PhD Misiura Ie.Iu. (доцент Місюра Є.Ю.)

Theme: Random variables and their economic interpretation. Basic laws of distribution

Part 1: discrete random variable

A definition of random variables and their classification

A variable is called ***random***, if it can receive real values with definite probabilities as a result of experiment.

A random variable / випадкова величина

A definition of random variables and their classification

A variable is called ***random***, if it can receive real values with definite probabilities as a result of experiment.

Definition

A random variable is a variable that takes values according to a certain probability distribution.

- Keys to know:
  - All the possible values it can take.
  - The probability distribution according to which it takes all possible values.

A definition of random variables and their classification

A variable is called ***random***, if it can receive real values with definite probabilities as a result of experiment.

In general, random variables can be **discrete** or **continuous**.

A discrete random variable /
дискретна випадкова величина

A continuous random variable /
неперервна випадкова величина

A definition of random variables and their classification

A random variable is a variable whose value is unknown, or a function that assigns values to each of an experiment's outcomes. Random variables are often designated by letters and can be classified as discrete, which are variables that have separated values, or continuous, which are variables that can have any values within a continuous range.

A random variable is denoted by X, Y, Z and so on and its possible values are denoted by x_i, y_i, z_i

For example, if X is a random variable, then its values are $x_1, x_2, \quad , x_n$ (these values form a complete group of events, therefore

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

DISCRETE RANDOM VARIABLE

The random variable X is called ***discrete***, if such non-negative function exists

$$P(X = x_i) = p_i \qquad i = \overline{1, n} \qquad \sum_{i=1}^{n} p_i = 1$$

which determines the correspondence between the value x_i of the variable X and the probability p_i that X receives this value.

A **discrete variable** is a variable whose value is obtained by counting.

Examples: number of students present

number of red marbles in a jar

number of heads when flipping three coins

students' grade level

Random Variables

According to the range of a random variable, we can define two types of random variables.

- Discrete random variables:
  - The range is finite or countably infinite.[1]
    - (1) The number of defective light bulbs in a box of ten (finite)
    - (2) The number of tails until the first heads comes up (countably infinite)
- Continuous random variables:
  - The range is uncountable (any value in an interval)
    - (1) Consider the experiment where a light bulb is tested until failure, and let X denote the time to failure.

$$\text{Range of } X = [0, \infty)$$

[1] A set of elements is countably infinite if the elements in the set can be put into one-to-one correspondence with the positive integers.

DISTRIBUTION LAW

The ***distribution law (row)*** of a random variable is called a set of all its possible values and probabilities which these values possess. It's often written in the form of a table

A distribution law (row) /
закон розподілу

DISTRIBUTION LAW OF A DISCRETE RANDOM VARIABLE

The ***distribution law (row)*** of a discrete random variable is called a set of all its possible values and probabilities which these values possess. It's often written in the form of a table

x_i	x_1	x_2	$\cdots$	x_n
p_i	p_1	p_2	$\cdots$	p_n

Example: Flip a Fair Coin Twice

Example: Flip a Fair Coin Twice

- The sample space is

$$\Omega = \{HH, HT, TH, TT\}$$

Example: Flip a Fair Coin Twice

- The sample space is

$$\Omega = \{HH, HT, TH, TT\}$$

- Let X be the number of heads.
  - The range of X is $\{0, 1, 2\}$

Example: Flip a Fair Coin Twice

- The sample space is

$$\Omega = \{HH, HT, TH, TT\}$$

- Let X be the number of heads.
  - The range of X is $\{0, 1, 2\}$
- For example,

$$\begin{aligned} P(X = 1) &= P(\{\omega : X(\omega) = 1\}) \\ &= P(\{HT\}) + P(\{TH\}) \\ &= \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \end{aligned}$$

Example. Two balls are drawn in succession without replacement from an urn containing 4 red balls and 3 black balls. Construct a distribution law.

Example. Two balls are drawn in succession without replacement from an urn containing 4 red balls and 3 black balls. Construct a distribution law.

Solution.

Let’s denote ***X*** as the number of red balls.

Example. Two balls are drawn in succession without replacement from an urn containing 4 red balls and 3 black balls. Construct a distribution law.

Solution.

Let's denote ***X*** as the number of red balls.

HOW MANY OUTCOMES ARE POSSIBLE?

X is the number of red balls

SAMPLE SPACE	x_i
RR	
RB	
BR	
BB	

X is the number of red balls

SAMPLE SPACE	x_i
RR	**2**
RB	**1**
BR	**1**
BB	**0**

Let’s find the probability of each value x_i

$$P(X=0)=P(BB)=\frac{3}{7}\cdot\frac{2}{6}=\frac{1}{7}$$

Let's find the probability of each value x_i

$$P(X=0)=P(BB)=\frac{3}{7}\cdot\frac{2}{6}=\frac{1}{7}$$

$$P(X=1)=P(RB+BR)=\frac{4}{7}\cdot\frac{3}{6}+\frac{3}{7}\cdot\frac{4}{6}=\frac{2}{7}+\frac{2}{7}=\frac{4}{7}$$

Let’s find the probability of each value x_i

$$P(X=0)=P(BB)=\frac{3}{7}\cdot\frac{2}{6}=\frac{1}{7}$$

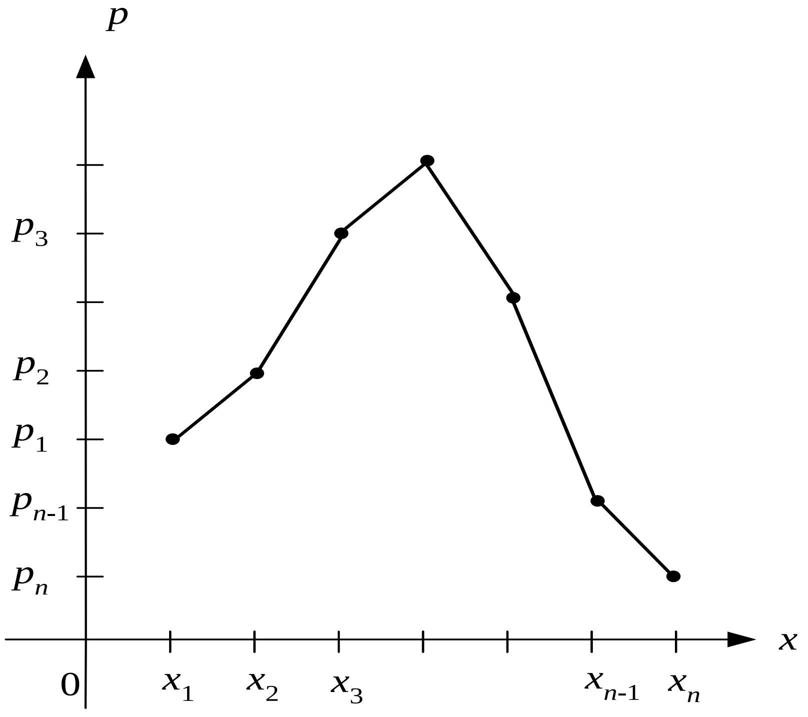
$$P(X=1)=P(RB+BR)=\frac{4}{7}\cdot\frac{3}{6}+\frac{3}{7}\cdot\frac{4}{6}=\frac{2}{7}+\frac{2}{7}=\frac{4}{7}$$

