§3.3. Функция распределения случайной величины

$P(X \le x)$ через $F(x)$

Функцией распределения называют функцию $F(x)$, определяющую вероятность того, что случайная величина X в результате испытания примет значение, меньше x, т.е. $F(x)= P(X \le x)$.

Функция распределения полностью характеризует дискретную и непрерывную случайную величины с вероятностной точки зрения, т.е. является одной из форм закона распределения.

Свойства функции распределения:

1. $0 \le F(x) \le 1$

2. $F(x_2) > F(x_1)$, если $x_2 > x_1$

Следствие 1. Вероятность того, что СВ примет значение в интервале от x_1 до x_2, равна приращению функции распределения на этом интервале: $P(x_1 \le X \le x_2) = F(x_2) - F(x_1)$.

Следствие 2. Вероятность того, что непрерывная СВ примет одно определенное значение, равна 0.

3. $F(-\infty)=0$

4. $F(\infty)=1$

$F(x)=P(X<x)=\sum\limits_{x_i<x} P(X<x)$

Для наглядности функцию распределения F(x) представляют в виде графика. Функция распределения F(x) в общем случае представляет собой график неубывающей функции с конечным числом точек разрывов (скачков), значения которой начинаются от 0 и кончаются 1.

Зная ряд распределения дискретной случайной величины, можно легко построить функцию распределения этой величины, определив ее выражением:

$$F(x)=P(X<x)=\sum_{x_i<x} P(X<x)$$

где $x_i<x$ под знаком суммы указывает, что суммирование распространяется на все значения x_i , которые меньше x. Когда текущая переменная x проходит через какое-нибудь из возможных значений дискретной случайной величины X, то функция распределения меняется скачкообразно, причем величина скачка равна вероятности этого значения.

Т.о., функция распределения любой дискретной случайной величины всегда есть разрывная ступенчатая функция, скачки которой происходят в точках, соответствующих возможным значениям случайной величины и равны вероятностям этих значений. Сумма всех скачков F(x) равна 1.

x	-10	5	10	15	20
p	0,30	0,40	0,25	0,04	0,01

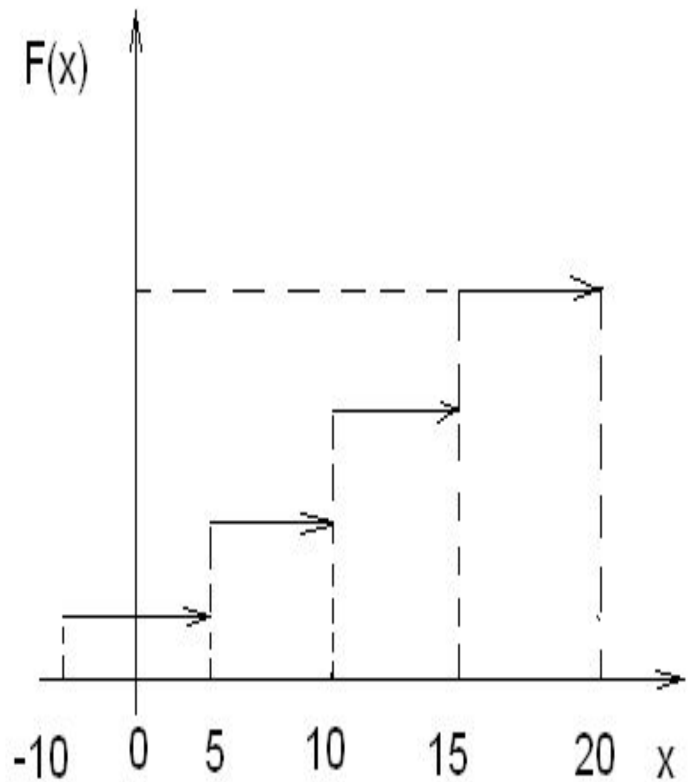


Рис.1

Пример. Брошена игральная кость. Случайная величина X –число выпавших очков. Написать закон распределения величины X и построить ее функцию распределения F(x).

