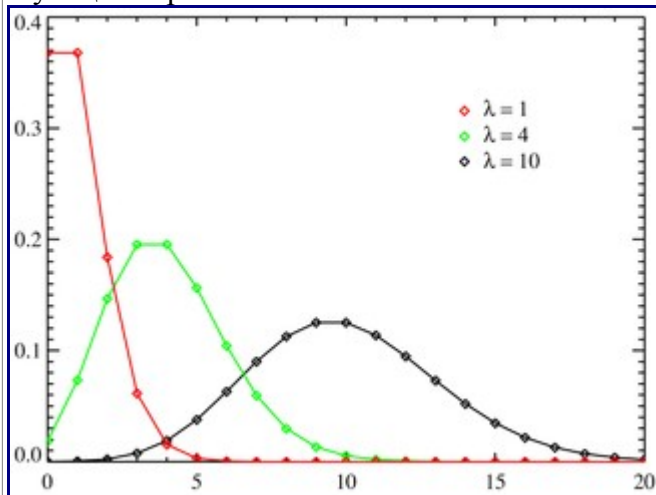
Примеры

- Случайная величина, принимающая значение 1, если выпал орёл, и 0, если выпала решка, имеет дискретное равномерное распределение. Она принимает оба значения с вероятностью $1/2$.
- Случайная величина, равная выпавшему числу на игральной кости, имеет дискретное равномерное распределение на $\{1,2,3,4,5,6\}$, и она принимает каждое значение с вероятностью $1/6$.

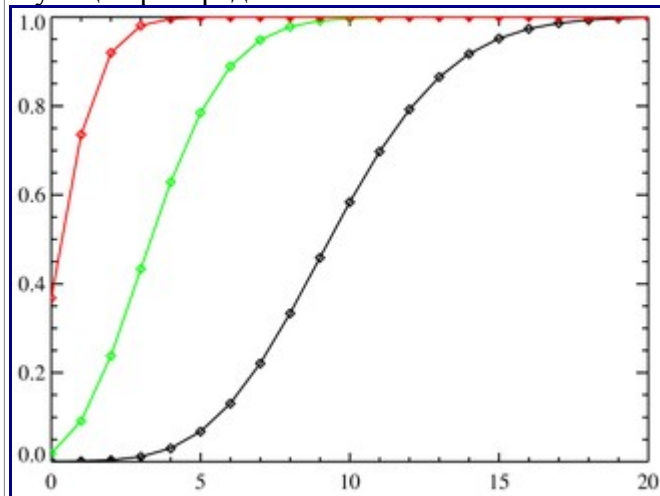
Распределение Пуассона

Распределение Пуассона

Функция вероятности



Функция распределения



Обозначение	$P(\lambda)$
Параметры	$\lambda \in (0, \infty)$
Носитель	$k \in \{0, 1, 2, \dots\}$

<u>Функция вероятности</u>	$\frac{e^{-\lambda} \lambda^k}{k!}$
<u>Функция распределения</u>	$\frac{\Gamma(k + 1, \lambda)}{k!}$
<u>Математическое ожидание</u>	λ
<u>Медиана</u>	$\approx \lfloor \lambda + 1/3 - 0.02/\lambda \rfloor$
<u>Мода</u>	$\lfloor \lambda \rfloor$
<u>Дисперсия</u>	λ
<u>Коэффициент асимметрии</u>	$\lambda^{-1/2}$
<u>Коэффициент эксцесса</u>	λ^{-1}
<u>Информационная энтропия</u>	$\lambda[1 - \ln(\lambda)] + e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k \ln(k!)}{k!}$

Производящая функция моментов

$$\exp(\lambda(e^t - 1))$$

Характеристическая функция

$$\exp(\lambda(e^{it} - 1))$$

Распределение Пуассона моделирует случайную величину, представляющую собой число событий, произошедших за фиксированное время, при условии, что данные события происходят с некоторой фиксированной средней интенсивностью и независимо друг от друга.

Распределение Пуассона играет ключевую роль в теории массового обслуживания.

Определение

Выберем фиксированное число $\lambda > 0$ и определим дискретное распределение, задаваемое следующей функцией вероятности:

$$p(k) \equiv \mathbb{P}(Y = k) = \frac{\lambda^k}{k!} e^{-\lambda},$$

где

$k!$ обозначает факториал,

$e = 2.718281828\dots$ — основание натурального логарифма.

Тот факт, что случайная величина Y имеет распределение Пуассона с параметром λ , записывается: $Y \sim P(\lambda)$.

