Національний технічний університет
ДНІПРОВСЬКА ПОЛІТЕХНІКА
1899

Пістунов І.М.

Probability Theory And Base of Mathematical Statistics

Probability Theory And Base of Mathematical Statistics : навч. наоч. посіб. / І.М. Пістунов. – Дніпро : НТУ «ДП», 2022. – 41 с.

Probability theory arose from the practical needs of the need to predict random events. We will call an event or coincidence a phenomenon that may or may not occur.
For economists, probability theory is one of the basic fundamental disciplines because financial activity in a society with a free economy is completely subject to the laws of chance, uncertainty. Already in the XVII century, the theory of probabilities was used in the activities of insurance companies, and today it is used to calculate risky financial transactions, planning of banking activities, macroeconomic planning. The tasks of the economic direction can also include demographic, agricultural, production calculations.
.
Designed for students of higher education specialties "Economics".

Рецензенти:
Кострицька С.І., завідувач кафедри іноземних мов НТУ “ДП”, проф.
Алексеєв М.О., декан факультету інформаційних технологій, проф.

CONTENT

PREFACE

Probability theory arose from the practical needs of the need to predict random events. We will call an event or coincidence a phenomenon that may or may not occur. And although this theory originates from the need to predict the results of gambling (the word "gambling", from the French "le hasard" means "case"), its further development has led to significant achievements in the theory of measurement, queuing, reliability, etc.
For economists, probability theory is one of the basic fundamental disciplines because financial activity in a society with a free economy is completely subject to the laws of chance, uncertainty.
Already in the XVII century, the theory of probabilities was used in the activities of insurance companies, and today it is used to calculate risky financial transactions, planning of banking activities, macroeconomic planning. The tasks of the economic direction can also include demographic, agricultural, production calculations.

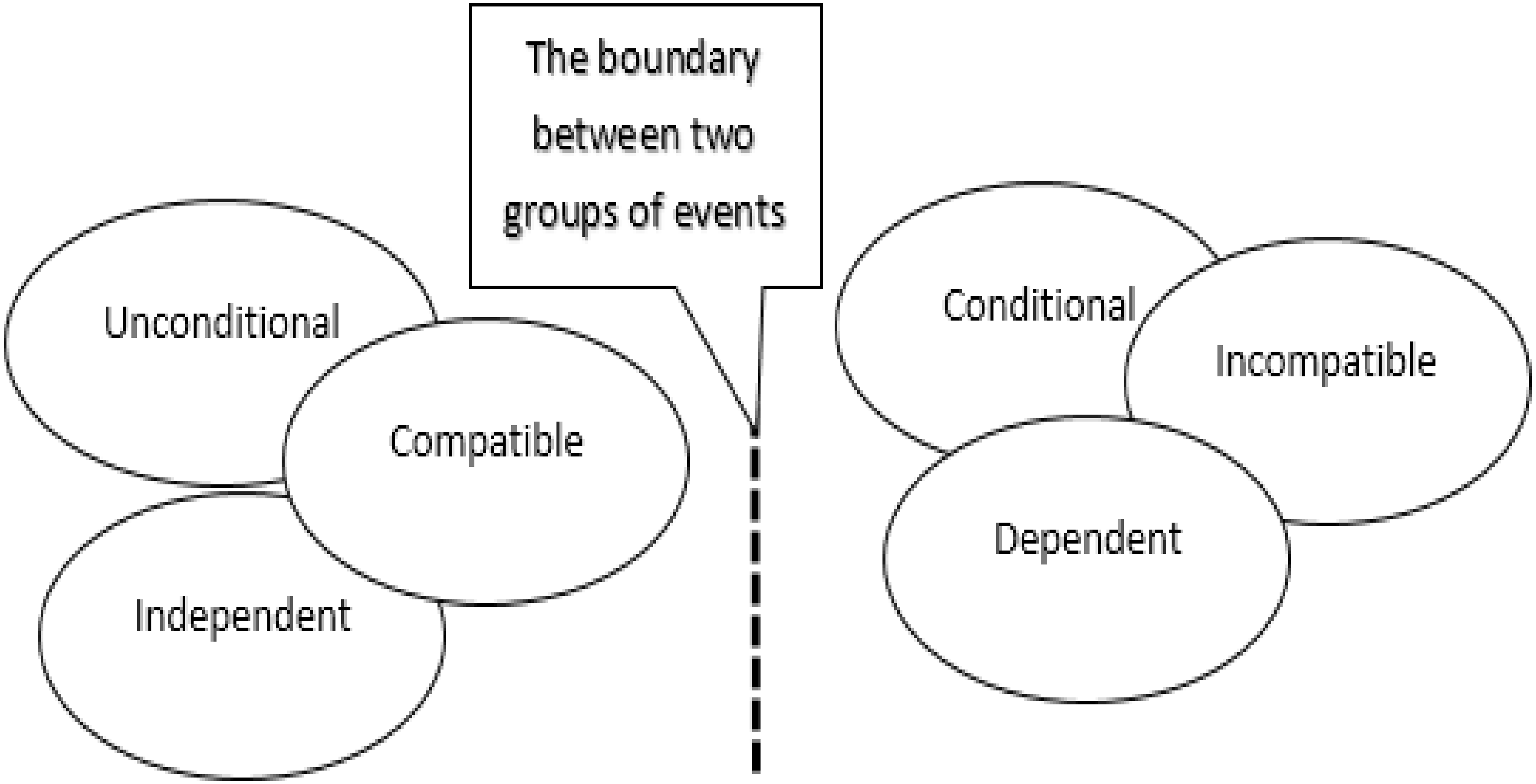
1 PROBABILITIES

$$p = \frac{a}{b}$$

$$a = p \cdot b$$

$$b = a/p$$

$$p = 100\frac{a}{b}\%$$



Impossible and reliable events

$$0 \leq P(A) \leq 1$$

RULES FOR COMPILATION OF LIABILITY

events A_1 ; A_2 ;…… A_n

their probabilities $P(A_1)$; $P(A_2)$;… $P(A_n)$

$P(A_1$ ***або*** A_2 …. ***або*** $A_n) =$

$$=P(A_1)+P(A_2)+\ldots+P(A_n) = \sum_{i=1}^{n} p(A_i)$$

only for incompatible events.

A complete system of events

$$P(A_1)+P(A_2)+\ldots+P(A_n) = 1$$

A complete system of events forms only incompatible events.

CONCEPT OF CONDITIONS OF CONDITIONS

$P_B(A)$ – the probability of the occurrence of event A, provided that event B occurred.

