The Ministry of Foreign Affairs Republic of Uzbekistan
University of World Economy and Diplomacy

Prof. Rasulov A.S.

INTRODUCTION TO THE THEORY OF PROBABILITY AND MATHEMATICAL STATISTICS

Tashkent - 2013

PREFACE

This study guide in English was prepared based on the textbook for the course «The theory of probability and mathematical statistics» which was approved by Teaching and Methodological Council of the UWED. The material is dedicated to the students specializing in economic studies for studying the «Theory of probability and mathematical statistics». The material contains the subjects such as random events and operations on them, various definitions of probability and conditional probability, additional and multiplication rules, random variables, and their distribution functions, the law of large numbers, correlation and regression analysis. Last part of this guide devoted to the analysis of statistical data and estimation unknown parameters of distribution and interval estimation, testing hypothesis. The material teaches how to systematize the statistical data and make scientific and practical decisions based on statistical data.

Material can be used as a handbook in theoretical and practical lessons on «The theory of probability and mathematical statistics» and is suitable for independent study of the students.

Grateful acknowledge for careful and accurate technical assistance to my students Utkur Zubaydullayev and Bahtier Hodjaev.

Short historical information

The first works, in which main conceptions of probability were engendered, were presented as attempts to create the theory of gambling (Cardano, Galileo, Pascal, Fermat, Leibnitz, etc. in the XVI-XVII century)

The next stage of development of the theory of probability is connected with name of Jacob Bernoulli(1654-1705). His work «The law of large numbers» was the first theoretical base.

Further researches were made by Karl Gauss, Pierre Laplace, Simon Poisson, etc.

New, more favorable period is connected with P.L. Chebishev (1821-1894) and his students A.A. Markov, A.M. Lyapunov. At this period theory of probability was fully recognized as a mathematical science.

Further contributions were made by soviet and Russian scientists: A.N. Kolmagorov, S.N. Bernshteyn, D.Y. Khinchin, N.V. Smirnov.

Our uzbek scientists such as T.A. Sarimsokov, Kori Niyoziy, S.H. Sirojiddinov, T.A. Azlarov, Sh..K. Farmonov have also devoted their investigations to the theory of probability and statistics which is well known in the world.

Part **I. ANALYSIS OF DATA AND ELEMENTARY PROBABILITY**

1.1 ORGANIZATION OF DATA

Example: There are 100 employees of Pepsi Company. Work levels range from 1 to 6.

X_i=5,1,4,3....... By ranging date we get:

1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 3, 3,...,3,...,6,6,...,6

4-times 6-times 12-times 18-times

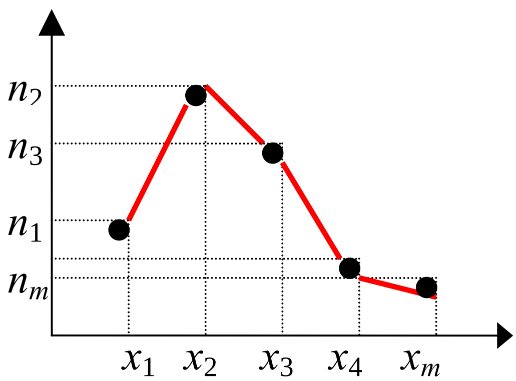
Let $n_1, n_2, n_3, \dots n_m \rightarrow$ Frequency of data and $n = \sum_{i=1}^{m} n_i$,

$m = 6$, $\frac{n_i}{n} \rightarrow$ Relative Frequency. This date could be represented as below table:

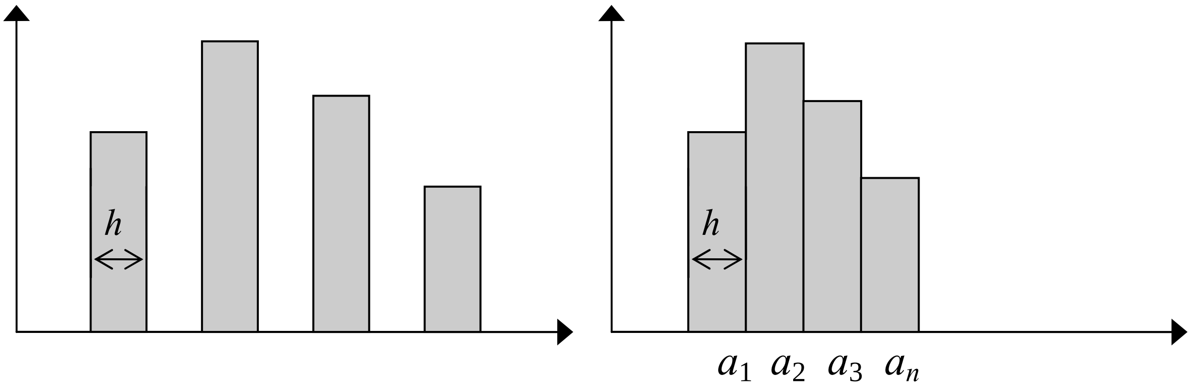
X_i	n_i	Cumulative Frequency $\sum$	Relative Frequency	Relative Cumulative Frequency
1	4	4	0.04	0.04
2	6	10	0.06	0.10
3	12	22	0.12	0.22
4	16	38	0.16	0.38
5	44	82	0.44	0.82
6	18	100	0.18	1.00

1.2 GRAPHICAL ILLUSTRATION OF DATA

1. ***Polygon*** *(relative polygon)*



2. ***Histogram*** *(bar chart)*



Sterdgess formula for the length of steps in construction bar chart will be

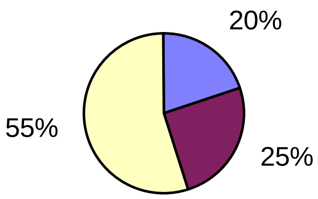
$$h = \frac{x_{\max} - x_{\min}}{1 + 3.32 \lg N}$$

Where, h-length of interval

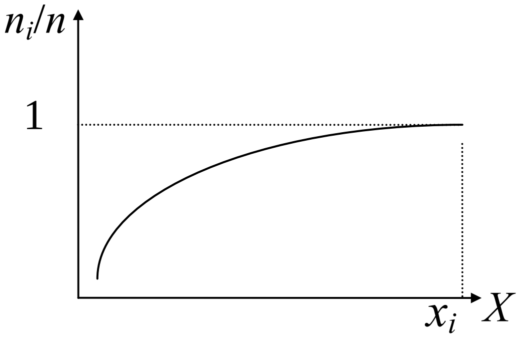
$a_1 = x_{min} - h/2;\ a_2 = a_1 + h;\ a_3 = a_2 + h = a_1 + 2h;\ \ldots\ a_n = a_1 + (n-1)h$

3. ***Stem and leaf displays (tree representation)***

4. ***Pie chart (diagram)***



5. ***Cumulative*** *(relatively cumulative polygon)*



1.3 MAIN NUMERICAL CHARACTERISTICS OF POPULATION

Let X be a set of $x_1, x_2, x_3, ..., x_n$, where n is a size of population. The main numerical characters of the population defined as follows:

1) $M_e = median = \begin{cases} x_{n+1/2}, & \text{where } n \text{ is odd;} \\ (x_{n/2} + x_{n/2-1}), & \text{where } n \text{ is even.} \end{cases}$

2) $M_d = mode = max \quad n_i = x_i;\ where \quad 1 \le i \le m$.

3) $The \ \ (arithmetical\ ) \ \ mean \quad \frac{1}{n}\sum_{i=1}^{m} x_i n_i = \overline{X}, \quad n = \sum_{i=1}^{m} n_i$

4) *The Variance:* The variance of a set of observations $x_1, x_2, x_3, ..., x_n$ denoted by S^2, is defined as following:

$$S^2 = \frac{1}{n-1}\sum (x_i - \overline{X})^2 n_i$$

where $\overline{X}$ is the *arithmetical* mean.