$$P(X=2)=P(RR)=\frac{4}{7}\cdot\frac{3}{6}=\frac{2}{7}$$

Let’s write the distribution law of this discrete random variable:

x_i	0	1	2
p_i	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{4}{7}$

Distribution law (row) can be graphically plotted. Values of a variable are marked on ***x***-axis, the corresponding probabilities are marked on ***p***-axis. The obtained points are connected with the help of segments. It results in *a* ***distribution polygon*** ***(багатокутник розподілу)***.



Example. Distribution law of a discrete random variable X

x_i	-2	2	6	10	14
p_i	0.05	0.16	0.35	0.31	0.13

is given. Draw a distribution polygon.

Example. Distribution law of a discrete random variable X

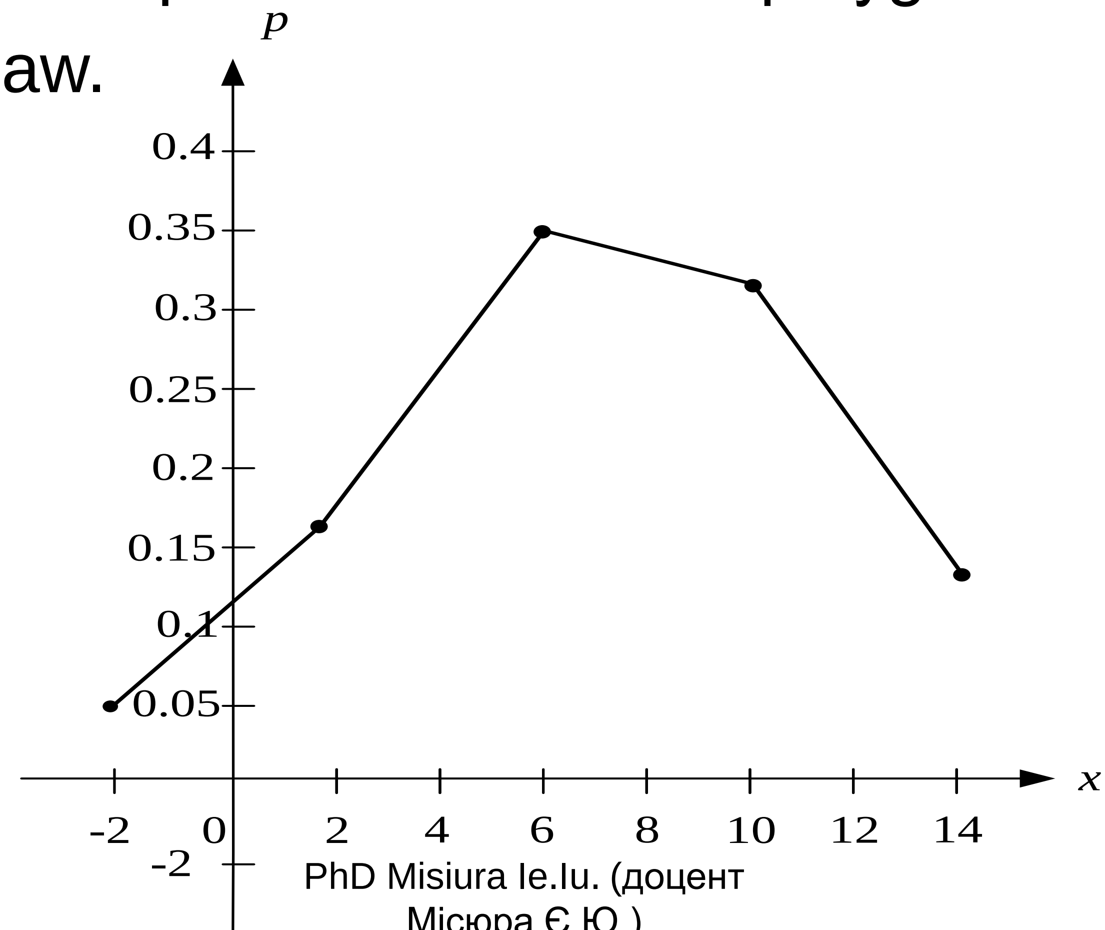
x_i	-2	2	6	10	14
p_i	0.05	0.16	0.35	0.31	0.13

is given. Draw a distribution polygon.

Solution. Let's plot the distribution polygon for the given distribution law.

x_i	-2	2	6	10	14
p_i	0.05	0.16	0.35	0.31	0.13

Solution. Let's plot the distribution polygon for the given distribution law.



A *discrete* random variable is characterized by one function:

a distribution function of probabilities

$$F(x)$$

The distribution function

The probability of the fact that a random variable X receives a value less than x, is called a ***cumulative distribution function*** of a random variable X and is marked as $F(x)$:

$$F(x) = P(X < x)$$

General properties of the cumulative distribution function:

1. $F(x)$ is a bounded function, i.e. $0 \le F(x) \le 1$
2. $F(x)$ is a non-decreasing function for $x \in (-\infty, \infty)$
3. $\lim\limits_{x \to -\infty} F(x) = F(-\infty) = 0$
4. $\lim\limits_{x \to +\infty} F(x) = F(+\infty) = 1$
5. The probability that a random variable X lies in the interval (x_1, x_2) is equal to the increment of its cumulative distribution function on this interval, i.e.

$$P(x_1 < X < x_2) = F(x_2) - F(x_1)$$

The distribution function

x_1	x_2	...	x_i	...
p_1	p_2	...	p_i	...

$$F(x) = \begin{cases} 0, & x < x_1, \\ p_1, & x_1 \le x < x_2, \\ p_1 + p_2, & x_2 \le x < x_3, \\ \ldots, & \\ p_1 + p_2 + \ldots + p_{n-1}, & x_{n-1} \le x < x_n, \\ 1, & x \ge x_n. \end{cases}$$

Example. Find the distribution function of the random variable X , which is defined by the distribution law:

x_i	1	2	3	4
p_i	0.7	0.21	0.063	0.027

Find the probability that the random variable X possesses a value less than 1 and more than 4.