Решение. Случайная величина *X* принимает возможные значения: {1, 2, 3, 4, 5, 6}. Вероятности этих значений p_k=P{X=k}=1/6, k=1.2.3.4.5.6. Следовательно, ряд распределений *X* имеет вид:

X	1	2	3	4	5	6
P	1/6	1/6	1/6	1/6	1/6	1/6

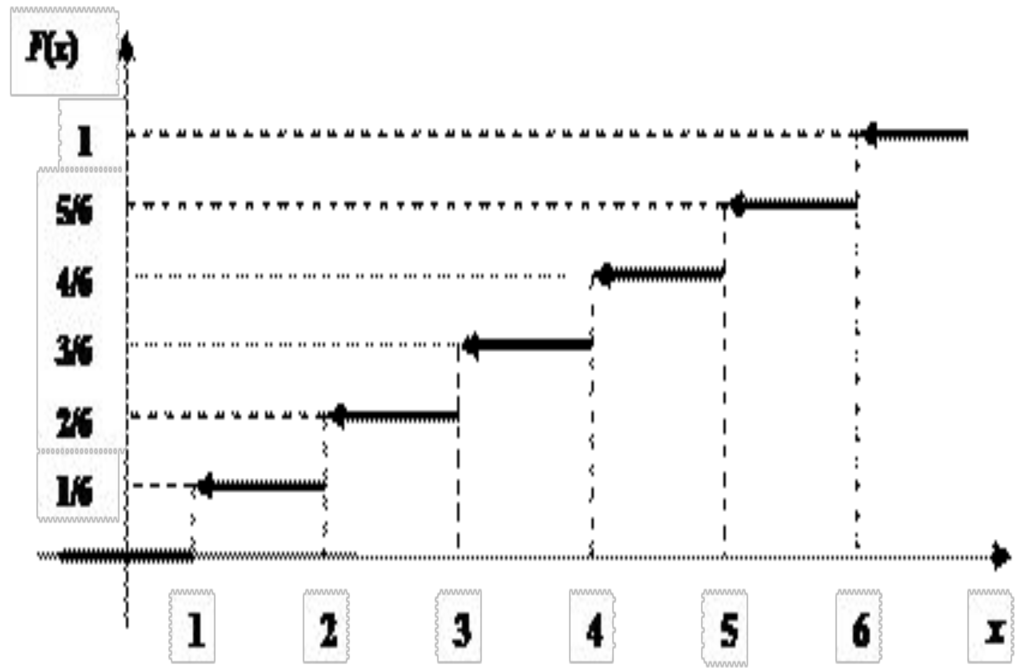
-

Вычислим функцию распределения величины X, используя формулу

$$F(x) = P\{X < x\} = \sum_{x_i < x} P\{X = x_i\} = \sum_{x_i < x} p_i$$

- $$F(x) = \begin{cases} 0 & \text{при } x \le 1 \\ p_1 = \frac{1}{6} & \text{при } 1 < x \le 2 \\ p_1 + p_2 = \frac{2}{6} & \text{при } 2 < x \le 3 \\ p_1 + p_2 + p_3 = \frac{3}{6} & \text{при } 3 < x \le 4 \\ p_1 + p_2 + p_3 + p_4 = \frac{4}{6} & \text{при } 4 < x \le 5 \\ p_1 + p_2 + p_3 + p_4 + p_5 = \frac{5}{6} & \text{при } 5 < x \le 6 \\ p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 1, & \text{при } x > 6 \end{cases}$$

Теперь построим график F(x):



Стрелки на графике F(x) обозначают одностороннюю непрерывность F(x) слева в каждой точке $x=x_i$.

§3.4. Плотность распределения случайной величины

от x до x+Δx   P(x ≤ X < x+Δx)=F(x+Δx)-F(x).

