LECTURE 8

Distribution function of a random variable

For a random variable X, the function F defined by

$$F(x) = P(X < x), \qquad -\infty < x < \infty$$

is called *the cumulative distribution function* or, more simply, *the distribution function* of X. Thus the distribution function specifies, for all real values x, the probability that the random variable is less than x.

Sometimes the distribution function $F(x)$ is said to be the *integral function of distribution* or the *integral law of distribution*.

Geometrically the distribution function is interpreted as the probability that a random variable X will hit to the left from a given point x.

Example 1. Let a series of distribution of a random variable be given:

$$X = \begin{pmatrix} 1 & 4 & 5 & 7 \\ 0{,}4 & 0{,}1 & 0{,}3 & 0{,}2 \end{pmatrix}$$

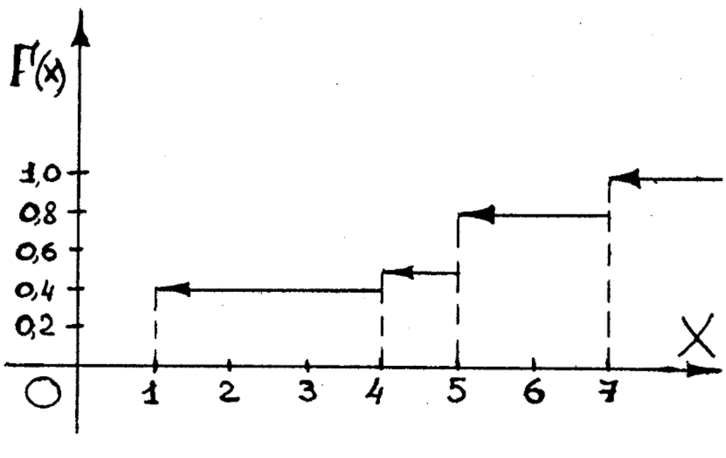
Find and represent graphically its distribution function.

Solution: Let's take different values x and find for them $F(x) = P(X < x)$.

1. If $x \le 1$ then obviously $F(x) = 0$ (and for $x = 1$, $F(1) = P(X < 1) = 0$).
2. Let $1 < x \le 4$ (for example, $x = 2$); $F(x) = P(X = 1) = 0{,}4$. Obviously, $F(4) = P(X < 4) = 0{,}4$.
3. Let $4 < x \le 5$ (for example, $x = 4{,}2$); $F(x) = P(X < x) = P(X = 1) + P(X = 4) = 0{,}4 + 0{,}1 = 0{,}5$. Obviously, $F(5) = 0{,}5$.
4. Let $5 < x \le 7$. $F(x) = [P(X = 1) + P(X = 4)] + P(X = 5) = 0{,}5 + 0{,}3 = 0{,}8$. Obviously, $F(7) = 0{,}8$.
5. Let $x > 7$. $F(x) = [P(X = 1) + P(X = 4) + P(X = 5)] + P(X = 7) = 0{,}8 + 0{,}2 = 1$.

Thus,

$$F(x) = \begin{cases} 0 & if \quad x \le 1, \\ 0{,}4 & if \ 1 < x \le 4, \\ 0{,}5 & if \ 4 < x \le 5, \\ 0{,}8 & if \ 5 < x \le 7, \\ 1{,}0 & if \quad x > 7. \end{cases}$$



This example allows making the following claim: *the distribution function of any discrete random variable is a discontinuous step function of which jumps take place in points corresponding to the possible values of the random variable and are equal to the probabilities of these values.*

The sum of all jumps is equal to 1.

Properties of a distribution function

1. The distribution function of a random variable is a non-negative function taking on values between 0 and 1: $0 \le F(x) \le 1$.

2. The distribution function of a random variable is a non-decreasing function for the entire numerical axis, i.e. if $x_1 < x_2$ then $F(x_1) \le F(x_2)$.

3. $F(-\infty) = \lim_{x\to-\infty} F(x) = 0, \quad F(+\infty) = \lim_{x\to+\infty} F(x) = 1.$

4. The probability of hit of a random variable in an interval $[x_1, x_2)$ (including x_1) is equal to the increment of its distribution function on this interval, i.e.

$$P(x_1 \le X < x_2) = F(x_2) - F(x_1).$$

Example 2. The distribution function of a random variable X has the following form:

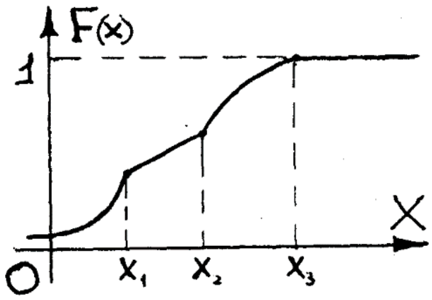
$$F(x) = \begin{cases} 0 & if \quad x \le 0, \\ x/2 & if \quad 0 < x \le 2, \\ 1 & if \quad x > 2. \end{cases}$$

Find the probability that the random variable will take on a value in the interval [1; 3).