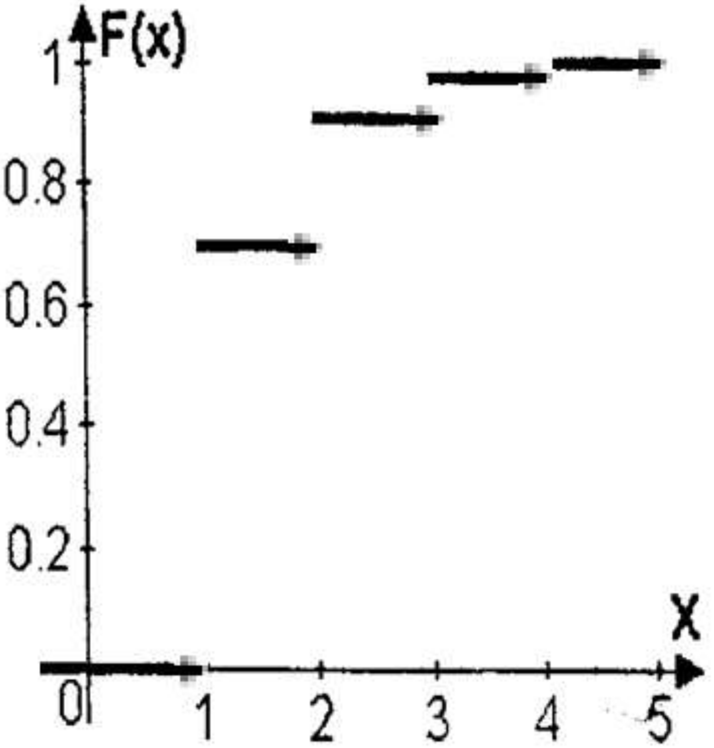
x_i	1	2	3	4
p_i	0.7	0.21	0.063	0.027

The distribution cumulative function is

$$F(x)=\begin{cases}0, & if\ x<1\\ 0.7, & if\ 1\le x<2\\ 0.91, & if\ 2\le x<3\\ 0.973, & if\ 3\le x<4\\ 1, & if\ x\ge 4\end{cases}$$

The distribution cumulative function is

$$F(x)=\begin{cases}0, & if\ x<1\\ 0.7, & if\ 1\le x<2\\ 0.91, & if\ 2\le x<3\\ 0.973, & if\ 3\le x<4\\ 1, & if\ x\ge 4\end{cases}$$



Let’s find the probability that a random variable possesses a value less than 1 and more than 4, i.e. :

$$P(x_1 < X < x_2) = F(x_2) - F(x_1)$$

$$P(1 < X < 4) = F(4) - F(1) = 1 - 0.7 = 0.3$$

$$F(x) = \begin{cases} 0, & if\ x < 1 \\ 0.7, & if\ 1 \le x < 2 \\ 0.91, & if\ 2 \le x < 3 \\ 0.973, & if\ 3 \le x < 4 \\ 1, & if\ x \ge 4 \end{cases}$$

Numerical characteristics of discrete random variables

The numerical characteristics of discrete random variables

The ***mathematical expectation*** of a discrete random variable X is called a sum of products of possible values x_i , which a variable X is taken, and their corresponding probabilities p_i .

$$M(X)=\sum_{i=1}^{n} x_i \cdot p_i = x_1 \cdot p_1 + x_2 \cdot p_2 + \ldots + x_n \cdot p_n$$

The numerical characteristics of distribution

The ***variance*** of a random variable is called a mathematical expectation of deviation square of a random variable X from its mathematical expectation $M(X)$, i.e.:.

$$D(X) = M\left[\left(X - M(X)\right)^2\right] = \sum_{i=1}^{n}\left(x_i - M(X)\right)^2 \cdot p_i$$

The numerical characteristics of distribution

or ***variance*** of a random variable X equals a mathematical expectation of its square minus a square of its mathematical expectation $M(X)$

i.e.:

$$D(X) = M(X^2) - [M(X)]^2$$

where

$$M(X^2) = \sum_{i=1}^{n} x_i^2 \cdot p_i$$

The numerical characteristics of a random variable

Root-mean-square deviation *(*or ***standard deviation****)* of a random variable X is the square root of its variance $D(X)$, i.e.:

$$\sigma(X) = \sqrt{D(X)}$$

4. Coefficient of variation is a ratio of a root-mean-square deviation and a mathematical expectation in percent:

$$\nu(X)=\frac{\sigma(X)}{M(X)}\cdot 100\,\%$$

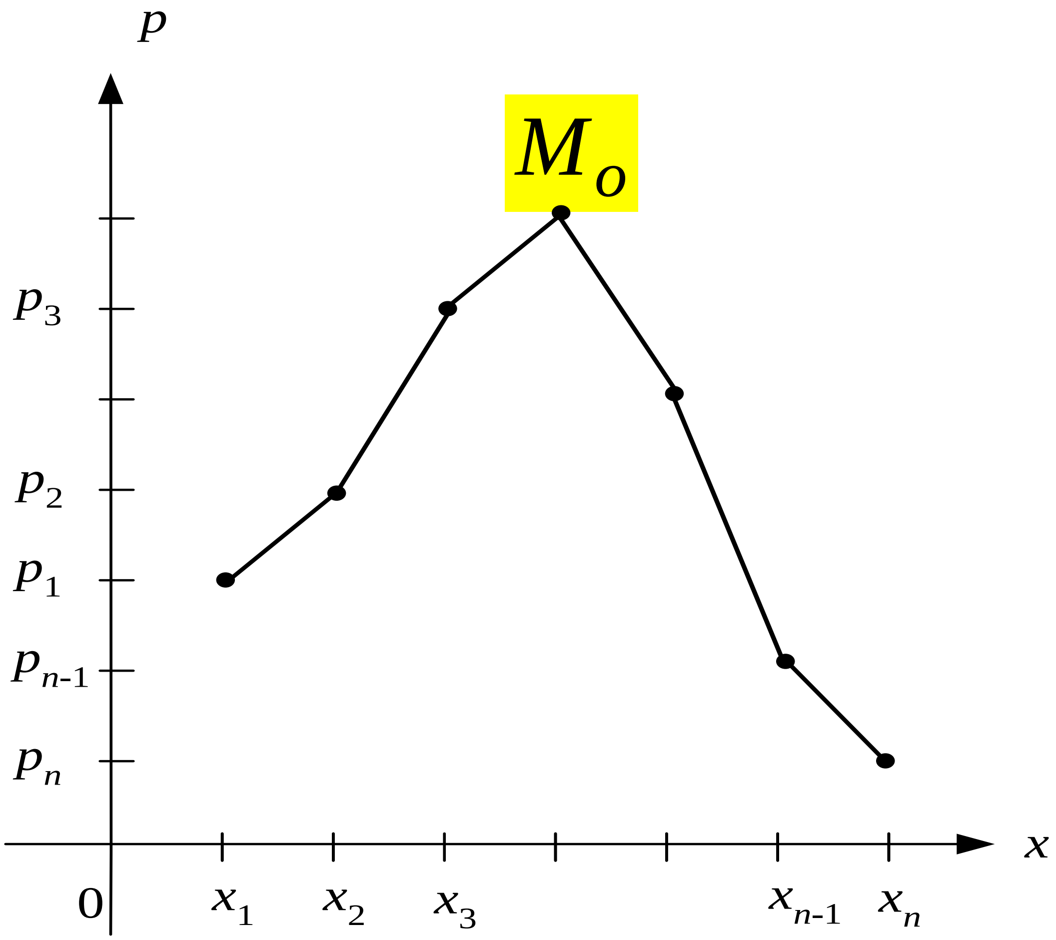
Coefficient of variation gives a possibility to compare a level of dispersion of values which have different character.

