Функции распределения

Определение 4.1. *Случайной величиной* $\xi = \xi(\omega)$ называется измеримое отображение измеримого пространства $(\Omega, \mathfrak{A})$ на *борелевскую прямую* $(\mathbb{R}, \mathfrak{B})$, т.е.

$$\xi^{-1}(B) \;=\; \{\omega :\; \xi(\omega) \in B\} \;\in\; \mathfrak{A}, \;\; \forall B \in \mathfrak{B},$$

или

$$\{\omega :\; \xi(\omega) < x\} \;\in\; \mathfrak{B}, \;\; \forall x \in \mathbb{R}.$$

Относительно ξ нас интересуют только вероятности $\mathbf{P}\{\xi \in B\}$ для разных подмножеств B числовой прямой.

Определение. *Распределением* случайной величины ξ называется вероятностная мера на борелевской сигма-алгебре

$$\mathsf{P}_\xi(B) \;=\; \mathbf{P}\{\xi \in B\}, \;\; B \in \mathfrak{B}.$$

Функцией распределения случайной величины ξ называется

$$F(x) = F_\xi(x) \;=\; \mathbf{P}\{\xi < x\}, \quad x \in \mathbb{R}.$$

Теорема (Осовные свойства функции распределения). *Функция* $F(x)$, $x \in$ R, *обладает следующими свойствами.*

$(F1)$ $F(x)$ *– неубывающая функция* $x \in \mathbb{R}$.

$(F2)$ $\lim\limits_{x\to-\infty} F(x) = 0, \ \lim\limits_{x\to+\infty} F(x) = 1.$

$(F3)$ *Функция* $F(x)$ *непрерывна слева:* $\lim\limits_{x\to a-} F(x) = F(a)$.

Полезные свойства функции распределения.

Теорема.

$(F4)$ *Вероятности попадания сл.в.* ξ *в интервалы* $\mathbb{R}$*:*

$$\mathbf{P}\{\xi \in [\,a,\ b)\} = F(b) - F(a), \qquad \mathbf{P}\{\xi \in [\,a,\ b\,]\} = F(b+) - F(a),$$
$$\mathbf{P}\{\xi \in (a,\ b\,]\} = F(b+) - F(a+), \qquad \mathbf{P}\{\xi \in (a,\ b)\} = F(b) - F(a+),$$
$$\mathbf{P}\{\xi = a\} \ = \ F(a+) - F(a).$$

$(F5)$ *Разрыв функции распределения только в точках, вероятность значения которых больше нуля:*

$$\mathbf{P}\{\xi = a\} \ = \ F(a+) - F(a-) > 0.$$

$(F6)$ *Функция* $F(x)$ *имеет не более чем счетное множество скачков.*

Доказательство

$(F1)$. Если $x_1 < x_2$, то

$$B_1 = (-\infty,\, x_1) \subseteq B_2 = (-\infty,\, x_2)$$
$$F(x_1) = \mathbf{P}\{\xi \in B_1\} \leqslant \mathbf{P}\{\xi \in B_2\} = F(x_2).$$

$(F2)$. Пусть $B_n = (-\infty; n), n = 1, 2 \ldots$.

Так как $B_1 \subset B_2 \subset \ldots \subset B_n \nearrow \mathbb{R}$, то в силу непрерывности вероятности при $n \to \infty$

$$F(n) = \mathbf{P}\{\xi \in B_n\} \nearrow \mathbf{P}\{\xi \in \mathbb{R}\} = 1.$$

В силу $(F1)$ отсюда следует, что $\lim\limits_{x\to\infty} F(x) = 1$. Аналогично для второго свойства $(F2)$.

$(F3)$. Пусть $x_n = x - 1/n \nearrow x$ при $n \to \infty$, следовательно

$$B_n = (-\infty,\, x_n) \nearrow B = (-\infty,\, x).$$

Снова в силу непрерывности меры

$$F(x_n) = \mathbf{P}\{\xi \in B_n\} \to \mathbf{P}\{\xi \in B\} = F(x)$$

что, по определению, означает непрерывность слева функции $F(x)$.