Solution: $P(1 \le X < 3) = F(3) - F(1) = 1 - 1/2 = 1/2$.

Continuous random variables. Probability density

A random variable X is *continuous* if its function of distribution is continuous at each point and differentiable everywhere but possibly finitely many points.



The distribution function of a continuous random variable X which is differentiable everywhere but three points of break is shown at the picture.

Theorem. The probability of a separately taken value of a continuous random variable X is equal to zero, i.e. $P(X = x_1) = 0$ for each value x_1 of the continuous random variable X.

Corollary. If X is a continuous random variable then probability of hit of the random variable in an interval (x_1, x_2) doesn't depend on whether the interval is open or closed, i.e.

$$P(x_1 < X < x_2) = P(x_1 \le X < x_2) = P(x_1 < X \le x_2) = P(x_1 \le X \le x_2)$$

The probability density (*distribution density* or simply *density*) $\varphi(x)$ (or $f(x)$) of a continuous random variable X is the derivative of its distribution function, i.e.

$$\boxed{f(x) = \varphi(x) = F'(x)}$$

Sometimes the probability density is said to be the *differential function* or the *differential law of distribution*.

The graph of probability density $\varphi(x)$ is said to be *distribution curve*.

Example. Let X be a random variable as in Example 2. Find the probability density of the random variable X.

Solution: The probability density $\varphi(x) = F'(x),$ i.e.

$$\varphi(x) = \begin{cases} 0 & if \quad x \le 0 \ or \ x > 2, \\ 1/2 & if \quad 0 < x \le 2. \end{cases}$$

Properties of probability density

1. The probability density is a non-negative function, i.e. $\varphi(x) \ge 0$.
2. The probability of hit of a continuous random variable in an interval [*a*, *b*] is equal to the definite integral of its density in limits from *a* to *b*, i.e.

$$\boxed{P(a \le X \le b) = \int_a^b \varphi(x)dx}$$

3. The distribution function of a continuous random variable can be expressed by the probability density:

$$\boxed{F(x) = \int_{-\infty}^{x} \varphi(t)dt}$$

4. The improper integral in infinite limits of the probability density of a continuous random variable is equal to 1:

$$\boxed{\int_{-\infty}^{+\infty} \varphi(x)dx = 1}$$

Formulas of mathematical expectation and dispersion of a continuous random variable *X* have the following form:

$$\boxed{a = M(X) = \int_{-\infty}^{+\infty} x\varphi(x)dx}$$ (if the integral converges absolutely)

$$\boxed{D(X) = \int_{-\infty}^{+\infty} (x-a)^2 \varphi(x)dx}$$ (if the integral converges).

Using $D(X) = M(X^2) - [M(X)]^2$, we have $D(X) = M(X^2) - a^2$ or

$$\boxed{D(X) = \int_{-\infty}^{+\infty} x^2 \varphi(x)dx - a^2}$$

<u>*Example*</u>. Let a function $\varphi(x)$ be given: $\varphi(x) = \begin{cases} 0 & if \ \ x \le 1, \\ A/x^4 & if \ \ x > 1. \end{cases}$

Find: a) the value of the constant *A* for which the function is the probability density of some random variable *X*; b) the expression of the distribution function $F(x)$; c) the probability that the random variable X will take on values in the segment [2; 3]; d) the mathematical expectation and the dispersion of the random variable X.

Solution: a) $\varphi(x)$ will be the probability density if $\varphi(x) \ge 0$, i.e. $A/x^4 \ge 0$ (and consequently, $A \ge 0$) and it satisfies the condition $\int_{-\infty}^{+\infty} \varphi(x)dx = 1.$ Consequently,

$$\int_{-\infty}^{+\infty} \varphi(x)dx = \int_{-\infty}^{1} 0 \cdot dx + \int_{1}^{+\infty} \frac{A}{x^4}dx = 0 + \lim_{b\to+\infty}\int_{1}^{b} \frac{A}{x^4}dx = \frac{A}{3}\lim_{b\to+\infty}\left(-\frac{1}{x^3}\Big|_1^b\right) =$$

$$= \frac{A}{3}\lim_{b\to+\infty}\left(1-\frac{1}{b^3}\right) = \frac{A}{3} = 1.$$

We have $A = 3$.