Моменты

Производящая функция моментов распределения Пуассона имеет вид:

$$E_Y(t) = e^{\lambda(e^t - 1)},$$

откуда

$$\begin{aligned} \mathbb{M}[Y] &= \lambda, \\ \mathbb{D}[Y] &= \lambda. \end{aligned}$$

Для факториальных моментов распределения справедлива общая формула:

$$\mathbb{M}Y^{[k]} = \lambda^k,$$

где $k = 1, 2, \dots$

А так как моменты и факториальные моменты линейным образом связаны, то часто для Пуассоновского распределения исследуются именно факториальные моменты, из которых при необходимости можно вывести и обычные моменты.

Свойства распределения Пуассона

Сумма независимых пуассоновских случайных величин также имеет распределение Пуассона. Пусть $Y_i \sim P(\lambda_i)$, $i = 1, \dots, n$. Тогда

$$Y = \sum_{i=1}^n Y_i \sim P\left(\sum_{i=1}^n \lambda_i\right).$$

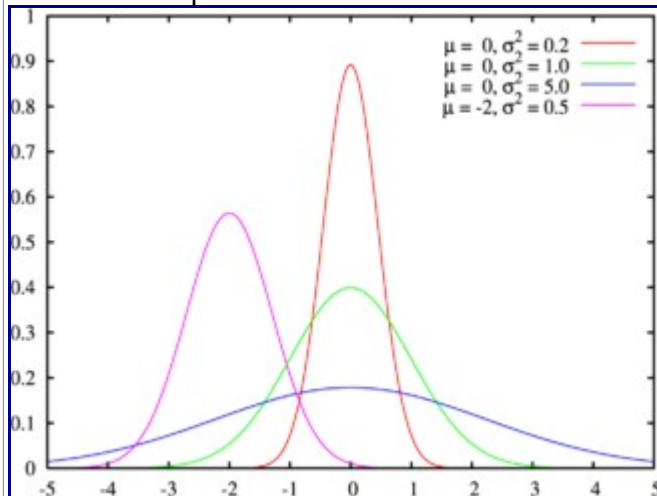
Пусть $Y_i \sim P(\lambda_i)$, $i = 1, 2$, и $Y = Y_1 + Y_2$. Тогда условное распределение Y_1 при условии, что $Y = y$, биномиально. Более точно:

$$Y_1 \mid Y = y \sim \text{Bin}\left(y, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right).$$

Нормальное распределение

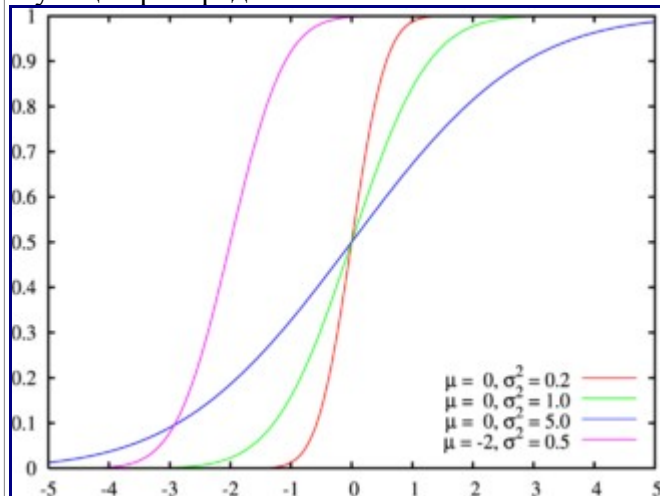
Нормальное распределение

Плотность вероятности



Зеленая линия соответствует стандартному нормальному распределению

Функция распределения



Цвета на этом графике соответствуют графику сверху

Обозначение	$N(\mu, \sigma^2)$
Параметры	μ - <u>коэффициент сдвига</u> (вещественное число) $\sigma > 0$ - <u>коэффициент масштаба</u> (вещественный, строго положительный)
<u>Носитель</u>	$x \in (-\infty; +\infty)$
<u>Плотность вероятности</u>	$\frac{1}{\sigma \sqrt{2\pi}} \exp \left(-\frac{(x - \mu)^2}{2\sigma^2} \right)$
<u>Функция распределения</u>	$\frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{x - \mu}{\sqrt{2\sigma^2}} \right) \right]$
<u>Математическое ожидание</u>	μ
<u>Медиана</u>	μ
<u>Мода</u>	μ

Дисперсия

$$\sigma^2$$

Коэффициент
асимметрии

$$0$$

Коэффициент эксцесса

$$0$$

Информационная
энтропия

$$\ln(\sigma\sqrt{2\pi e})$$

Производящая функция
моментов

$$M_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$$

Характеристическая
функция

$$\phi_X(t) = \exp\left(\mu i t - \frac{\sigma^2 t^2}{2}\right)$$

Нормальное распределение, также называемое **гауссовским распределением** или **распределением Гаусса** — распределение вероятностей, которое задается функцией плотности распределения:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

где параметр μ — среднее значение (математическое ожидание) случайной величины и указывает координату максимума кривой плотности распределения, а σ^2 — дисперсия.