Probability multiplication rule

or conditional events

$$P(A \text{ i } B) = P(B)\, P_B(A) = P(A)\, P_A(B)$$

for independent events

$$P(A \text{ i } B) == P(A)P(B)$$

$$p(A_1 \text{ i } \; A_2 \text{ i } \ldots \text{ i } A_n) = P(A_1)P(A_2)\ldots P(A_n) =$$

$$= \prod_{i=1}^{n} p(A_i)$$

Probability of occurrence of at least one of the independent events

$p(A_1$ or A_2 or … or $A_n) =$

$$= 1 - \prod_{i=1}^{n} [1 - p(A_i)]$$

When the probabilities are the same for all events

$p(A_1$ or A_2 or … or $A_n) =$

$$= 1 - (1 - p)^n$$

Generalization of the rules of compilation and multiplication of probabilities

P(A or B) = P(A) + P(B) - P(A i B).

If events A and B are incompatible,

then ***P(A and B)*= 0.**

If events A and B are mutually independent, then ***P(A and B) = P(A)P(B)***.

Because ***0 ≤ P(A and B),*** then

P(A or B) ≤ P(A) + P(B)

Formulas of combinatorics

- Permutations A^n_m $= n(n-1)(n-2)....(n-m+1) =$

$$= \frac{n!}{(n-m)!}$$

- Placing $P_m = m!$

- Combination C^n_m $= \frac{A_n^m}{P_m} = \frac{n(n-1)(n-2)....(n-m+1)}{m!} =$

$$= \frac{n!}{m!(n-m)!}$$

Let, in a urn *a* white and *b* black balls; from the urn of catch-up take out *k* bullets. Find the probability that among them they will be *l* white, and, therefore, *k*-*l* black ($l \le a$, $k- l \le b$).

$$P(A)=\frac{C_a^l \cdot C_b^{k-l}}{C_{a+b}^k}$$

Формула повної ймовірності

Events A_1 ; A_2 ;...... A_n

Their probabilities $P(A_1)$; $P(A_2)$;... $P(A_n)$

Probability of the result *К*

$P_{A1}(K)$;$P_{A2}(K)$; … ; $P_{An}(K)$

for any possible result K of this operation, the probability of its onset will be

$$P(K)=\sum_{i=1}^{n}P(A_i)P_{A_i}(K)$$

Bayes formula

Let events A_1, A_2,.., A_n represent a complete system of events. If then K means the arbitrary result of this operation, then the probability that this arbitrary result was due to *q* operation($1 \le q \le n$)

$$P_K(A_q) = \frac{P(A_q)P_{A_q}(K)}{\sum_{i=1}^{n} P(A_i)P_{A_i}(K)}$$

Bernoulli's Formula

$$P_n(k) = C_n^k \cdot p^k \cdot (1 - p)^{n-k}$$

Clarification of the Bayes formula for multiple tests

$$P_{K_m}(A_{qs}) = \frac{P_q p_q^m (1 - p_q)^{s-m}}{\sum_{i=1}^{n} P_i p_i^m (1 - p_i)^{s-m}}$$

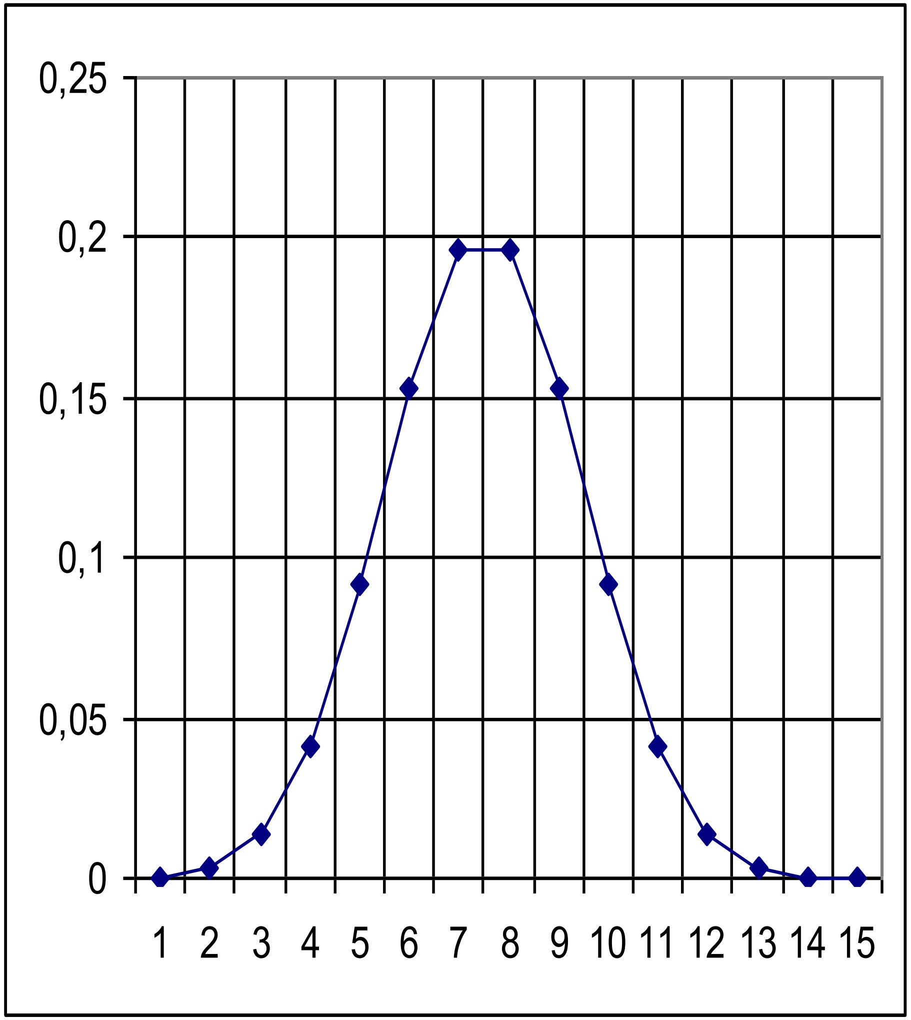
Most likely is the occurrence of events k_0

$$np - (1 - p) \leq k_0 \leq np + p$$

Moreover:

a) if the number $np - q$ is fractional, then there is one most likely number k_0;