$(F4)$–$(F5)$ Выводятся аналогично — на самостоятельную работу.

$(F6)$.

Рассмотрим некоторое целое $n \geqslant 1$ и множество $B_n = \{a : \ F(a+) - F(a-) \geqslant \frac{1}{n}\}$ — множество точек разрыва функции $F(x)$ с величиной скачка, не меньшей $1/n$.

В силу $(F4)$ имеем

$$\mathbf{P}\{\xi \in B_n\} \geqslant \sum_{a\in B_n} \mathbf{P}\{\xi = a\} \geqslant \sum_{a\in B_n} \frac{1}{n}.$$

Последняя сумма будет больше 1, если количество элементов B_n больше n. Следовательно, количество элементов B_n не превосходит n.

Очевидно, все скачки F входят в объединение $\bigcup_n B_n$. Как известно, счётное объединение конечных множеств не более чем счётно.

Если мы определим только вероятности элементов борелевского поля $\mathfrak{B}$, имеющих вид интервалов, то сможем ли на основании их вычислять вероятности других событий из $\mathfrak{B}$? Ответ на этот вопрос даёт

Теорема (о мере Лебега–Стилтьеса). *Пусть функция* $F(x),\ x \in \mathbb{R}$, *обладает свойствами* $(F1), (F2), (F3)$. *Тогда на борелевской прямой* $(\mathbb{R}, \mathfrak{B})$ *существует единственная вероятность* P, *для которой* $\mathsf{P}\{(-\infty, x)\} = F(x)$ *для всех* $x \in \mathbb{R}$.

Замечание. Формально говоря, утверждение теоремы означает, что для функции F, удовлетворяющей указанным свойствам, найдётся вероятностное пространство $(\Omega, \mathfrak{A}, \mathbf{P})$ и случайная величина ξ на нём, для которой функция F есть функция распределения этой сл.в. (в данном случае, $\Omega = \mathbb{R}, \mathfrak{A} = \mathfrak{B}, \mathbf{P} = \mathsf{P}$, а сл.в. $\xi(\omega) = \omega, \quad \forall \omega \in \mathbb{R}$)

Доказательство опирается на следующую теорему о продолжении меры.

Теорема (Каратеодори). *Пусть на множестве* Ω *выделен класс подмножеств* $\mathcal{K}$, *образующий полукольцо. Если на полукольце* $\mathcal{K}$ *задана сигма-конечная, сигма-аддитивная функция множеств* μ, *то существует единственная сигма-аддитивная мера* μ^* *на сигма-алгебре* $\sigma(\mathcal{K})$, *порождённой полукольцом* $\mathcal{K}$, *которая совпадает с* μ *на* $\mathcal{K}:\ \mu^*(B) = \mu(B), \forall B \in \mathcal{K}$.

Рассмотрим полукольцо $\mathcal{K}$ всех конечных интервалов вида $[a;b)$ и определим функцию множеств $\mu[a;b) = F(b) - F(a)$.

Так как $\mathbb{R} = \bigcup\limits_{n=-\infty}^{\infty} [n; n+1)$ и $F(n+1) - F(n) < \infty$, то эта функция множеств сигма-конечна.

Заметим, что эта функция аддитивна, т.к. если

$$[a;b) = [c;d) + [e;f), \quad \text{то} \quad a = c \;<\; d = e \;<\; b = f$$

и

$$\mu[a;b) = F(b) - F(a) = \big(F(d) - F(a)\big) + \big(F(b) - F(d)\big) = \mu[c;d) + \mu[e;f)$$