Нормальное распределение играет важнейшую роль во многих областях знаний, особенно в статистической физике. Физическая величина, подверженная влиянию значительного числа независимых факторов, могущих вносить с равной погрешностью положительные и отрицательные отклонения, вне зависимости от природы этих случайных факторов, часто подчиняется нормальному распределению, поэтому из всех распределений в природе чаще всего встречается нормальное (отсюда и произошло одно из названий этого распределения вероятностей).

Нормальное распределение зависит от двух параметров — *смещения* и *масштаба*, то есть является с математической точки зрения не одним распределением, а целым их семейством. Значения параметров соответствуют значениям среднего (математического ожидания) и разброса (стандартного отклонения).

Стандартным нормальным распределением называется нормальное распределение с математическим ожиданием 0 и стандартным отклонением 1

Свойства

Если случайные величины X_1 и X_2 независимы и имеют нормальное распределение с математическими ожиданиями μ_1 и μ_2 и дисперсиями σ_1^2 и σ_2^2 соответственно, то $X_1 + X_2$ также имеет нормальное распределение с математическим ожиданием $\mu_1 + \mu_2$ и дисперсией $\sigma_1^2 + \sigma_2^2$.

Моделирование нормальных случайных величин

Простейшие, но неточные методы моделирования основываются на центральной предельной теореме. Именно, если сложить много независимых одинаково распределённых величин с конечной дисперсией, то сумма будет распределена *примерно* нормально. Например, если сложить 12 независимых базовых случайных величин, получится грубое приближение стандартного нормального распределения. Тем не менее, с увеличением слагаемых распределение суммы стремится к нормальному.

Использование точных методов предпочтительно, поскольку у них практически нет недостатков. В частности, преобразование Бокса — Мюллера является точным, быстрым и простым для реализации методом генерации.

Центральная предельная теорема

Нормальное распределение часто встречается в природе. Например, следующие случайные величины хорошо моделируются нормальным распределением:

- отклонение при стрельбе
- погрешности измерений
- рост живых организмов

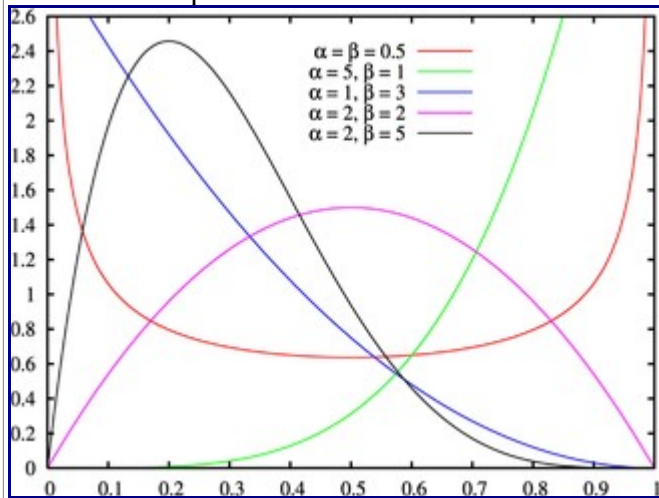
Такое широкое распространение закона связано с тем, что он является предельным законом, к которому приближаются многие другие (например, биномиальный).

Доказано, что сумма очень большого числа случайных величин, влияние каждой из которых близко к 0, имеет распределение, близкое к нормальному. Этот факт является содержанием [центральной предельной теоремы](#).