5. Mode

A ***mode*** M_o of a discrete random variable is the value of X that is most likely to occur. Consequently, the **mode** is equal to the value of x at which the probability distribution function reaches a maximum

$$P(M_o)$$

5. Mode



6. Median

A ***median*** m_e of a discrete random variable is the “middle” value. It is the value of ***X*** for which $P(X \le x)$ is greater than or equal to 0.5 and $P(X \ge x)$ is greater than or equal to 0.5.

7. Initial moments

The ***k-th moment*** satisfies the relation

$$\nu_k = M\left(X^k\right)$$

$$\nu_k = \sum_{i=1}^{n} x_i^k \cdot p_i$$

7. Initial moments

$$\nu_1 = M(X)$$

$$\nu_2 = M\left(X^2\right)$$

8. Central moments

The ***k-th central moment*** satisfies the relation:

$$\mu_k = \sum_{i=1}^{n} \left(x_i - M(X)\right)^k \cdot p_i$$

8. Central moments

$$\mu_k = M\left(\left(X - M(X)\right)^k\right)$$

$$\mu_1 = 0$$

$$\mu_2 = D(X)$$

EXAMPLE (DRV)

Distribution of the discrete random variable ***X*** is given:

x_i	1	2	3
p_i	0,4	0,3	0,3

EXAMPLE (DRV)

Distribution of the discrete random variable ***X*** is given:

x_i	1	2	3
p_i	0,4	0,3	0,3

Find all numerical characteristics.

EXAMPLE (DRV)

x_i	1	2	3
p_i	0,4	0,3	0,3

1. The mathematical expectation:

$$M(X) = x_1 p_1 + x_2 p_2 + x_3 p_3 = 1 \cdot 0{,}4 + 2 \cdot 0{,}3 + 3 \cdot 0{,}3 = 1{,}9$$

EXAMPLE (DRV)

x_i	1	2	3
p_i	0,4	0,3	0,3

2. The variance:

$$D(X) = (x_1 - M(X))^2 p_1 + (x_2 - M(X))^2 p_2 + (x_3 - M(X))^2 p_3 =$$

EXAMPLE (DRV)

x_i	1	2	3
p_i	0,4	0,3	0,3

2. The variance:

$$D(X) = (x_1 - M(X))^2 p_1 + (x_2 - M(X))^2 p_2 + (x_3 - M(X))^2 p_3 =$$

$$= (1 - 1{,}9)^2 \cdot 0{,}4 + (2 - 1{,}9)^2 \cdot 0{,}3 + (3 - 1{,}9)^2 \cdot 0{,}3 \approx 0{,}69$$

EXAMPLE (DRV)

x_i	1	2	3
p_i	0,4	0,3	0,3

2. The variance:

$$D(X) = x_1^2 p_1 + x_2^2 p_2 + x_3^2 p_3 - (M(X))^2 = 1 \cdot 0,4 + 4 \cdot 0,3 + 9 \cdot 0,3 - 1,9^2 = 0,69$$

EXAMPLE (DRV)

x_i	1	2	3
p_i	0,4	0,3	0,3

3. The root-mean-square deviation:

$$\sigma(X) = \sqrt{D(X)} = \sqrt{0{,}69} \approx 0{,}84$$

EXAMPLE (DRV)

x_i	1	2	3
p_i	0,4	0,3	0,3

4. The mode:

$$M_o = 1 \quad \left(max\ p_i = 0{,}4\right)$$

EXAMPLE (DRV)

x_i	1	2	3
p_i	0,4	0,3	0,3

5. The median:

1 1 1 1 2 2 2 3 3 3

$$m_e = \frac{2+2}{2} = 2$$

EXAMPLE (DRV)

x_i	1	2	3
p_i	0,4	0,3	0,3

6. The coefficient of variation:

$$v(X) = \frac{\sigma(X)}{M(X)} \cdot 100\,\%$$

$$v(X) = \frac{0{,}84}{1{,}9} \cdot 100\,\% = 44\,\%$$

EXAMPLE (DRV)

x_i	1	2	3
p_i	0,4	0,3	0,3

7. The initial moment:

$$\nu_1 = M(X) \qquad \nu_2 = M\left(X^2\right)$$

$$\nu_1 = 1{,}9 \qquad \nu_2 = 4{,}3$$

EXAMPLE (DRV)

x_i	1	2	3
p_i	0,4	0,3	0,3

8. The central moment:

$$\mu_1 = 0 \qquad \mu_2 = D(X)$$

$$\mu_1 = 0 \qquad \mu_2 = 0{,}69$$

Basic distribution laws of discrete random distributions and their numerical characteristics

Binomial distribution law

A random variable X has the ***binomial distribution*** with parameters (n, p) if

$$P_n(x = k) = C_n^k p^k q^{n-k}$$

$k = 0, 1, \ldots, n$

where

$0 < p < 1$ $\qquad q = 1 - p$ $\qquad n \geq 1$

Binomial distribution law

$$P_n(x=k)=C_n^k p^k q^{n-k}$$

The numerical characteristics are given by the formulas:

$$M(X)=np \qquad D(X)=npq$$

$$\sigma(X)=\sqrt{npq}$$

Binomial distribution law

x_i	0	1	…	$n-1$	n
p_i	q^n	$C_n^1 pq^{n-1}$	…	$C_n^{n-1} p^{n-1} q$	p^n

Example. The probability of passing an exam excellently for each of three students (a 5-point system) equals 0.4. Make up a distribution law of a number of excellent marks which are got by the students at the exam. Find a mathematical expectation, a variance and a root-mean square deviation of a discrete random variable.