b) Find $F(x)$: If $x \le 1$ then $F(x) = \int_{-\infty}^{x} \varphi(x)dx = \int_{-\infty}^{x} 0 \cdot dx = 0.$

If $x > 1$ then $F(x) = 0 + \int_{1}^{x} \frac{3}{x^4}dx = -\frac{1}{x^3}\Big|_1^x = 1 - \frac{1}{x^3}\cdot$ Thus,

$$F(x) = \begin{cases} 0 & if \;\; x \le 1, \\ 1 - 1/x^3 & if \;\; x > 1. \end{cases}$$

c) $P(2 \le X \le 3) = \int_{2}^{3} \frac{3}{x^4}dx = -\frac{1}{x^3}\Big|_2^3 = \frac{1}{2^3} - \frac{1}{3^3} = \frac{19}{216}.$

The probability $P(2 \le X \le 3)$ can also be found as the increment of the distribution function, i.e.

$$P(2 \le X \le 3) = F(3) - F(2) = \left(1-\frac{1}{3^3}\right) - \left(1-\frac{1}{2^3}\right) = \frac{19}{216}.$$

d) $a = M(X) = \int_{-\infty}^{+\infty} x\varphi(x)dx = \int_{-\infty}^{1} 0 \cdot dx + \int_{1}^{+\infty} x \cdot \frac{3}{x^4}dx = 0 + 3\lim_{b\to+\infty}\int_{1}^{b} \frac{dx}{x^3} =$

$$= 3\lim_{b\to+\infty}\left(-\frac{1}{2x^2}\Big|_1^b\right) = \frac{3}{2}.$$

Since $D(X) = M(X^2) - a^2$, $M(X^2) = \int_{-\infty}^{+\infty} x^2\varphi(x)dx = \int_{-\infty}^{1} x^2 \cdot 0\,dx + \int_{1}^{+\infty} x^2 \cdot \frac{3}{x^4}dx =$

$= \lim_{b\to+\infty}\left(-\frac{3}{x}\Big|_1^b\right) = \lim_{b\to+\infty}\left(3 - \frac{3}{b}\right) = 3.$ Then $D(X) = 3 - \left(\frac{3}{2}\right)^2 = \frac{3}{4}.$

Glossary

distribution function – функция распределения
jump of a function – скачок функции
step discontinuous function – ступенчатая разрывная функция
increment – приращение; **distribution curve** – кривая распределения
probability density – плотность вероятности

Exercises for Seminar 8

8.1. Let the law of distribution of a discrete random variable be given:

X	1	4	6	8
P	0,1	0,3	0,4	0,2

Find the integral function of the random variable X and construct its graph.

8.2. Find the integral function of distribution of the random variable X – the number of hits in a target if three shots were made with the probability of hit in the target equal 0,8 for each shot.

8.3. A continuous random variable X is given by the integral function:

$$F(x) = \begin{cases} 0 & if \quad x \le 0, \\ \dfrac{x^3}{125} & if \quad 0 < x \le 5, \\ 1 & if \quad x > 5. \end{cases}$$

Determine:

a) the probability of hit of the random variable into the interval (2; 3);

b) the mathematical expectation, the dispersion and the mean square deviation of the random variable X.

8.4. A random variable X is given by the integral function:

$$F(x) = \begin{cases} 0 & if \quad x \le -2, \\ \dfrac{x}{4} + \dfrac{1}{2} & if \quad -2 < x \le 2, \\ 1 & if \quad x > 2. \end{cases}$$

Find the probability that in result of a trial the random variable X will take on the value:

(a) less than 0;
(b) less than 1;
(c) no less than 1;
(d) being in the interval (0; 2).

8.5. The amount of time, in hours, that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f(x) = \begin{cases} \lambda e^{-x/100} & if \quad x \ge 0 \\ 0 & if \quad x < 0 \end{cases}$$

What is the probability that

(a) a computer will function between 50 and 150 hours before breaking down;
(b) it will function less than 100 hours?

Direction: Take e equal 2,718281.

The answer: a) 0,3834; b) 0,632.

8.6. The lifetime in hours of a certain kind of radio tube is a random variable having a probability density function given by

$$f(x) = \begin{cases} 0 & if \quad x \le 100 \\ \dfrac{100}{x^2} & if \quad x > 100 \end{cases}$$

What is the probability that exactly 2 of 5 such tubes in a radio set will have to be replaced within the first 150 hours of operation? Assume that the events E_i, i = 1, 2, 3, 4, 5, that the i^{th} such tube will have to be replaced within this time, are independent.

The answer: 80/243.