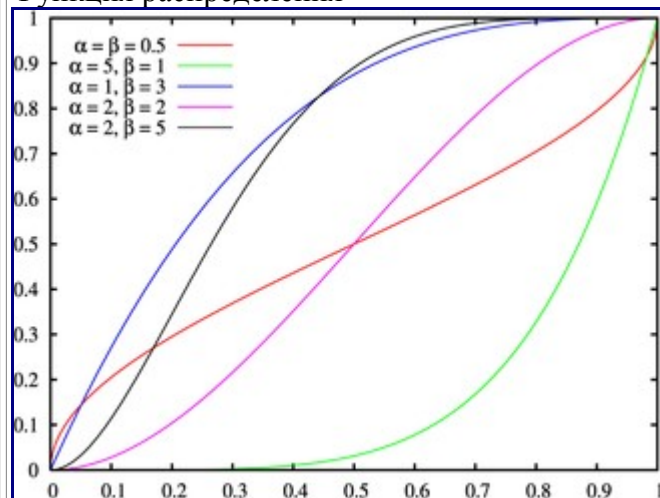
Бета-распределение

Бета-распределение

Плотность вероятности



Функция распределения



Обозначение	$\{\{\{\text{notation}\}\}\}$
Параметры	$\alpha > 0$ $\beta > 0$
<u>Носитель</u>	$x \in [0, 1]$
<u>Плотность вероятности</u>	$\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$
<u>Функция распределения</u>	$I_x(\alpha, \beta)$
<u>Математическое ожидание</u>	$\frac{\alpha}{\alpha + \beta}$
<u>Медиана</u>	
<u>Мода</u>	$\frac{\alpha - 1}{\alpha + \beta - 2} \text{ для } \alpha > 1, \beta > 1$

<u>Дисперсия</u>	$\frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$
<u>Коэффициент асимметрии</u>	$\frac{2(\beta - \alpha)\sqrt{\alpha + \beta + 1}}{(\alpha + \beta + 2)\sqrt{\alpha\beta}}$
<u>Коэффициент эксцесса</u>	$6 \frac{\alpha^3 - \alpha^2(2\beta - 1) + \beta^2(\beta + 1) - 2\alpha\beta(\beta + 2)}{\alpha\beta(\alpha + \beta + 2)(\alpha + \beta + 3)}$
<u>Информационная энтропия</u>	
<u>Производящая функция моментов</u>	$1 + \sum_{k=1}^{\infty} \left(\prod_{r=0}^k \frac{\alpha + r}{\alpha + \beta + r} \right) \frac{t^k}{k!}$
<u>Характеристическая функция</u>	${}_1F_1(\alpha; \alpha + \beta; i t)$

Бéта распределéние в [теории вероятностей](#) и [статистике](#) — двухпараметрическое семейство [абсолютно непрерывных распределений](#). Используется для описания случайных величин, значения которых ограничены конечным интервалом.

Определение

Пусть распределение случайной величины X задаётся плотностью вероятности f_X , имеющей вид:

$$f_X(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1},$$

где

- $\alpha, \beta > 0$ произвольные фиксированные параметры, и

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx$$

- — бета-функция.

Тогда случайная величина X имеет бета-распределение. Пишут: $X \sim B(\alpha, \beta)$.

Форма графика

Форма графика плотности вероятности бета-распределения зависит от выбора параметров α и β .

- $\alpha < 1, \beta < 1$ — график выпуклый и уходит в бесконечность на границах (красная кривая);
- $\alpha < 1, \beta \geq 1$ или $\alpha = 1, \beta > 1$ — график строго убывающий (синяя кривая)
 - $\alpha = 1, \beta > 2$ — график строго выпуклый;
 - $\alpha = 1, \beta = 2$ — график является прямой линией;
 - $\alpha = 1, 1 < \beta < 2$ — график строго вогнутый;
- $\alpha = 1, \beta = 1$ график совпадает с графиком плотности стандартного непрерывного равномерного распределения;
- $\alpha = 1, \beta < 1$ или $\alpha > 1, \beta \leq 1$ — график строго возрастающий (зелёная кривая);
 - $\alpha > 2, \beta = 1$ — график строго выпуклый;
 - $\alpha = 2, \beta = 1$ — график является прямой линией;
 - $1 < \alpha < 2, \beta = 1$ — график строго вогнутый;

- $\alpha > 1, \beta > 1$ — график унимодальный (пурпурная и чёрная кривые)

В случае, когда $\alpha = \beta$, плотность вероятности симметрична относительно $1/2$ (красная и пурпурная кривые), то есть

$$f_X(x - 1/2) = f_X(x + 1/2), \quad x \in [0, 1/2].$$

Моменты

Математическое ожидание и дисперсия случайной величины X , имеющей бета-распределение, имеют вид:

$$\mathbb{E}[X] = \frac{\alpha}{\alpha + \beta},$$

$$D[X] = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

Связь с другими распределениями

- Стандартное непрерывное равномерное распределение является частным случаем бета-распределения:

$$U[0, 1] \equiv B(1, 1).$$

- Бета-распределение широко используется в байесовской статистике, так как оно является сопряжённым априорным распределением для распределения Бернулли, биномиального и геометрического распределений.
- Если X, Y — независимые гамма распределённые случайные величины, причём $X \sim \Gamma(\alpha, 1)$, а $Y \sim \Gamma(\beta, 1)$, то

$$\frac{X}{X + Y} \sim B(\alpha, \beta)$$

Распределение Дирихле

Распределение Дирихле является обобщением [Бета-распределения](#) на многомерный случай. То есть, его функция плотности вероятности возвращает доверительную вероятность того, что вероятность каждого из K взаимоисключающих событий равна x_i при условии, что каждое событие наблюдалось $\alpha_i - 1$ раз.