Solution. Let a discrete random variable be a number of students with the mark "5" (a 5-point system). It has such possible values:

$x_1 = 0$

(no student passed the exam with the mark "5");

$x_2 = 1$

(one student passed the exam with the mark "5");

$x_3 = 2$

(two students passed the exam with the mark "5");

$x_4 = 3$

(three students passed the exam with the mark "5").

$$x_1 = 0$$

$$P_3(0) = C_3^0 p^0 q^{3-0} = 1 \cdot 1 \cdot q^3 = 0.6^3 = 0.216$$

$$x_2 = 1$$

$$P_3(1) = C_3^1 p^1 q^{3-1} = 3 \cdot p \cdot q^2 = 3 \cdot 0.4 \cdot 0.6^2 = 0.432$$

$$x_3 = 2$$

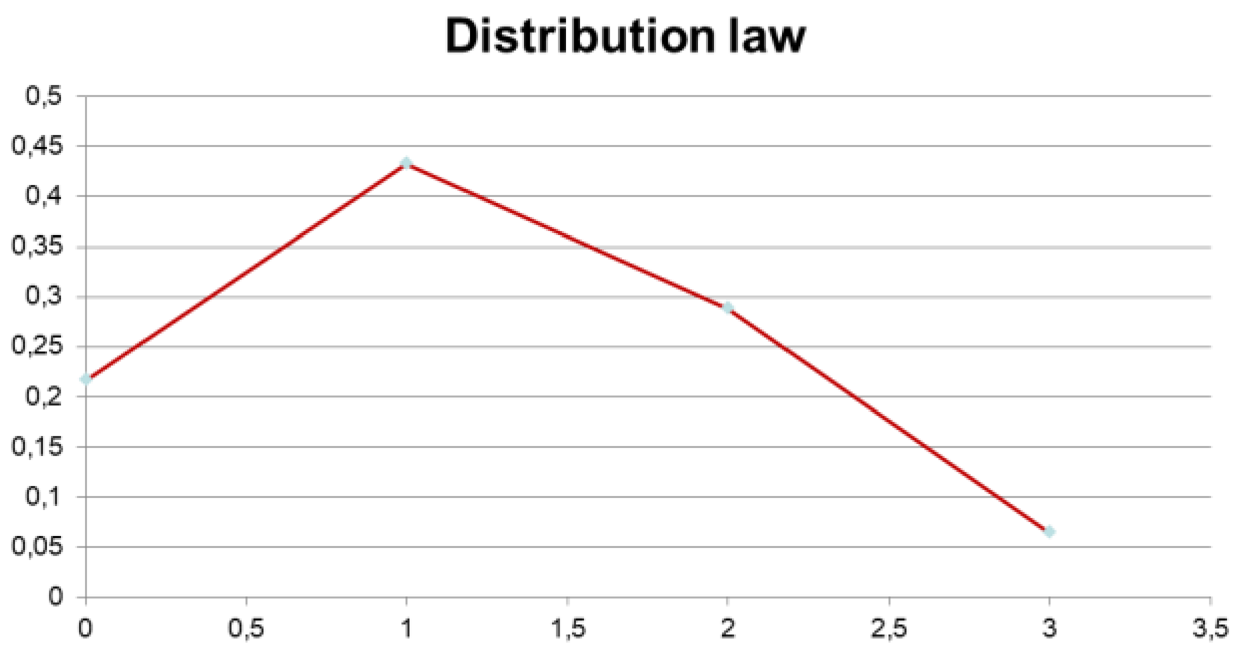
$$P_3(2) = C_3^2 p^2 q^{3-2} = 3 \cdot p^2 \cdot q = 3 \cdot 0.4^2 \cdot 0.6 = 0.288$$

$$x_4 = 3$$

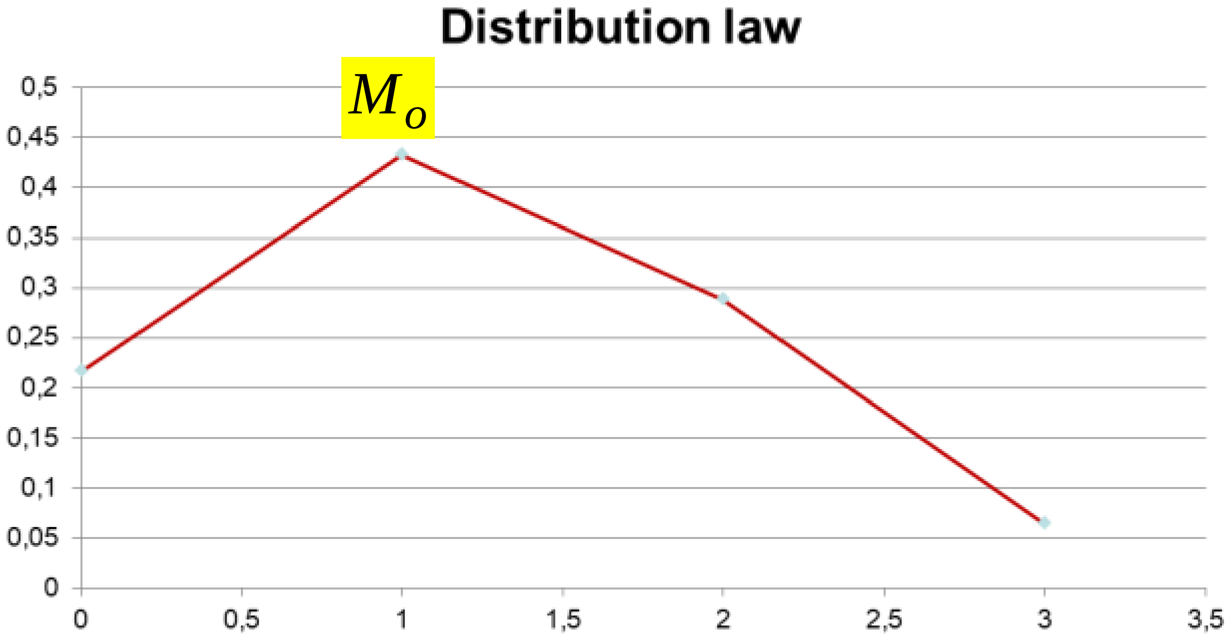
$$P_3(3) = C_3^3 p^3 q^{3-3} = 1 \cdot p^3 \cdot q^0 = 1 \cdot 0.4^3 \cdot 1 = 0.064$$