Рассмотрим отношение этой вероятности к длине участка Δx. При x→0

$$p(x) = \lim_{\Delta x \to 0} \frac{P(x \le X < x + \Delta x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x} = \frac{dF(x)}{dx}$$

dF=p(x)dx

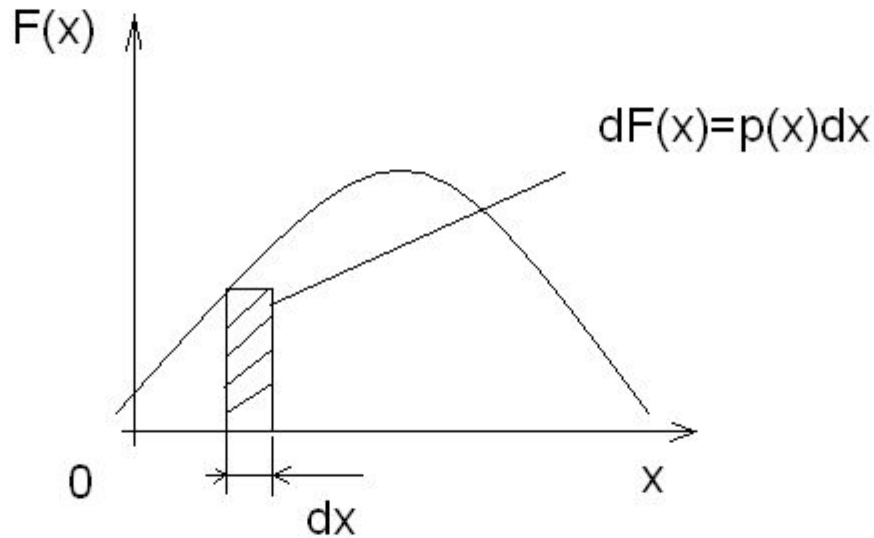


Рис.2

Свойства плотности распределения.

1. p(x)≥0. p(x)=F(x)≥0 F(x)= $\frac{F(x+\Delta x)-F(x)}{\Delta x}\geq 0$

2. $\int\limits_{-\infty}^{\infty} p(x)dx = 1$

$$\int\limits_{-\infty}^{\infty} p(x)dx = \int\limits_{-\infty}^{\infty} dF(x)F(x) = F(x)\Big|_{-\infty}^{\infty} = F(\infty)-F(-\infty)=1$$

(-∞≤X≤∞).

3. F(x)= $\int\limits_{-\infty}^{\infty} p(y)dy$

$$\int\limits_{-\infty}^{\infty} p(y)dy = \int\limits_{-\infty}^{\infty} dF(y) = F(x)-F(-\infty)=F(x)$$