8.7. Let X be a random variable with probability density function

$$f(x) = \begin{cases} C(1-x^2) & if \quad -1 < x < 1 \\ 0 & otherwise \end{cases}$$

(a) What is the value of C?

(b) What is the cumulative distribution function of X?

The answer: a) 3/4.

8.8. Compute $M(X)$ if X has a density function given by

$$f(x) = \begin{cases} \frac{1}{4} x e^{-x/2} & if \quad x > 0 \\ 0 & otherwise \end{cases}$$

The answer: 4.

8.9. A random variable X is given by the differential function:

$$f(x) = \begin{cases} 0 & if \quad x \le 0, \\ \frac{4a-2x}{3a^2} & if \quad 0 < x \le a, \\ 0 & if \quad x > a. \end{cases}$$

Find:

(a) the integral function;

(b) the probability of hit of the random variable into the interval (a/6; a/3).

The answer: b) 7/36.

8.10. A random variable X is given by the integral function:

$$F(x) = \begin{cases} 0 & if \quad x \le 2, \\ \frac{x^3-8}{19} & if \quad 2 < x \le 3, \\ 1 & if \quad x > 3. \end{cases}$$

Find:

(a) the differential function;

(b) the probability of hit of the random variable X into the interval (2,5; 3);

(c) the mathematical expectation, the dispersion and the mean square deviation of the random variable X.

The answer: b) 0,599; c) $M(X)$ = 2,566; $D(X)$ = 0,079.

Exercises for Homework 8

8.11. Let the law of distribution of a discrete random variable be given:

X	– 2	5	7	9
p	0,4	0,3	0,2	0,1

Find the integral function of the random variable X and construct its graph.

8.12. The probability of passing the first exam by a student is 0,7, the second exam – 0,6 and the third exam – 0,8. Find the integral function of the random variable *X* – the number of exams passed by the student. Determine *M(X).*

The answer: M(X) = 2,1.

8.13. The density function of *X* is given by

$$f(x)=\begin{cases} a+bx^2 & if \quad 0\le x\le 1 \\ 0 & otherwise \end{cases}$$

If *M(X)* = 3/5, find *a* and *b*.

The answer: a = 3/5; *b* = 6/5.

8.14. A system consisting of one original unit and a spare can function for a random amount of time *X*. If the density of *X* is given (in units of months) by

$$f(x)=\begin{cases} Cxe^{-x/2} & if \quad x>0 \\ 0 & if \quad x\le 0 \end{cases}$$

what is the probability that the system functions for at least 5 months (a spare – запасной элемент)?

The answer: 0,616.

8.15. Suppose that *X* is a continuous random variable whose probability density function is given by

$$f(x)=\begin{cases} C(4x-2x^2) & if \quad 0<x<2 \\ 0 & otherwise \end{cases}$$

(a) What is the value of *C*?

(b) Find *P*(*X* > 1).

The answer: a) 3/8; b) 1/2.

8.16. Find *M(X)* and *D(X)* when the density function of *X* is

$$f(x)=\begin{cases} 2x & if \quad 0\le x\le 1 \\ 0 & otherwise \end{cases}$$

The answer: M(X) = 2/3; *D(X)* = 1/18.

8.17. Let *X* be a random variable with probability density function

$$f(x)=\begin{cases} C(2x-x^3) & if \quad 0<x<\dfrac{5}{2} \\ 0 & otherwise \end{cases}$$

(a) What is the value of *C*?

(b) What is the cumulative distribution function of *X*?

The answer: a) – 64/225.

8.18. Compute *M(X)* if *X* has a density function given by

$$f(x)=\begin{cases} \dfrac{50}{x^3} & if \quad x>5 \\ 0 & otherwise \end{cases}$$

The answer: 10.

8.19. A random variable X is given by the differential function:

$$f(x) = \begin{cases} 0 & if \quad x \le 0, \\ x^3 + x & if \quad 0 < x \le \sqrt{\sqrt{5} - 1}, \\ 0 & if \quad x > \sqrt{\sqrt{5} - 1}. \end{cases}$$

Find:

(a) the integral function;

(b) the probability of hit of the random variable into the interval (1; 1,1).

The answer: b) 0,221.

8.20. A random variable X is given by the integral function:

$$F(x) = \begin{cases} 0 & if \quad x \le 1, \\ \frac{x^2}{2} - \frac{x}{2} & if \quad 1 < x \le 2, \\ 1 & if \quad x > 2. \end{cases}$$

Find:

(a) the differential function;

(b) the probability of hit of the random variable X into the interval (1; 1,5);

(c) the mathematical expectation, the dispersion and the mean square deviation of the random variable X.

The answer: b) 0,375; c) $M(X)$ = 19/12; $D(X)$ = 11/144.