x_i	0	1	2	3
p_i	0.216	0.432	0.288	0.064

x_i	0	1	2	3
p_i	0.216	0.432	0.288	0.064



x_i	0	1	2	3
p_i	0.216	0.432	0.288	0.064



Numerical characteristics

$n = 3$ $p = 0.4$ $q = 1 - p = 0.6$

$$M(X) = np = 3 \cdot 0.4 = 1.2$$

$$D(X) = npq = 3 \cdot 0.4 \cdot 0.6 = 0.72$$

$$\sigma(X) = \sqrt{npq} = \sqrt{0.72} \approx 0.85$$

Poisson distribution law

A random variable X has the ***Poisson distribution*** with parameters $\lambda\ (\lambda > 0)$ if

$$P_n(x=k)=\frac{\lambda^k e^{-\lambda}}{k!}$$

where $0 < p < 1$

$q = 1 - p$ $\qquad k = 0, 1, 2, \ldots$

$\lambda = np$ $\qquad k! = 1 \cdot 2 \cdot \ldots \cdot k$

Poisson distribution law

$$P_n(x=k)=\frac{\lambda^k e^{-\lambda}}{k!}$$

$$(p \to 0)$$

$$q = 1 - p \to 1$$

The numerical characteristics are given by the formulas:

$$M(X)=np=\lambda \qquad D(X)=npq \approx \lambda$$

$$\sigma(X)=\sqrt{npq}=\sqrt{\lambda}$$

Poisson distribution law

x_i	0	1	2	...	n
p_i	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2 e^{-\lambda}}{2!}$	...	$\frac{\lambda^n e^{-\lambda}}{n!}$

Example. The probability of finding a mistake on a book page is equal to 0.004. 500 pages are checked. Make up a distribution law of a number of finding a mistake on a book page. Find a mathematical expectation, a variance and a root-mean square deviation of a discrete random variable.

Example. The probability of finding a mistake on a book page is equal to 0.004. 500 pages are checked. Make up a distribution law of a number of finding a mistake on a book page. Find a mathematical expectation, a variance and a root-mean square deviation of a discrete random variable.

Solution. Let ***X*** be a number of finding a mistake on a book page, then the possible values of are 0, 1, 2, 3, …, 500. Here

$$p = 0.004$$

$$\lambda = 500 \cdot 0.004 = 2$$

$$n = 500$$

$$P_{500}(x=0)=\frac{2^0\cdot e^{-2}}{0!}\approx 0.13534$$

$$P_{500}(x=1)=\frac{2^1\cdot e^{-2}}{1!}\approx 0.27067$$

$$P_{500}(x=2)=\frac{2^2\cdot e^{-2}}{2!}\approx 0.27067$$

$$P_{500}(x=3)=\frac{2^3\cdot e^{-2}}{3!}\approx 0.18045$$

$$P_{500}(x=4)=\frac{2^4\cdot e^{-2}}{4!}\approx 0.09022$$

$$P_{500}(x=5)=\frac{2^5\cdot e^{-2}}{5!}\approx 0.03609$$

$$P_{500}(x=6)=\frac{2^6\cdot e^{-2}}{6!}\approx 0.01203$$

$$P_{500}(x=7)=\frac{2^7\cdot e^{-2}}{7!}\approx 0.00344$$

$$P_{500}(x=8)=\frac{2^8\cdot e^{-2}}{8!}\approx 0.00086$$

$$P_{500}(x=9)=\frac{2^9\cdot e^{-2}}{9!}\approx 0.00019$$

$$P_{500}(x=10)=\frac{2^{10}\cdot e^{-2}}{10!}\approx 0.00004$$

.

At $k\geq 11$ we have that $P_{500}(x\geq 11)\approx 0$

x_i	**0**	**1**	**2**	**3**	**4**	**5**
p_i	**0.13534**	**0.27067**	**0.27067**	**0.18045**	**0.09022**	**0.03609**
x_i	**6**	**7**	**8**	**9**	**10**	**…**
p_i	**0.01203**	**0.00344**	**0.00086**	**0.00019**	**0.00004**	**…**

Let's find the numerical characteristics by the formulas:

$$M(X) = \lambda = 2$$

$$D(X) \approx \lambda = 2$$

$$\sigma(X) = \sqrt{\lambda} = \sqrt{2} \approx 1.41421$$

INDEPENDENT WORK

Geometric distribution law

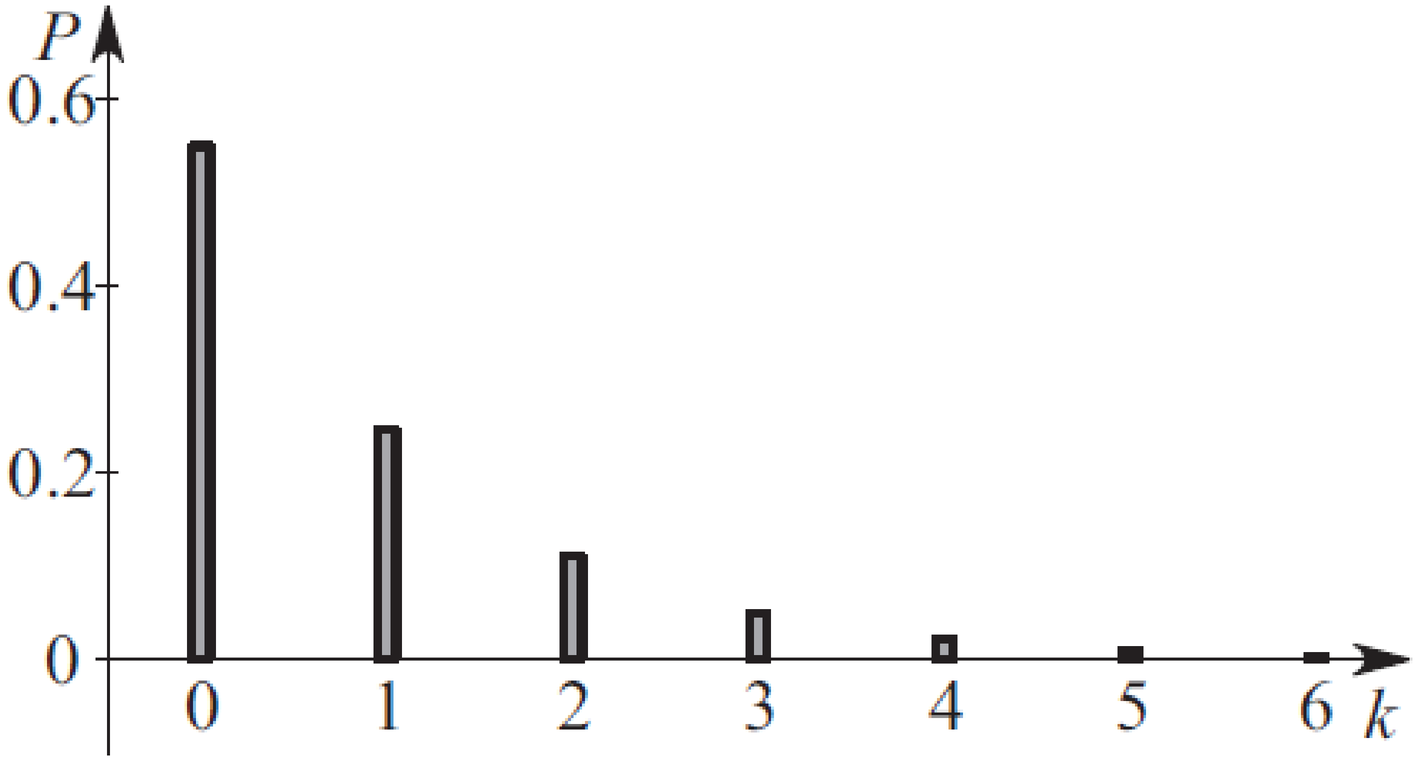
A random variable ***X*** has a *geometric distribution* with parameters (n, p) if