5) *Standard Deviation:* The positive square root of the variance is called the standard deviation and is denoted by S:

$$\sigma = \sqrt{S^2}$$

6) *Coefficient of Variation* is the ratio of the standard deviation to the mean expressed as a percentage:

$$C = \frac{\sigma}{\overline{X}} 100\,\%$$

7) *The Geometric Mean:*

$$\overline{X}_G = \sqrt[n]{\prod_{i=1}^{n} x_i}$$

8) *Harmonic mean:*

$$X_{harmonic} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

9) $Range = x_{max} - x_{min}$

1.4 INITIAL MOMENT OF A SAMPLE

Initial moment q order of a sample is

$$V_q = \frac{1}{n}\sum_{i=1}^{n} x_i^q n_i$$

where, $V_0 = 1, V_1 = \overline{X}$

1.5 CENTRAL MOMENT OF A SAMPLE

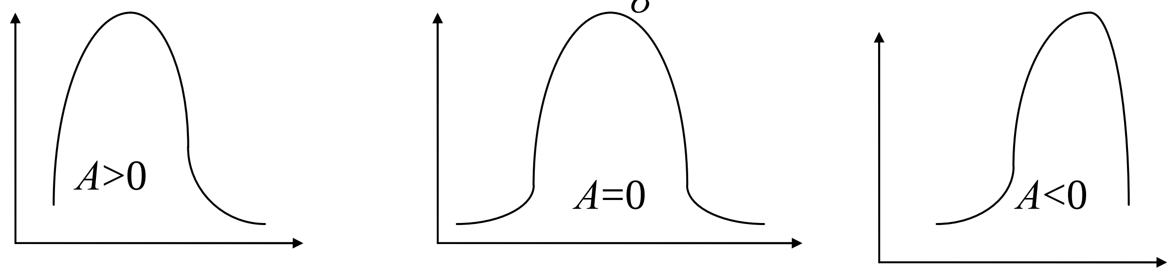
Central moment *q* order of a sample is

where $M_0 = 1,\ M_1 = 0,\ M_2 = \sigma^2$

1.6 SHAPE OF DISTRIBUTION

1) *Coefficient of Asymmetry:*

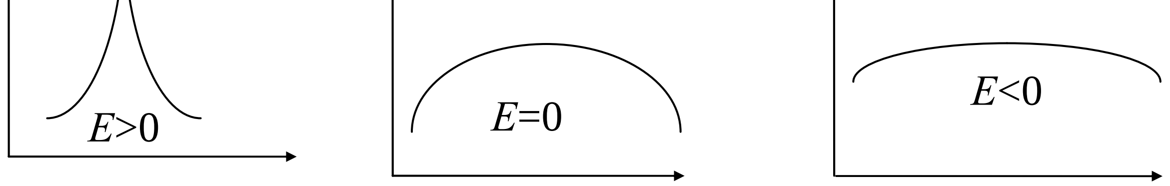
$$A = \frac{M_3}{\sigma^3}$$



Skewedness to the left or normal or to the right

2) *Excess:*

$$E = \frac{M_4}{\sigma^4} - 3$$



Flatness : sharp *normal* *gently sloping*

Example: Calculation of average annual growth rate. Let conditional profit per year will be $a_1 = 5,\ a_2 = 6,\ a_3 = 8,\ a_4 = 10,\ a_5 = 15.$

Usually the rate will calcullate as $\rightarrow \frac{a_{i+1}}{a_i}$, and $\prod_{i=1}^{n} \frac{a_{i+1}}{a_i} \rightarrow$ growth rate for n years. $\overline{T}$ - average annual growth rate. $\overline{T} - const,\ n = 4$. Calculation