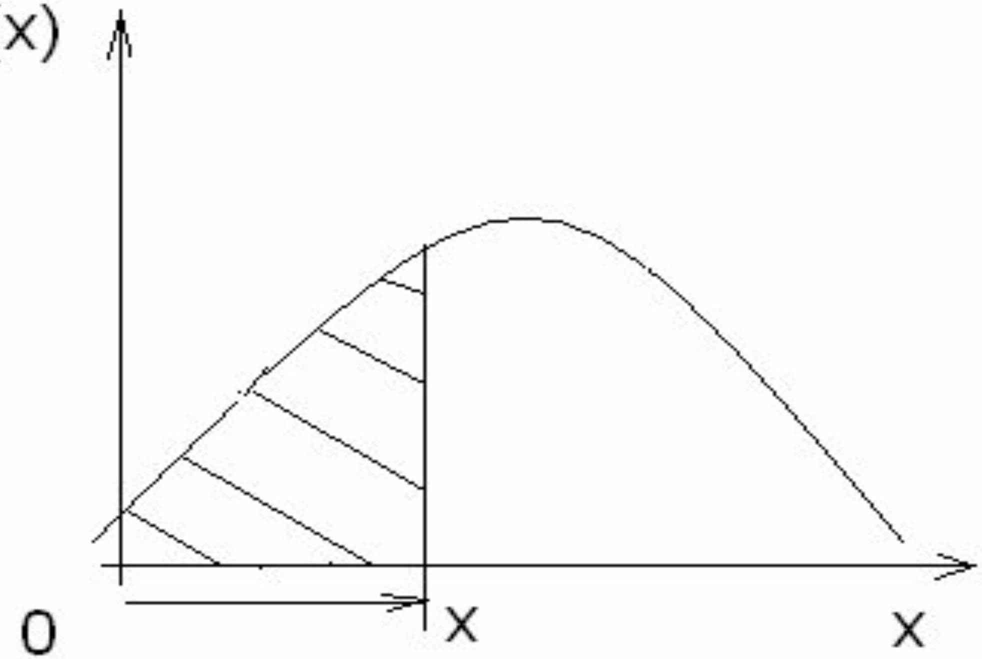


Рис.3

4. [a, b) P(a≤X<b)= $\int_a^b p(x)dx$

$$\int_a^b p(x)dx = \int_a^b dF(x) = F(x)\Big|_a^b = F(b) - F(a) = P(a \le X < b)$$

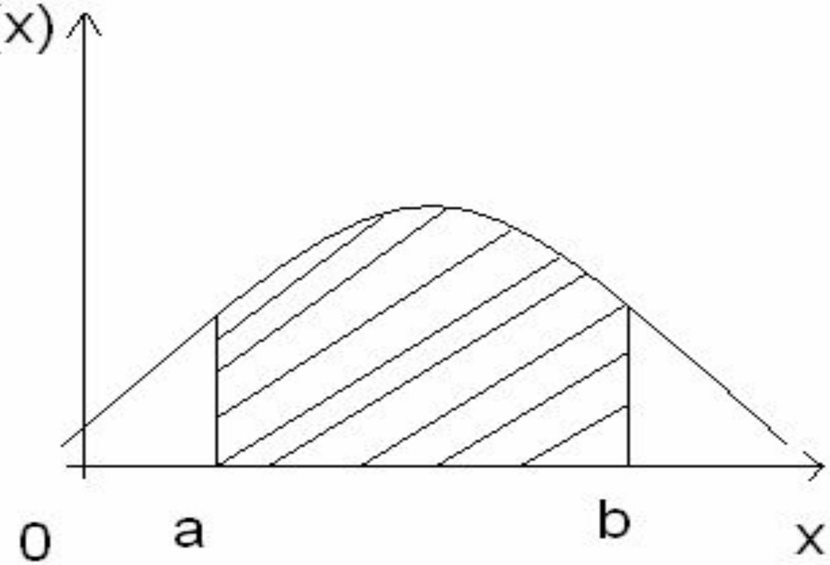


Рис.4

Вместо закона или функции распределения для описания СВ используется также так называемая характеристическая функция (ХФ). ХФ ϕ(u) СВ X определяется как математическое ожидание (МО) СВ e^{iuX}, т.е.

$$\varphi(u) = M\left\lfloor e^{iuX} \right\rfloor = \int_{-\infty}^{\infty} e^{iux} p(x) dx$$

где u – вещественная переменная, i= $\sqrt{-1}$ - мнимая единица.

Для распределения, обладающего плотностью p (x), ХФ является преобразованием Фурье функции p(x). Если p(x) удовлетворяет некоторым условиям, подробно рассматриваемым в теории интеграла Фурье, то p(x) можно восстановить по формуле:

p(x)= $\frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi(u)e^{-iux}du$

§3.5. Квантили

При решении практических задач часто требуется найти значение x, при котором функция распределения $F_x(x)$ случайной величины X принимает заданное значение α, т.е. требуется решить уравнение $F(x) = \alpha$. Решения такого уравнения (соответствующие значения x) в теории вероятностей называются *квантилями.*

α-квантиль (квантиль порядка α) – это числовая характеристика закона распределения случайной величины. α-квантиль – такое число, что данная случайная величина попадает левее его с вероятностью, не превосходящей α.