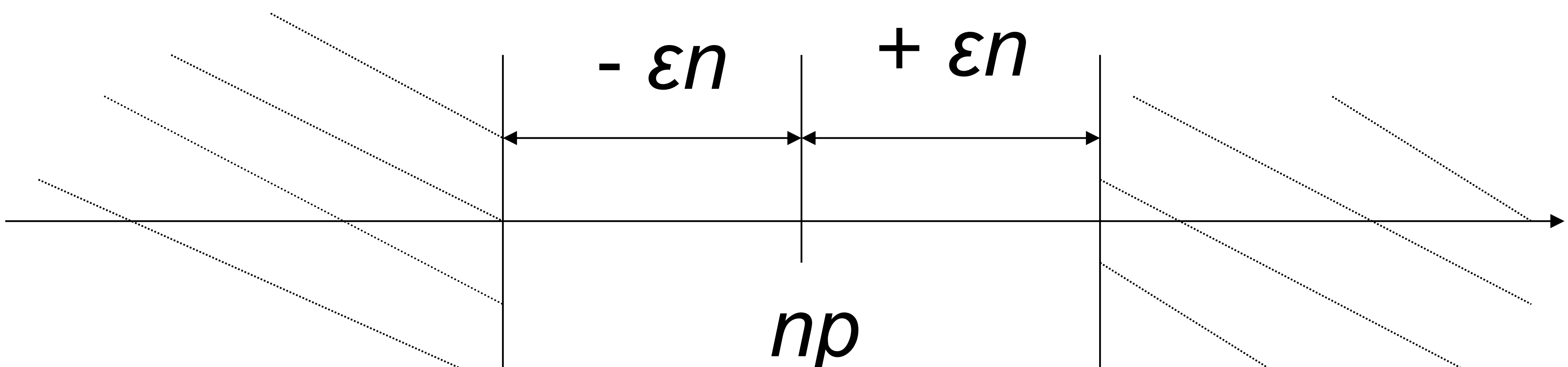
- b) If the number $np - q$ is an integer, then there are two most probable numbers, namely : k_0 and k_0+1;
- c) if the number np is an integer, then the most likely number $k_0 = np$.



Bernoulli's theorem. The first form of the law of large numbers

$$P(|k - np| > \varepsilon n) < \frac{p(1-p)}{\varepsilon^2 n}$$

$$P(|k - np| \le \varepsilon n) \ge 1 - \frac{p(1-p)}{\varepsilon^2 n}$$



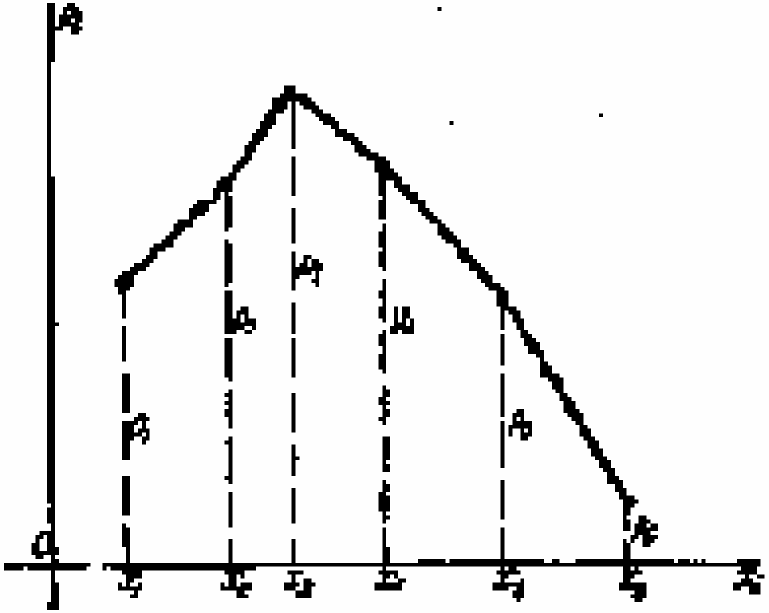
2. DISCRETE RANDOM VALUES

The law of distribution

x_1	x_2	...	x_n
p_1	p_2	...	p_n

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

Polygon distribution or distribution of probabilities



Numerical characteristics of a discrete random variable

Beginning moments

$$\alpha_S[X] = \sum_{i=1}^{n} X_i^S P_i$$

Central moments

$$\mu_S[X] = \sum_{i=1}^{n} (X_i - \alpha_1)^S P_i$$

Value of beginning and central moments

$\mu_1 = 0;$
$\mu_2 = \alpha_2 - \alpha_1^2;$
$\mu_3 = \alpha_3 - 3\alpha_1\alpha_2 + 2\alpha_1^3;$
$\mu_4 = \alpha_4 - 4\alpha_1^2\alpha_3 + 3\alpha_2\alpha_3 -$
$-4\alpha_1^4$.

Average

$\alpha_1 = m_x, M_x, M[X]$, $\bar{x}$

Dispersion

$\mu_2 = D_x, D[X], Q_x^2, q^2, \sigma_x^2$

$$D_X = \sum_{i=1}^{n} X_i^2 P_i - M^2[X] \qquad \sigma_x = \sqrt{D_x} \qquad Var_x = \frac{\sigma_x}{M_x}$$

Theorems on the properties of the mean and dispersion

$M(a + X) = a + M(X)$.

$M(a \cdot X) = a \cdot M(X)$.

$M(X_1 + X_2 + X_3 ..) = M(X_1) + M(X_2) + M(X_3) + ..$

$M(X_1 \cdot X_2 \cdot X_3 ...) = M(X_1) \cdot M(X_2) \cdot M(X_3) \cdot ...$

$D(a + X) = D(X)$.

$D(a \cdot X) = a^2 \cdot D(X)$.

$D(X_1 + X_2 + X_3 ...) = D(X_1) + D(X_2) + D(X_3) + ..$

Theorems on mean square deviation (standard)

We have random values $x_1, x_2, \ldots, x_n$

with standards $q_1, q_2, \ldots, q_n$

arithmetic mean $\xi = (x_1 + x_2 + \ldots + x_n)/n$

results of n measurements.

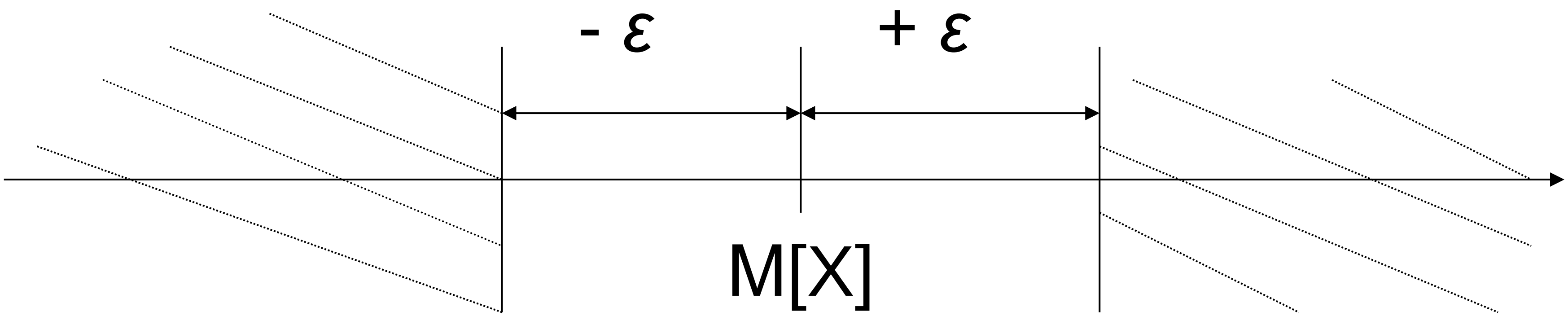
Then the standard of this average, provided that all standards are the same

$$\frac{Q}{n} = \frac{q}{\sqrt{n}}$$

Chebyshev's inequality

$$P\left(\left|X - M(X)\right| < \varepsilon\right) \geq 1 - \frac{D(X)}{\varepsilon^2}$$

$$P\left(\left|\xi - M(X)\right| \leq \varepsilon\right) \geq 1 - \frac{q^2}{\varepsilon^2 \cdot n}$$