$$P_n(x = k) = pq^k$$

$$0 < p < 1 \qquad k = 0, 1, 2, \ldots$$

$$q = 1 - p \qquad n \geq 1$$

Geometric distribution law



$p = 0.55 \qquad n = 6$

Geometric distribution law

$$P_n(x=k)=pq^k$$

The numerical characteristics can be calculated by the formulas

$$M(X)=\frac{1-p}{p}=\frac{q}{p} \qquad D(X)=\frac{1-p}{p^2}=\frac{q}{p^2}$$

$$\sigma(X)=\sqrt{\frac{q}{p^2}}$$

Example

A doctor is seeking an anti-depressant for a newly diagnosed patient. Suppose that, of the available anti-depressant drugs, the probability that any particular drug will be effective for a particular patient is p=0.6. What is the probability that the first drug found to be effective for this patient is the first drug tried, the second drug tried, and so on? What is the expected number of drugs that will be tried to find one that is effective?

Example

Consider the anti-depressant example above. The probability that any given drug is effective (success) is $p=0.6$.

The probability that a drug will not be effective (fail) is $q = 1 – p = 1 – 0.6 = 0.4$.

Here are probabilities of some possible outcomes.

Example

(a) The first drug works. There are zero failures before the first success. Y = 0 failures. The probability p (zero failures before first success) is simply the probability that the first drug works.

$$P\ (Y = 0) = q^0 p\ = 0.4^0 * 0.6 = 1 * 0.6 = 0.6.$$

Example

(b) The first drug fails, but the second drug works. There is one failure before the first success. Y=1 failure. The probability for this sequence of events is p(first drug fails) * p(second drug is success) which is given by

$$\mathrm{P}\ (Y = 1) = q^1\, p\ = 0.4^1 * 0.6 = 0.4 * 0.6 = 0.24.$$

Example

(c) The first drug fails, the second drug fails, but the third drug works. There are two failures before the first success. Y= 2 failures. The probability for this sequence of events is p(first drug fails) * p(second drug fails) *p(third drug is success)

$$\mathrm{P}\ (Y = 2) = q^2\,p, = 0.4^2 * 0.6 = 0.096.$$

Example

The general formula to calculate the probability of k failures before the first success, where the probability of success is p and the probability of failure is q = 1 - p, is

$$\mathrm{P}\ (Y = k) = q^k\, p.$$

for k = 0, 1, 2, 3, ...

TASK

A representative from the National Football League's Marketing Division randomly selects people on a random street in Kansas City, Kansas until he finds a person who attended the last home football game. Let p, the probability that he succeeds in finding such a person, equal 0.20. And, let X denote the number of people he selects until he finds his first success.

TASK

A representative from the National Football League's Marketing Division randomly selects people on a random street in Kansas City, Kansas until he finds a person who attended the last home football game. Let p, the probability that he succeeds in finding such a person, equal 0.20. And, let X denote the number of people he selects until he finds his first success.

(a) What is the probability that the marketing representative must select 4 people before he finds one who attended the last home football game?

(b) What is the probability that the marketing representative must select more than 6 people before he finds one who attended the last home football game?

(c) How many people should we expect (that is, what is the average number) the marketing representative needs to select before he finds one who attended the last home football game? And, while we're at it, what is the variance?

TASK

(a) What is the probability that the marketing representative must select 4 people before he finds one who attended the last home football game?

Solution. To find the desired probability, we need to find $P(X = 4)$, which can be determined readily using the p.m.f. of a geometric random variable with $p = 0.20$, $1-p = 0.80$, and $x = 4$:

$$P(X = 4) = 0.80^3 \times 0.20 = 0.1024$$

There is about a 10% chance that the marketing representative would have to select 4 people before he would find one who attended the last home football game.

TASK

(b) What is the probability that the marketing representative must select more than 6 people before he finds one who attended the last home football game?

Solution. To find the desired probability, we need to find $P(X > 6) = 1 - P(X \leq 6)$, which can be determined readily using the c.d.f. of a geometric random variable with $1-p = 0.80$, and $x = 6$:

$$P(X > 6) = 1 - P(X \leq 6) = 1 - [1 - 0.8^6] = 0.8^6 = 0.262$$

There is about a 26% chance that the marketing representative would have to select more than 6 people before he would find one who attended the last home football game.

TASK

(c) How many people should we expect (that is, what is the average number) the marketing representative needs to select before he finds one who attended the last home football game? And, while we're at it, what is the variance?

Solution. The average number is:

$$\mu = E(X) = \frac{1}{p} = \frac{1}{0.20} = 5$$

That is, we should expect the marketing representative to have to select 5 people before he finds one who attended the last football game. Of course, on any given try, it may take 1 person or it may take 10, but 5 is the average number. The variance is 20, as determined by:

$$\sigma^2 = Var(X) = \frac{1-p}{p^2} = \frac{0.80}{0.20^2} = 20$$

Example

A doctor is seeking an anti-depressant for a newly diagnosed patient. Suppose that, of the available anti-depressant drugs, the probability that any particular drug will be effective for a particular patient is p=0.6. What is the probability that the first drug found to be effective for this patient is the first drug tried, the second drug tried, and so on? What is the expected number of drugs that will be tried to find one that is effective?