Докажем её сигма-аддитивность, т.е. для любых $c < d$

$$F(d) - F(c) = \sum_{k=1}^{\infty} [\,F(b_k) - F(a_k)\,], \tag{1}$$

если интервал

$$[\,c;d) = \sum_{1}^{\infty} [\,a_k;\, b_k).$$

Для любого конечного $n \geqslant 1$

$$F(d) - F(c) \geqslant \sum_{1}^{n} [\,F(b_j) - F(a_j)\,],$$

ибо по свойству полукольца $\mathcal{K}$ разность

$$[\,c,\, d) \setminus \sum_{1}^{n} [\,a_j;\, b_j)$$

можно представить в виде конечного объединения (M) не пересекающихся интервалов того же вида (возможно отличных от интервалов $[\,a_k;\, b_k)$), т.е.

$$[\,c,\, d) = \sum_{1}^{n} [\,a_k;\, b_k) + \sum_{1}^{M} [\,u_k;\, t_k).$$

В силу аддитивности μ

$$F(d) - F(c) = \mu[\,c,\, d) = \sum_{1}^{n} \mu[\,a_j;\, b_j) + \sum_{1}^{M} \mu[\,u_j;\, t_j) \geqslant \sum_{1}^{n} [\,F(b_j) - F(a_j)\,].$$

Устремляя $n \to \infty$, получаем

$$F(d) - F(c) \geqslant \sum_1^{\infty} [\, F(b_j) - F(a_j)\,].$$

Покажем теперь, что имеет место противоположное неравенство, и, следовательно, справедливо свойство сигма-аддитивности.

Исходный интервал $[\, c;\, d)$ сузим до замкнутого интервала $[\, c;\, d'\,]$ так, чтобы $d' < d$ и $F(d') \geqslant F(d) - \varepsilon$. Этого всегда можно добиться в силу непрерывности слева функции F.

Аналогично, каждый из интервалов $[\, a_n; b_n)$ расширим до открытого интервала $(a'_n;\, b_n)$ так, чтобы $a'_n < a_n$ и

$$F(a'_n) \geqslant F(a_n) - \varepsilon/2^n.$$

В результате получим покрытие

$$[\, c;\, d'\,] \subset \bigcup_{n=1}^{\infty} (a'_n;\, b_n)$$

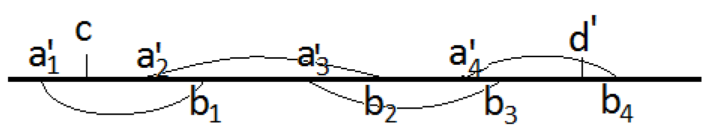
ограниченного замкнутого множества семейством открытых интервалов.

В силу известной леммы Гейне-Бореля найдётся конечное покрытие

$$[\, c; d'\,) \subset [\, c; d'\,] \subset \bigcup_{i=1}^{N} (a'_i;\, b_i) \subset \bigcup_{i=1}^{N} [a'_i;\, b_i),$$

в котором можно выбрать

$$a'_1 < c < a'_2 < b_1 < a'_3 < \ldots < b_N$$



Точки $b_1, \ldots, b_N$ образуют разбиение интервала $[\, a'_1,\, b_N)$, который содержит интервал $[\, c,\, d')$, и поэтому

$$F(d') - F(c) \leqslant F(b_N) - F(a'_1) =$$
$$= \big[F(b_1) - F(a'_1)\big] + \big[F(b_2) - F(b_1)\big] + \ldots + \big[F(b_N) - F(b_{N-1})\big] \leqslant$$
$$\leqslant \sum_{j=1}^{N} \big[F(b_j) - F(a'_j)\big] \leqslant \sum_{j=1}^{\infty} \big[F(b_j) - F(a'_j)\big]$$

Из построения интервалов следует, что

$$F(d) - F(c) \leqslant F(d') - F(c) + \varepsilon$$

и

$$F(b_n) - F(a'_n) \leqslant F(b_n) - F(a_n) + \varepsilon/2^n,$$

откуда

$$F(d) - F(c) \leqslant \sum_{n=1}^{\infty} [F(b_n) - F(a_n)] + 2\varepsilon. \tag{2}$$

Устремляя $\varepsilon \to 0$, получаем окончательное доказательство равенства (1).