3. CONTINUOUS RANDOM VALUES

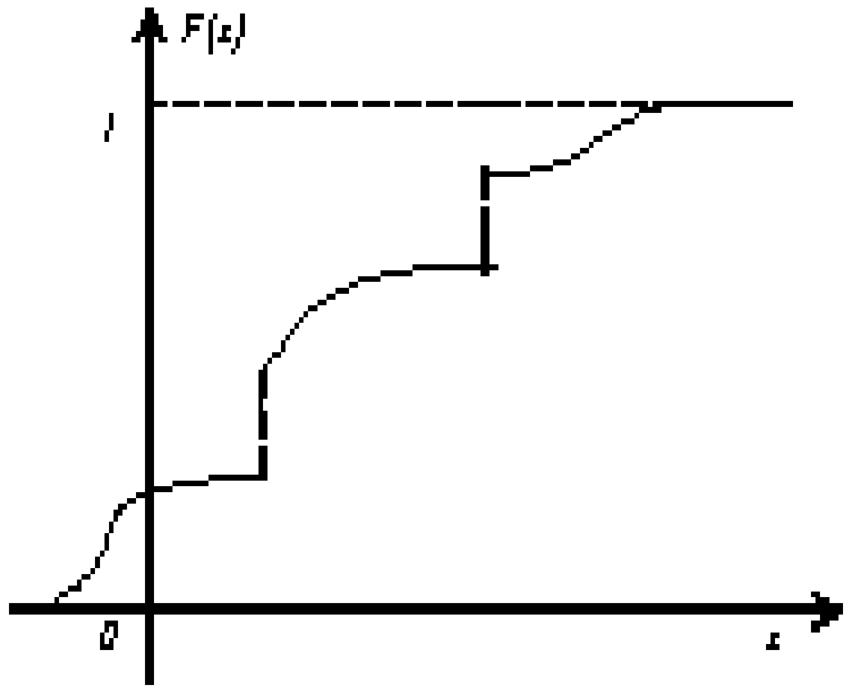
Distribution function $F(x) = P(X<x)$.

Properties:

1) $x_2 > x_1$, $F(x_2)>F(x_1)$.

2)

3) $$F(-\infty)=0$$
$$F(+\infty)=1$$



Typical distribution function graph

Distribution density

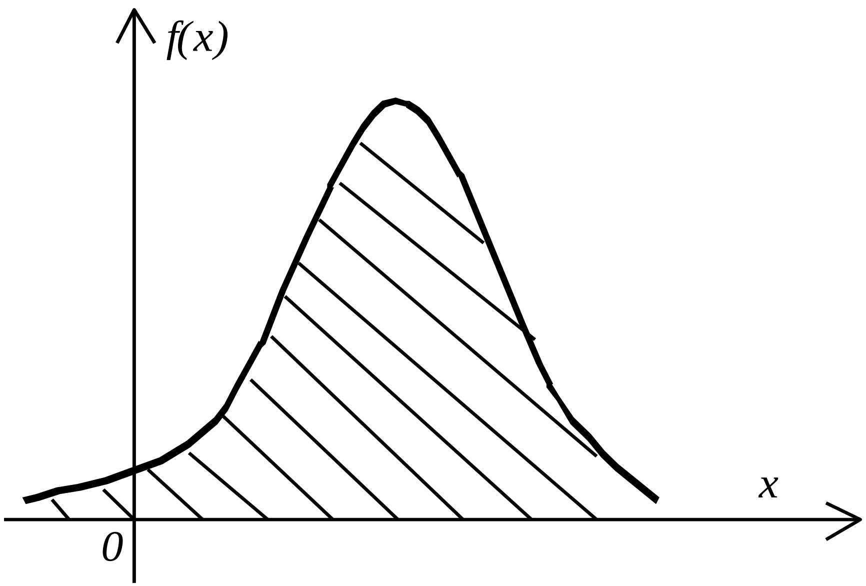
$$\frac{dF(x)}{dx} = f(x)$$

$$F(x) = \int_{-\infty}^{\alpha} f(x)dx$$

Properties

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

2) $f(x) \geq 0$



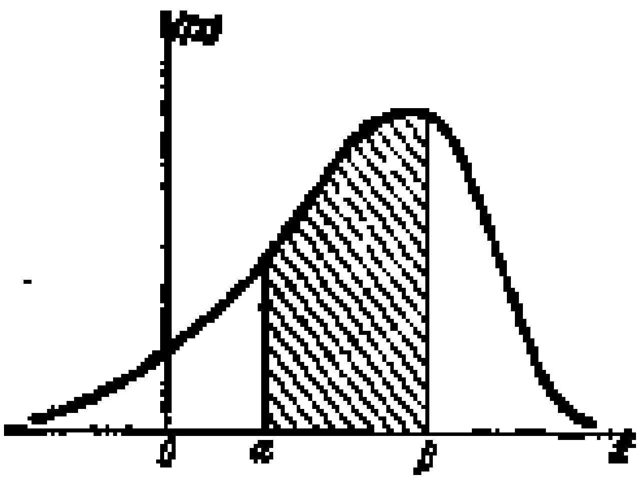
Typical form of distribution density

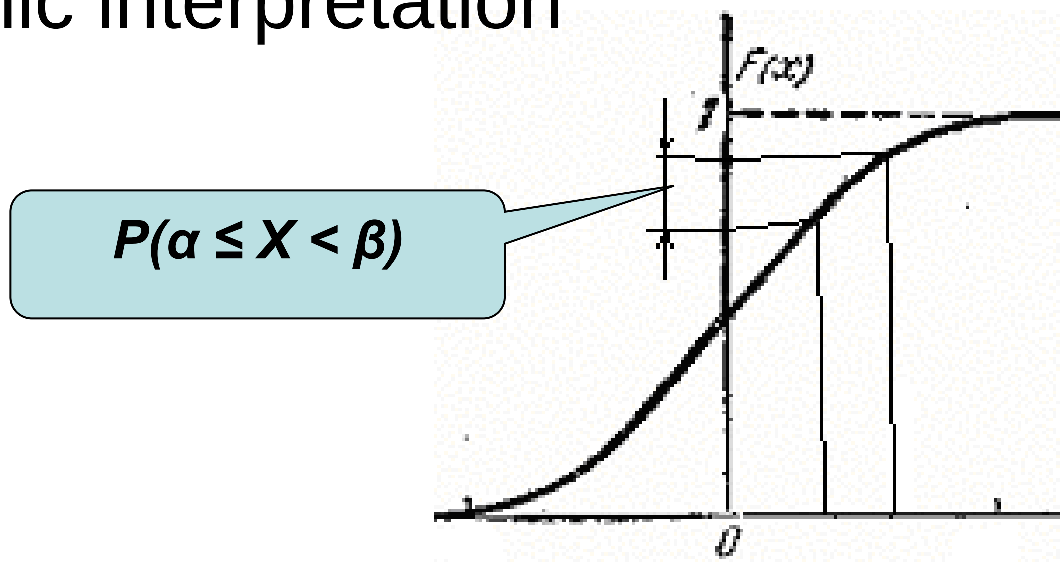
The probability of accidental access to a given site