Функция плотности вероятности

Функция плотности вероятности для распределения Дирихле порядка K есть:

$$f(x_1, \dots, x_K; \alpha_1, \dots, \alpha_K) = \frac{1}{B(\alpha)} \prod_{i=1}^K x_i^{\alpha_i - 1}$$

где $x_i \geq 0$, $\sum_{i=1}^K x_i = 1$, и $\alpha_i \geq 0$.

Свойства

Пусть $X = (X_1, \dots, X_K) \sim \text{Dir}(\alpha)$ и $\alpha_0 = \sum_{i=1}^K \alpha_i$, тогда

$$\mathbb{E}[X_i|\alpha] = \frac{\alpha_i}{\alpha_0},$$

$$\text{Var}[X_i|\alpha] = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0 + 1)},$$

$$\text{Cov}[X_i X_j|\alpha] = \frac{-\alpha_i \alpha_j}{\alpha_0^2(\alpha_0 + 1)}.$$

Модой распределения является вектор $x = (x_1, \dots, x_K)$ с

$$x_i = \frac{\alpha_i - 1}{\alpha_0 - K}, \quad \alpha_i > 1.$$

Распределение Дирихле является сопряжённым априорным распределением к мультиномиальному распределению, а именно: если

$$\beta|X = (\beta_1, \dots, \beta_K)|X \sim \text{Mult}(X),$$

где β_i — число вхождений i в выборку из n точек дискретного распределения на $\{1, \dots, K\}$ определенного через X , то

$$X|\beta \sim \text{Dir}(\alpha + \beta).$$

Эта связь используется в Байесовской статистике для того, чтобы оценить скрытые параметры, X , дискретного вероятностного распределения имея набор из n выборок. Очевидно, если априорное распределение обозначено как $\text{Dir}(\alpha)$, то $\text{Dir}(\alpha + \beta)$ есть апостериорное распределение после серии наблюдений с гистограммой β .

Связи с другими распределениями

Если для $i \in \{1, 2, \dots, K\}$,

$Y_i \sim \text{Gamma}(\text{shape} = \alpha_i, \text{scale} = 1)$ независимо, то

$$V = \sum_{i=1}^K Y_i \sim \text{Gamma}(\text{shape} = \sum_{i=1}^K \alpha_i, \text{scale} = 1),$$

и

$$(X_1, \dots, X_K) = (Y_1/V, \dots, Y_K/V) \sim \text{Dir}(\alpha_1, \dots, \alpha_K).$$

Несмотря на то, что X_i не являются независимыми друг от друга, они могут быть сгенерированы из набора из K независимых гамма случайных величин. К несчастью, так как сумма V теряется в процессе формирования $X = (X_1, \dots, X_K)$, становится невозможно восстановить начальные значения гамма случайных величин только по этим значениям. Тем не менее, благодаря тому, что работать с независимыми случайными величинами проще, это преобразование параметров может быть полезно при доказательстве свойств распределения Дирихле.

Генерация случайных чисел

Метод построения случайного вектора $x = (x_1, \dots, x_K)$ для распределения Дирихле размерности K с параметрами $(\alpha_1, \dots, \alpha_K)$ следует непосредственно из этой связи. Сначала получим K независимых случайных выборок y_1, \dots, y_K из [гамма-распределений](#), каждое из которых имеет плотность

$$\frac{y_i^{\alpha_i-1} e^{-y_i}}{\Gamma(\alpha_i)},$$

а затем положим

$$x_i = y_i / \sum_{j=1}^K y_j.$$

Наглядная трактовка параметров

В качестве примера использования распределения Дирихле можно предложить задачу, в которой требуется разрезать нитки (каждая начальной длины 1.0) на K частей с разными длинами так, чтобы все части имели заданную среднюю длину, но с возможностью некоторой вариации относительных длин частей. Значения α/α_0 определяют средние длины частей нитки, получившиеся из распределения. Дисперсия вокруг среднего значения обратно пропорциональна α_0 .