Осталось показать, что полученная мера вероятностная, т.е. $\mathsf{P}\{\mathbb{R}\} = 1$. Проделайте это самостоятельно.

Замечание. Можно отказаться от условия $(F2)$ на функцию распределения. В этом случае мы получим способ задания произвольной меры с помощью генерирующей функции. Такие меры называются мерами Лебега–Стилтьеса. Заметим, что любая мера, которая принимает конечные значения на конечных интервалах будет мерой Лебега–Стилтьеса.

Замечание. Так как сигма-алгебра, порождённая интервалами, совпадает с борелевской сигма-алгеброй, то все борелевские множества измеримы относительно любой меры Лебега–Стилтьеса. Однако класс измеримых относительно конкретной меры Лебега–Стилтьеса может быть шире борелевской сигма-алгебры. Например, существуют измеримые по Лебегу не борелевские множества.

Обозначение:

$$\xi \sim F$$

означает, что сл.в. ξ имеет функцию распределения F (имеет распределение, описываемое с помощью ф.распределения F).

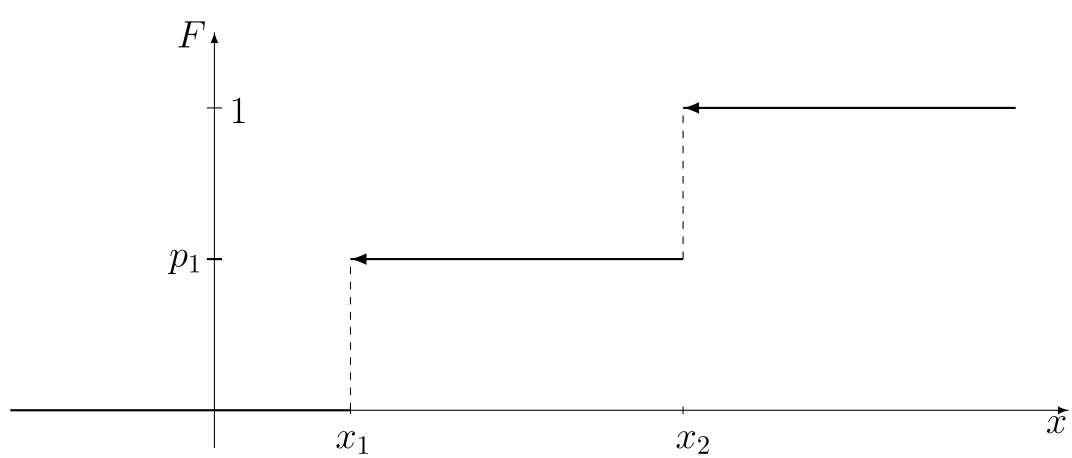
Дискретный тип распределения.

Определение. Говорят, что сл.в. имеет дискретный тип распределения, если найдётся конечное или счётное число точек $\mathfrak{X} = \{x_k, k = \overline{1, N}\}, N \leqslant \infty$, что

$$\sum_{1}^{N} \mathbf{P}\{\xi = x_k\} = 1.$$

Функция $p(x) = \mathbf{P}\{\xi = x\}$, которая отлична от нуля только для $x \in \mathfrak{X}$, называется *функцией вероятностей* дискретной сл.в., а множество её возможных значений $\mathfrak{X}$ называется *носителем сл.в.*

График дискретной ф.р. выглядит наподобие ступенек



Поэтому часто дискретный тип распределения определяют как распределение со ступенчатой ф.распределения.

Дискретные сл.в. гораздо удобнее описывать с помощью функции вероятностей, а не с помощью функции распределения.

Примеры дискретных распределений: Классическое равномерное, Биномиальное, Геометрическое, Гипергеометрическое, Пуассона.

Примеры не дискретных распределений.

Равномерное распределение в отрезке.

Датчики случайных чисел выдают, вообще говоря, дискретную сл.в., у которой носитель $\mathfrak{X} = \{k \cdot 10^{-N}, \quad k = 1, \ldots, 10^N\}$ и все значения этого носителя равновероятны $1/10^N$.