α-квантиль случайной величины X с функцией распределения $F(x) = P(X<x)$ – это любое число α, удовлетворяющее двум условиям:

$$F(x_\alpha) \leq \alpha,\ F(x_\alpha+0) \geq \alpha.$$

Данные условия эквивалентны следующим:

$$P(X<x_{\alpha}) \leq \alpha, \; P(X>x_{\alpha}) \geq 1 - \alpha \; .$$

Если $F(X)$– непрерывная строго монотонная функция, то существует единственный квантиль x_{α} любого порядка $\alpha \in (0,1)$, который однозначно определяется из уравнения $F(x_{\alpha}) = \alpha$ и выражается через функцию, обратную к функции распределения: $x_{\alpha} = F^{-1}(\alpha)$.

Кроме рассмотренного случая, когда уравнение $F(x_\alpha) = \alpha$ имеет единственное решение и дает соответствующий квантиль, возможны следующие случаи:

– уравнение $F(x_\alpha) = \alpha$ не имеет решений. Значит, существует единственная точка x_α, в которой функция распределения имеет разрыв, которая удовлетворяет данному определению и является квантилем порядка α. Для этой точки выполнены соотношения:

$$P(X<x_\alpha)< P(X>x_\alpha)\leq 1 - \alpha \ ;$$

– уравнение $F(x_\alpha) = \alpha$ имеет более одного решения. Значит, все его решения образуют интервал, на котором функция распределения постоянна. В качестве квантиля α может быть взята любая точка этого интервала. Содержательные выводы, в которых участвует квантиль, от этого существенно не изменятся, поскольку вероятность попадания случайной величины X в данный интервал равна нулю.

Если возникает необходимость отделить сверху, снизу или с обеих сторон области, вероятности попадания в которые малы, то используется следующая терминология:

– нижний (односторонний) квантиль уровня α. Обычный квантиль порядка α;

– верхний (односторонний) квантиль уровня α. Обычный квантиль порядка 1-α;

– двусторонние квантили уровня α. Пара (нижний + верхний) односторонних квантилей уровня α/2.

Квантилью α (α – квантилью, квантилью уровня α) случайной величины X, имеющей функцию распределения $F(x)$, называют решение x_α уравнения $F(x) = \alpha$, $\alpha \in (0, 1)$.

Квантиль – это общее понятие.

Частными случаями квантиля являются: квартили; децили; процентили.

Квантили, наиболее часто встречающиеся в практических задачах, имеют свои названия:

– *медиана* – квантиль уровня 0.5 – $x_{1/2}$ – средний показатель распределения;

– *квартиль* –$x_{p/4}$, где p=1, 2, 3. Указывает на место расположения данных распределения. Когда значение находится в зоне, где расположено менее 25% наблюдаемых значений переменной, то говорят, что оно расположено в нижнем квартале (*нижняя квартиль*).

Если же оно расположено там, где находятся верхние 25% значений, то говорят, что оно расположено в верхнем квартале (*верхняя квартиль* – квантиль уровня 0.75);

– *дециль* – $x_{p/10}$, где p=1, …, 9. Граница десятой части распределения. Например, если все доходы сгруппированы в убывающем порядке, первым децилем будет доход, выше которого находятся 10% представленных в списке доходов, а ниже – остальные

90% доходов (квантили уровней 0.1, 0.2, …, 0.9);

– *процентили* – $x_{p/100}$, где p=1, …, 99. Значения, выделяющие сотые части распределения, выстроенные в ряд по их величине. Например, 99-я процентиль распределения дохода представляет собой такой уровень дохода, когда только один процент населения имеет больший доход (квантили уровней 0.01, 0.02, …, 0.99).