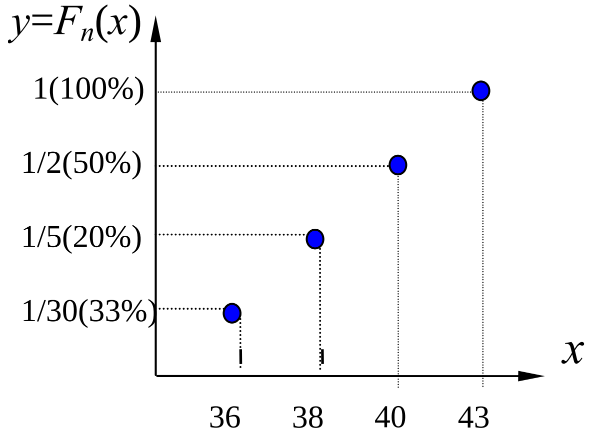
$$T_1 T_2 T_3 T_4 = \prod_{1}^{4} \frac{a_{i+1}}{a_i} \quad , \qquad \overline{T} = \sqrt[4]{\prod_{i=1}^{n} \frac{a_{i+1}}{a_i}} = \sqrt[4]{15/5} = 1.2$$

1.7 EMPIRICAL DISTRIBUTION OF FUNCTIONS

Relative cumulative frequency (r.q.f) .

$F_n(x) = \frac{m(x)}{n}$, $m(x)$ quantity of x, where $x \le A$,

36	1	1	1/30
37	1	2	2/30
38	4	6	6/30
39	4	10	10/30
40	5	15	15/30
41	8	23	23/30
42	5	28	28/30
43	2	30	1
X	**n**	**c.f**	**r.c.f**



$x<40 \rightarrow 50\%$

1.8 PROPERTY OF ARITHMETICAL AVERAGE (MEAN)

Let $\begin{pmatrix} x_1 & x_2 & x_3 \ldots x_m \\ n_1 & n_2 & n_3 \ldots n_m \end{pmatrix}$ is given, $\overline{X} = \frac{1}{n}\sum_{i=1}^{m} x_i n_i,$

where $n = \sum_{i=1}^{m} n_i,\ n_i \rightarrow$ frequency of date x_i

if all n_i=1, then (Arithmetical Average) $\overline{X} = \frac{1}{n}\sum x_i$

Then we have the following properties:

1. If $c = const$, then $\overline{c\overline{X}} = \frac{1}{n}\sum cx_i n_i$
2. $(\sum (x_i \pm c)n_i) = \overline{X} \pm c$
3. $\frac{1}{n}\sum (x_i - \overline{x})n_i = \overline{X}$, $\overline{X} = 0.$
4. $\frac{\sum x_i k n_i}{\sum k n_i} = \overline{X}$

Example from *a* cotton factory:

S_i – number of the group

X_i – the level of salary

p_i, q_i, r_i – frequency

x_i	S_1	S_2	...	S_l	
x_1	p_1	q_1	...	r_1	$p_1+q_1+\dots r_1=n_1$
...	...	...	...	...	...
x_m	p_m	q_m	...	r_m	$p_m+q_m+\dots r_m=n_m$
	N_1	N_2	...	N_l	$\Sigma N_i = N$

$\overline{x}_1 = \frac{1}{N_1}\sum_{i=1}^{m} x_i p_i$, ..., $\overline{x}_l = \frac{1}{N_l}\sum_{i=1}^{m} x_i r_i$

$\overline{X} = \frac{1}{N}\sum_{i=1}^{m} x_i n_i = \frac{1}{N}\sum_{i=1}^{l} \overline{x}_i N_i$

$\Sigma N_i = N$, Home task: Please proof the following equation:

If $z_i = x_i \pm y_i$; then $\overline{Z} = \overline{X} \pm \overline{Y}$.

1.9 VARIANCE σ^2

If $\begin{pmatrix} x_1 & x_2 & x_3 \dots x_m \\ n_1 & n_2 & n_3 \dots n_m \end{pmatrix}$ then the variance of the observations will be

$\sigma^2 = \frac{1}{n}\sum_{i=1}^{m}(x_i - \overline{X})^2 n_i$

Property of the variance:

1) If kx_i, then $k^2\sigma^2$

2) If $x_i \pm c$, then σ^2 stays constant

3) If kn_i, then σ^2 stays constant

From property 2 we can obtain the following:

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{m}(x_i - c)^2 n_i - (\overline{X} - c)^2$$

If $c = 0$, then $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 n_i - \overline{X}^2$

Simplified method of computing $\overline{X}, \sigma^2$

$$\overline{\overline{X}} = (\Sigma((x_i \pm c)/k)n_i)k/n \pm c$$

$$\sigma^2 = (\Sigma\left(\left(x_i - c\right)/k\right)^2 n_i)k^2/n - (\overline{x} - c)^2$$

1.10 ELEMENTARY PROBABILITY

Classification of events

Definition 1. An event ***E*** is some fact, which happens as a **result of** an experiment.