Найдём для этой сл.в. вероятность попадания в интервал $[a; b)$:

$$\mathbf{P}\{\xi \in [a; b)\} = \#\Big\{k:\ a \leqslant \frac{k}{10^N} < b\Big\}\frac{1}{10^N} = \\ = \#\big\{k:\ a10^N \leqslant k < b10^N\big\}\frac{1}{10^N}.$$

Ясно, что количество $(\#)$ целых чисел, попадающих в какой-то интервал, равно разности целых частей границ этого интервала:

$$\mathbf{P}\{\xi \in [a; b)\} \approx \frac{\big[b10^N\big] - \big[a10^N\big]}{10^N}.$$

Так как целая часть любого числа x удовлетворяет неравенствам

$$x - 1 < \big[x\big] < x + 1,$$

то при $N \to \infty$

$$b - a \leftarrow \frac{b10^N - a10^N - 2}{10^N} < \mathbf{P}\{\xi \in [a; b)\} < \frac{b10^N - a10^N + 2}{10^N} \to b - a.$$

Таким образом, датчик случайных чисел приближённо можно описать с помощью распределения, для которого вероятность попадания в любой интервал внутри отрезка $[0; 1]$ равна его длине. Хорошо известно, что такая мера совпадает с мерой Лебега на $[0; 1]$.

Легко понять, что ф.распределения такой сл.в. равна

$$F(x) = \begin{cases} 0, \text{ если } & x \leqslant 0, \\ x, \text{ если } & 0 \leqslant x \leqslant 1, \\ 1, \text{ если } & x \geqslant 1. \end{cases}$$

Такое распределение называется *равномерным* в отрезке $[0; 1]$:

$$\xi \sim \mathrm{Unif}(0, 1).$$

В «синей книге» вводится показательное распределение из условия отсутствия последействия (условия нестарения) — вероятность того, что прибор прослужит ещё время t не зависит от того, сколько он уже прослужил. То есть, если ξ — время службы прибора, то

$$\mathbf{P}\{\xi \geqslant s+t \,|\, \xi \geqslant s\} = \mathbf{P}\{\xi \geqslant t\}, \quad \forall s, t \geqslant 0.$$

Дабы не повторяться, мы взглянем не это чуть с другой стороны. Найдём так называемую интенсивность отказов:

$$G(t) \;=\; \lim_{\Delta\to 0} \frac{1}{\Delta}\mathbf{P}\{t \leqslant \xi < t+\Delta \,|\, \xi \geqslant t\},$$

т.е. долю приборов, вышедших из строя в единицу времени, среди тех приборов, которые остались «живы» к моменту t.

По формуле условной вероятности

$$\mathbf{P}\{t \leqslant \xi < t+\Delta \,|\, \xi \geqslant t\} \;=\; \frac{\mathbf{P}\{t \leqslant \xi < t+\Delta\}}{\mathbf{P}\{\xi \geqslant t\}} = \frac{F(t+\Delta)-F(t)}{1-F(t)}$$

Если предположить, что функция распределения F имеет производную $F'(t)$ в любой точке t, то получим, что

$$G(t) \;=\; \frac{F'(t)}{1-F(t)}, \quad t \geqslant 0. \tag{$*$}$$

Определение. *Показательным* называется распределение, у которого интенсивность отказов не зависит от срока службы: $G(t)=\gamma, \;\; \forall t \geqslant 0.$

Легко понять, что правая часть ($*$) представляет собой так называемую логарифмическую производную функции $1-F$ (с точностью до знака). Поэтому условие постоянства интенсивности отказов эквивалентно равенству

$$\frac{d\ln(1-F(t))}{dt} = -\gamma, \;\; t \geqslant 0.$$

Тривиальными выкладками отсюда получаем, что показательная функция распределения

$$F(t) \;=\; 1-e^{-\gamma t}, \;\; t \geqslant 0.$$

При $t \leqslant 0$ полагаем $F(t)=0$.