Geometric distribution law

$$P_n(x = k) = pq^k$$

The geometric distribution describes a random variable ***X*** equal to the number of failures before the first success in a sequence of Bernoulli trials with probability ***p*** of success in each trial.

Hypergeometric distribution law

A random variable has the *hypergeometric distribution* with parameters (*N, p, n*) if

$$P_n\left(x=k\right)=\frac{C_{Np}^{k}\,C_{Nq}^{n-k}}{C_N^n}$$

$$0<p<1 \qquad k=0,1,\ldots,n$$

$$q=1-p \qquad 0\le n\le N$$

$$N>0$$

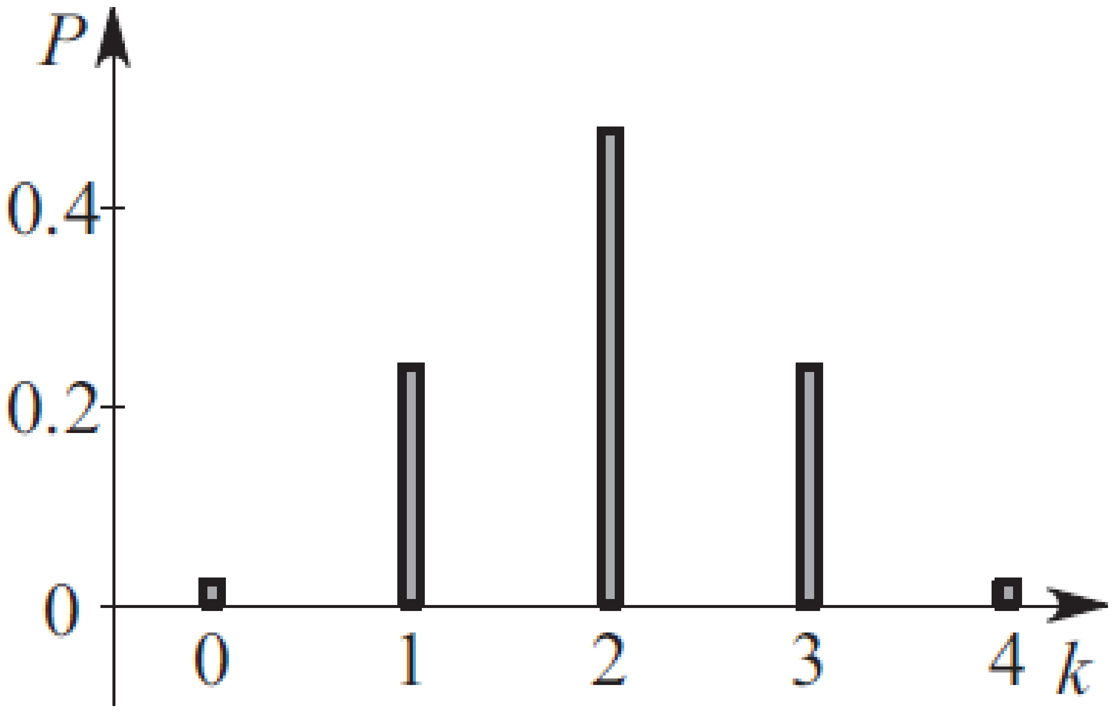
Hypergeometric distribution law

If $n << N$ (in practice, $n < 0{,}1N$), then

$$\frac{C_{Np}^{k} C_{Nq}^{n-k}}{C_{N}^{n}} \approx C_n^k p^k q^{n-k}$$

i.e., the hypergeometric distribution tends to the binomial distribution.

Hypergeometric distribution law



$p = 0.5$ $N = 10$ $n = 4$

Hypergeometric distribution law

$$P_n(x=k)=\frac{C_{Np}^{k}\,C_{Nq}^{n-k}}{C_N^n}$$

The numerical characteristics are given by the formulas

$$M(X)=np$$

$$D(X)=\frac{N-n}{N-1}npq$$

$$\sigma(X)=\sqrt{\frac{N-n}{N-1}npq}$$

Hypergeometric distribution law

$$P_n\left(x=k\right)=\frac{C_{Np}^{k}\,C_{Nq}^{n-k}}{C_{N}^{n}}$$

A typical scheme in which the hypergeometric distribution arises is as follows: *n* elements are randomly drawn without replacement from a population of *N* elements containing exactly *Np* elements of type *I* and *Nq* elements of type *II*. The number of elements of type *I* in the sample is described by the hypergeometric distribution.

Example

A crate contains 50 light bulbs of which 5 are defective and 45 are not. A Quality Control Inspector randomly samples 4 bulbs without replacement. Let X = the number of defective bulbs selected. Make up the distribution law of the discrete random variable X.

TASK 1

Suppose we randomly select 5 cards without replacement from an ordinary deck of playing cards. What is the probability of getting exactly 2 red cards (i.e., hearts or diamonds)?

TASK 1

Suppose we randomly select 5 cards without replacement from an ordinary deck of playing cards. What is the probability of getting exactly 2 red cards (i.e., hearts or diamonds)?

Solution. This is a hypergeometric experiment in which we know the following:

N = 52; since there are 52 cards in a deck.

k = 26; since there are 26 red cards in a deck.

n = 5; since we randomly select 5 cards from the deck.

x = 2; since 2 of the cards we select are red.

TASK 2

A box of 20 marbles contains 15 blue and 5 red. You need to draw a lot of 10 marbles at random. Find the probability of drawing 6 blue marbles in the lot drawn.

TASK 3

Automobiles arrive in a dealership in lots of 10. Five out of each 10 are inspected. For one lot, it is know that 2 out of 10 do not meet prescribed safety standards.

What is probability that at least 1 out of the 5 tested from that lot will be found not meeting safety standards?

TASK 4

Suppose we select 5 cards from an ordinary deck of playing cards. What is the probability of obtaining 2 or fewer hearts?

Solution: This is a hypergeometric experiment in which we know the following:

N = 52; since there are 52 cards in a deck.

k = 13; since there are 13 hearts in a deck.

n = 5; since we randomly select 5 cards from the deck.

x = 0 to 2; since our selection includes 0, 1, or 2 hearts.

Thank You for your time and attention! Hope this presentation was educational and helpful to you ☺