$$P(\alpha \leq X < \beta) = F(\beta) - F(\alpha).$$

$$P(\alpha \leq X < \beta) = \int_{\alpha}^{\beta} f(x)dx$$

Graphic interpretation





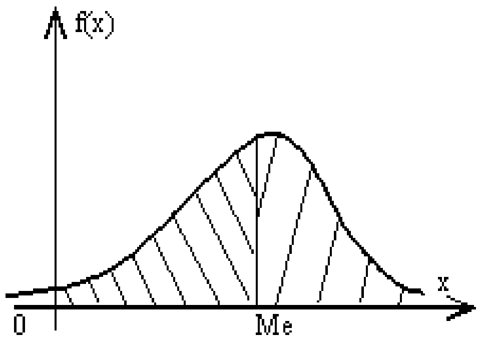
Numerical characteristics of continuous random variables

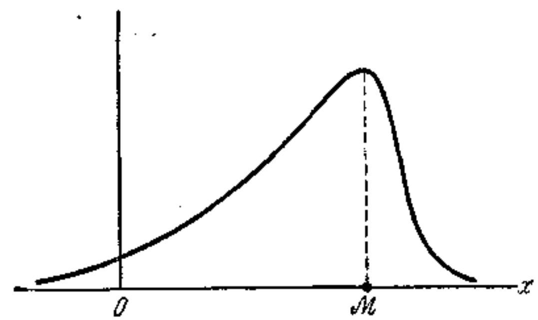
Beginning moments

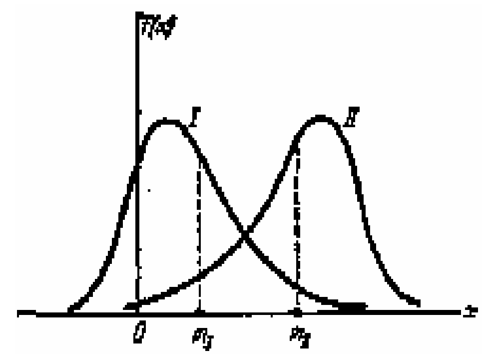
$$\alpha_S[X] = \int_{-\infty}^{\infty} x^S f(x)dx$$

Central moments

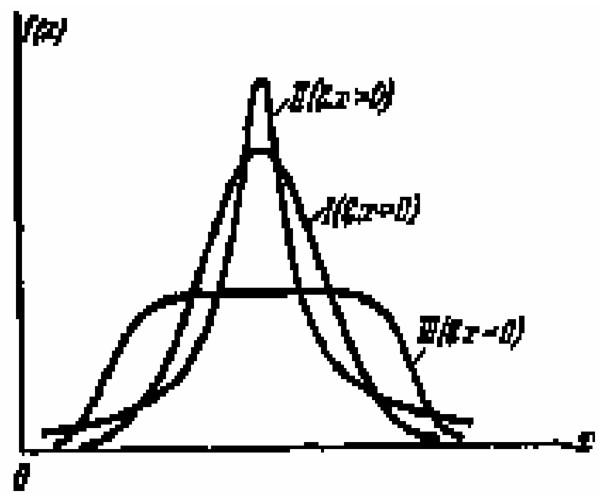
$$\mu_S[X] = \int_{-\infty}^{\infty} (x-\alpha_1)^S f(x)dx$$







$$\int_{-\infty}^{Me} f(x)dx = \int_{Me}^{+\infty} f(x)dx$$



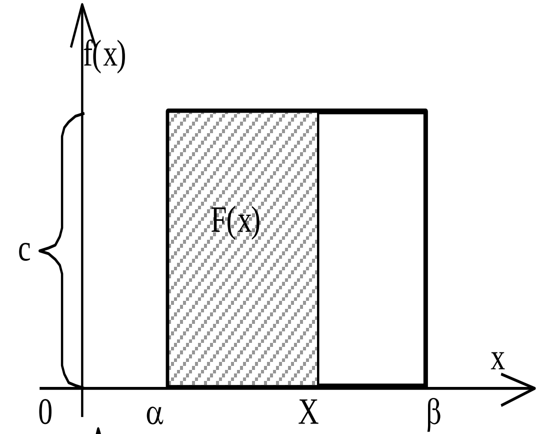
$$E_X = \mu_4 / \sigma^4 - 3$$

$S_K = \mu_3/\sigma^3$

The law of uniform density

f(x)=1/(*β*-*α*), при $\alpha \le x \le \beta$

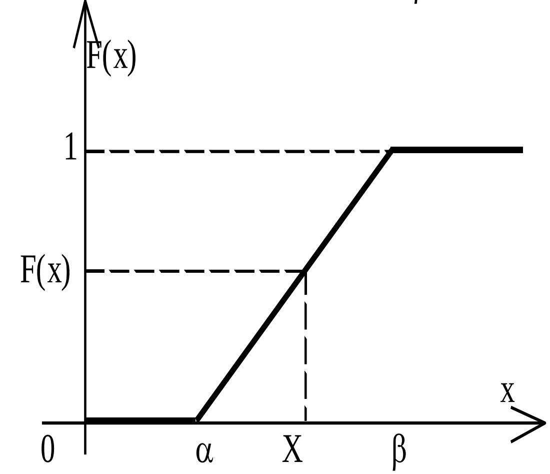
f(x)=0 при *x < α* або *x > β*



F(x) = 0, при *x < α;*

F(x) = (x-α)/(β-α), при $\alpha \le x \le \beta$

F(x) = 1 при *x > β*



$$m_x = \int_\alpha^\beta \frac{x}{\beta-\alpha} \, dx = \frac{(\alpha+\beta)}{2}$$

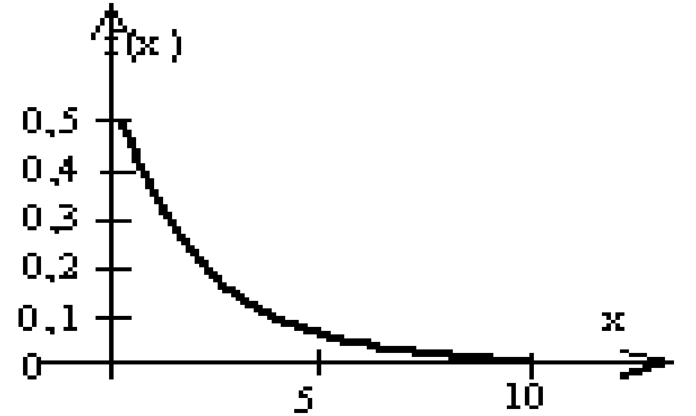
$$D_x = \alpha_2 = \frac{1}{\beta-\alpha} \int_\alpha^\beta \left(x - \frac{\alpha+\beta}{2} \right)^2 dx = \frac{(\beta-\alpha)^2}{12}$$

$$\sigma_x = \sqrt{D_x} = \frac{\beta-\alpha}{\sqrt{12}} \quad E_X = \mu_4 / \sigma^4 - 3 = -1{,}2$$

$$P(a < x < b) = \frac{b-a}{\beta-\alpha}$$

Exponential distribution law

$$f(x)=\begin{cases}0, & npu \quad x<0\\ \lambda\cdot e^{-\lambda x}, & npu \quad x\ge 0\end{cases}$$



$$F(x)=\begin{cases}0, & npu \quad x<0\\ \int_0^x \lambda\cdot e^{-\lambda x}dx=1-e^{-\lambda x}, & npu \quad x\ge 0\end{cases}$$

$$M(x)=\int_0^\infty \lambda x e^{-\lambda x}dx=\frac{1}{\lambda}$$

$P(a<x<b)=e^{-\lambda a}-e^{-\lambda b}$

$$D_x=\int_0^\infty \left(x-\frac{1}{\lambda}\right)^2 \lambda e^{-\lambda x}dx=\frac{1}{\lambda^2}$$

$M_e=-\ln 0.5/\lambda\approx 0.69/\lambda$

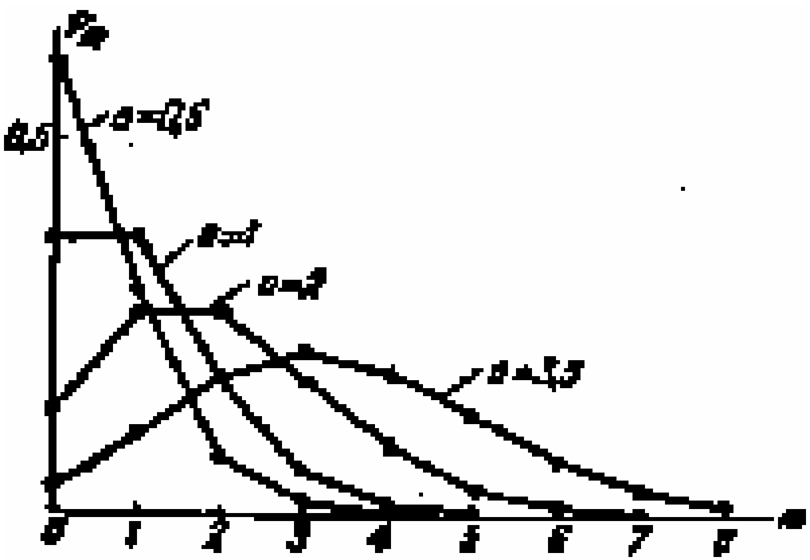
Poisson's law

$$P_m = \frac{a^m}{m!} e^{-a}$$

$$m_x = \sum_{m=0}^{\infty} m \frac{a^m}{m!} e^{-a} = a$$

Poisson's Law for Various Values *a*

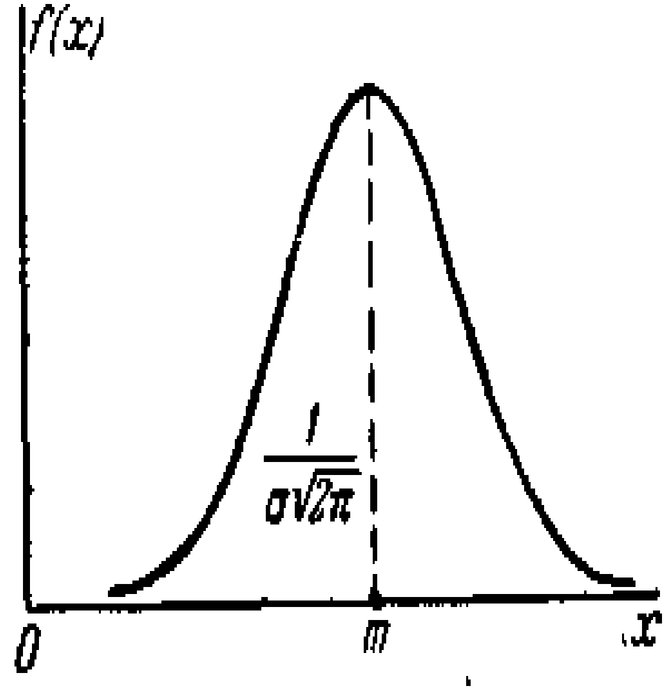
$$D_x = \sum_{m=0}^{\infty} m^2 \frac{a^m}{m!} e^{-a} - a^2 = a$$



$$R_1 = 1 - P_m = 1 - e^{-a}$$

Normal law and its parameters

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



$$M(X) = \int_{-\infty}^{\infty} xf(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} xe^{-\frac{(x-m_x)^2}{2\sigma^2}} dx = m_x$$

$$D(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x - m_x)^2 e^{-\frac{(x-m_x)^2}{2\sigma^2}} dx = \sigma^2$$

$$P(\alpha < x < \beta) = \int_{\alpha}^{\beta} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{(x-m_x)^2}{2\sigma^2}} dx$$

For odd

$$\mu_S(X) = (S-2)\sigma^2 \mu_{S-2}(X)$$

for paired central moments

$$\mu_S(X) = (S-1)!!\sigma^S$$

Laplace's function

A fragment of the function values table

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

for quantile $z = \dfrac{\beta - m}{\sigma}$

z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$
0	0	0,1	0,08	0,2	0,159	0,3	0,236	0,4	0,311
0,5	0,383	0,6	0,451	0,7	0,516	0,8	0,576	0,9	0,632
1,0	0,683	1,1	0,729	1,2	0,770	1,3	0,806	1,4	0,838
1,5	0,866	1,6	0,890	1,7	0,911	1,8	0,928	1,9	0,943
2,0	0,955	2,5	0,988	3,0	0,997	4,0	0,9999	5,0	1

Laplace theorems

Local

$$P_n(k) = \frac{1}{\sqrt{npq}} \cdot \varphi(x)$$

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

$$x = \frac{k - np}{\sqrt{npq}}$$

Integral

$$P(k_1; k_2) = \Phi(k_2) - \Phi(k_1)$$

Deviation of the relative frequency from the constant probability of independent research

$$x = \varepsilon \sqrt{\frac{n}{pq}}$$

$$P\left(\left| \frac{m}{n} - p \right| \le \varepsilon \right) = 2\Phi\left(\varepsilon \sqrt{\frac{n}{pq}} \right)$$

Other distribution functions

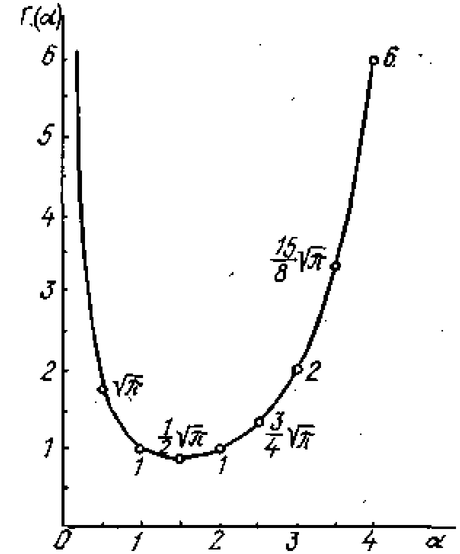
Gamma function

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$$

1) $\Gamma(1) = \Gamma(2) = 1$

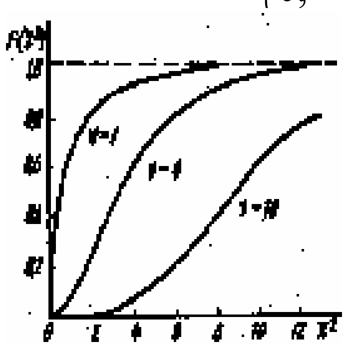
2) $\Gamma(\alpha + 1) = \alpha \cdot \Gamma(\alpha) \quad при \quad \alpha > 0;$

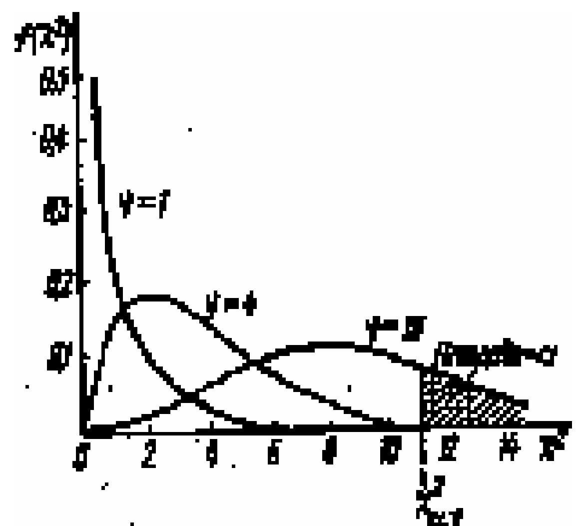
3) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$



Hi-square (χ^2)

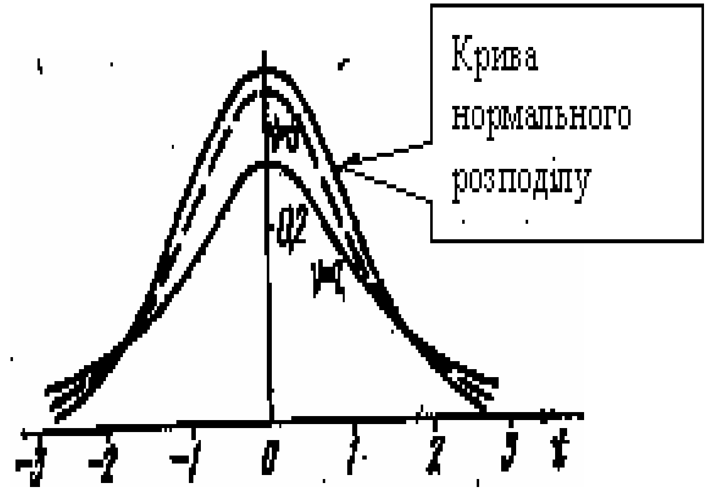
$$F(\chi^2) = P(\chi^2 < \chi_0^2) = \begin{cases} \dfrac{1}{2^{\frac{v}{2}}\Gamma\left(\frac{v}{2}\right)} \displaystyle\int_0^{\chi_0^2} \left(\chi^2\right)^{\frac{v}{2}-1} e^{-\frac{\chi^2}{2}} d\left(\chi^2\right), & якщо \quad \chi^2 \geq 0 \\ 0, & якщо \quad \chi^2 < 0 \end{cases}$$





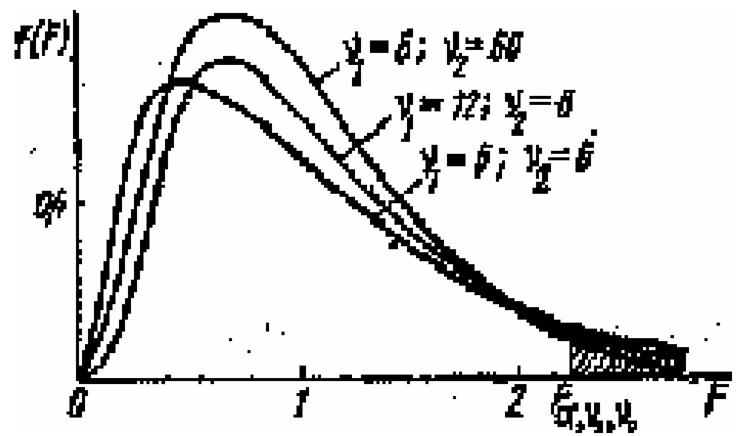
Student distributing

$$f(t) = S(t,\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\cdot\Gamma\left(\frac{\nu}{2}\right)}\cdot\left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad -\infty$$



Fisher distribution

$$f(F)= \begin{cases} \dfrac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\dfrac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} \dfrac{F^{\frac{\nu_1}{2}-1}}{\left(1+\frac{\nu_1}{\nu_2}F\right)^{\frac{\nu_1+\nu_2}{2}}} & F>0 \\ 0 & F?0 \end{cases}$$



The concept of the theory of mass service

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \qquad P_0(t) = e^{-\lambda t} \qquad \rho = \lambda / \upsilon$$

With $1 \le k \le n$ $\qquad \rho_k = \frac{\rho^k}{k!} \rho_0;$

With $k \ge n$ $\qquad \rho_k = \frac{\rho^k}{n! n^{n-k}} \rho_0$

Where $\rho_0 = [\sum_{k=0}^{n} \frac{\rho^k}{k!} + \frac{\rho^{n+1}}{n!(n-\rho)}]^{-1}$ $\qquad \rho < n$

$\rho_0 = 0$ $\qquad$ For $\rho \ge n$

The concept of the theory of reliability

$$F(t) = P(T < t) = 1 - e^{-\lambda t} \quad (\lambda > 0) \quad P(t) = e^{-\lambda t}$$

The probability of failure-free operation in the interval t_0 - t

$$P(AB) = e^{-\lambda(t_0 + t)} = e^{-\lambda t_0} e^{-\lambda t}$$

The probability of failure-free operation in the interval $t_0 – t$, if he has already worked without fail in the previous interval 0- t_0

$$P_A(B) = \frac{P(AB)}{P(A)} = \frac{e^{-\lambda t_0} \cdot e^{-\lambda t}}{e^{-\lambda t_0}} = e^{-\lambda t}$$

The probability of failure-free operation with the reserve

$$P_n(m) = \sum_{i=0}^{m} C_{m+n}^{n+i} p^{m+i} (1 - p)^{m-i}$$

Central boundary theorem

Lyapunov's theorem

$$P(Y_n < y) \underset{n \to \infty}{\overset{p}{\to}} \frac{1}{\sqrt{2\pi} \cdot \sigma_{Y_n}} \int_{-\infty}^{y} e^{-\frac{(Y_n - M(Y_n))^2}{2\sigma_{Y_n}^2}} dy$$

Muavr-Laplace theorem

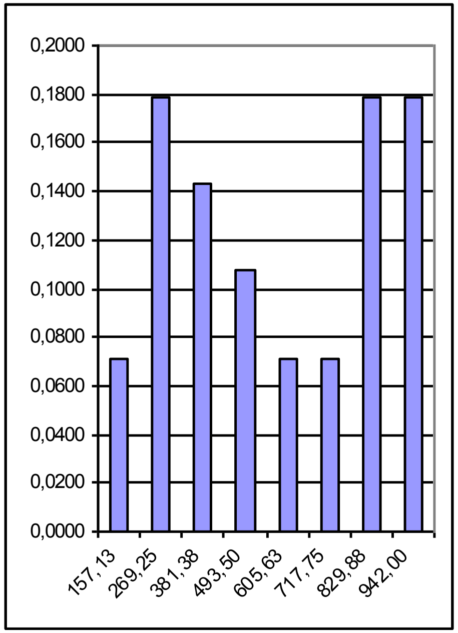
$$P(a \le Y_n \le b) \approx \frac{1}{2}\left(\Phi\left(\frac{b-np}{\sqrt{npq}}\right) - \Phi\left(\frac{a-np}{\sqrt{npq}}\right)\right)$$

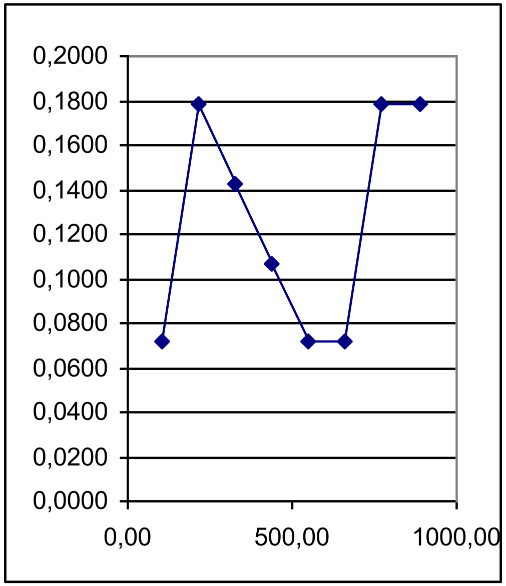
4. BASIC CONCEPTS OF MATHEMATICAL STATISTICS

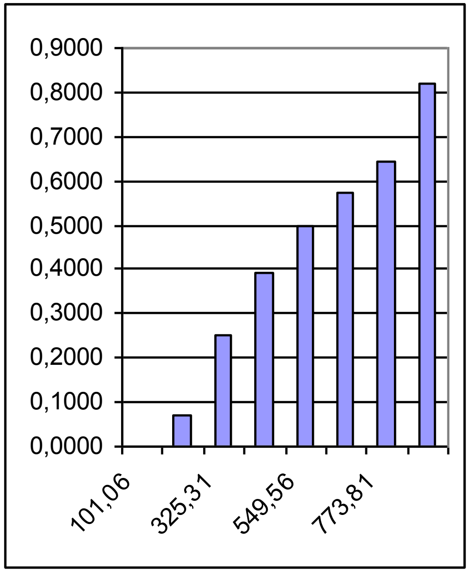
$$d_{\max}(i) = x_{\min} + \frac{(x_{\max} - x_{\min}) \cdot i}{d}$$

$$d_{op} = \frac{x_{\max} - x_{\min}}{1 + 3{,}332 \cdot \ln N}$$

$$F(d_i) = \sum_{l=0}^{i-1} k_l$$







Histogram, polygon and cumulants.

Estimates of the numerical characteristics of a random variable

$$\tilde{\alpha}_S[X] = \frac{1}{N}\sum_{i=1}^{N} x_i^{\,S}$$

$$\tilde{\mu}_S[X] = \frac{1}{N}\sum_{i=1}^{N}(x_i - \tilde{\alpha}_1)^S$$

$$\tilde{M}[X] = \tilde{\alpha}_1, \quad \tilde{D}[X] = \frac{N}{N-1}\tilde{\mu}_2$$

$$\mu_2 = \frac{1}{N}\sum_{i=1}^{N} x_i^2 - \tilde{\alpha}_1^2$$

$$\text{var}(X) = \frac{\tilde{D}(x)}{\tilde{M}(X)}$$

$$K\,\text{var}(X) = \frac{\tilde{\sigma}(x)}{\tilde{M}(X)}$$

$$\text{cov}(X,Y) = R_{XY} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \tilde{m}_x)(y_i - \tilde{m}_y)$$

$$cor(X,Y) = r_{XY} = \frac{R_{XY}}{\tilde{\sigma}_X \cdot \tilde{\sigma}_Y}$$

The law of large numbers. Chebyshev's theorem

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} x_i - \frac{1}{n}\sum_{i=1}^{n} M(x_i)\right| < \varepsilon\right) = 1$$

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} x_i - M(X)\right| < \varepsilon\right) = 1$$

Trust interval

$$P(|\Theta[X]-\tilde{\Theta}[X]|<\varepsilon)=\beta \qquad \tilde{\Theta}[X]-\varepsilon<\Theta[X]<\tilde{\Theta}[X]+\varepsilon$$

$$\varepsilon_m=\tilde{\sigma}_x Л(\beta) \qquad \varepsilon_D=\tilde{D}_x Л(\beta)\sqrt{\frac{0,8N+1,2}{N(N-1)}}$$

$$\varepsilon_{p_i}=Л(\beta)\sqrt{\frac{k_i(1-k_i)}{N}} \qquad \tilde{\sigma}_m=\sqrt{\frac{D_x}{N}}$$

Де $Л(\beta)$– зворотне значення функції Лапласа для квантиля таблиці z= $\frac{x-m_x}{\sigma_x}$

Assigning a random variable to a particular distribution law

$$P(x_i < x < x_{i+1}) = F(x_{i+1}) - F(x_i)$$

$$\chi^2 = n\sum_{i=1}^{d}\frac{(p_i - k_i)^2}{p_i}$$

$$r = d - s - 1$$