Definition 2. If ***E*** happens **compulsory** then ***E*** is called a **significant** event. Sometimes it is signed as ***E=U.***

Definition 3. If **E** does not happen ever, then it is called **impossible** event and ***E=∅.***

Definition 4. Another case is **E** could happen or does not happen as a result of some experiment, in this case ***E*** is called a **random event.**

Definition 5. Let's have a list of events ***E***: $A_1, A_2, ..., A_n$.. If one of them occurs as a result of any experiment, the list of ***E*** is called (or consist of) a full group of events.

Definition 6. If *A* happens and the other event $\overline{A}$ does not happen then $\overline{A}$ called an **opposite event.**

Definition 7. *A* and *B* are called **incompatible** events, if the occurrence of one of them (**except** or **eliminate), excludes** the occurrence of the other.

A and $\overline{A}$ are incompatible (complementary each other) events.

Definition 8. The situation is called **favorable** for some event, if appearance of that situation involves the happening of that event.

Definition 9. The **probability** of event A defined by formula $P(A)=\frac{m}{n}$, where m is the quantity of **favorable** situations, n-the **common number** of situations.

1.11 PROPERTY OF PROBABILITY

P.1. $0 \le P(A) \le 1;$

P.2. $P(U)=1 \Rightarrow P(\Omega)=1;$

P.3. $P(\varnothing)=0;\ P(\overline{A})=\frac{n-m}{n}=1-\frac{m}{n}=1-P(\overline{A});$

P.4. $P(A)+P(\overline{A})=1.$

Example 1. We have 25 balls, 10 of them white, 15 black. What is the probability of $P(A)$ =? , if A-white, B-black. Solution: $P(A)$=10 / 25.

Example 2. By tossing a coin, we get a head (H) or a tail (T). If event $A=H$, then $P(H)$=1/2. If $A=T$, then $P(T)$=1/2

1.12 SPACE OF ELEMENTARY EVENTS

Lets Ω - space, $x \in \Omega, x$ is the element of Ω.

$\{\Omega, U, P\}$ - called a probability space.

Example.

a) For a coin $\Omega=\{H, T\}$

b) For a die $\Omega=\{w_1, w_2, \ldots w_6\}$

U-is σ algebra of events if for $\forall\ A, B, \Omega \in U,\ A\cap B \in U,\ A\cup B\in U,\ A \setminus B \in U$.

For a die U will be $U=\{\varnothing,\{w_1\},\ldots\{w_6\}, \{w_1, w_2\},\ldots\{w_1,\ldots w_6\}=\Omega\}$.

Let's introduce the following *axioms*:

1) $\forall\ A\in U,\ P(A) \ge 0$

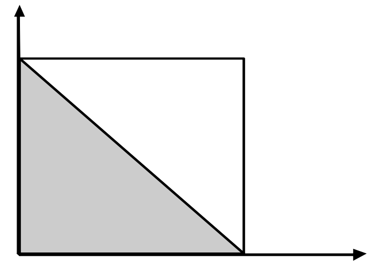
2) $P(\Omega)=1$

3)If A, B are incompatible

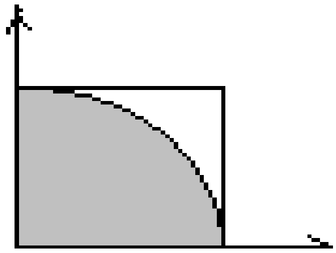
$(A\cap B)=\varnothing$, then $P(A\cup B)=P(A)+P(B)$

Definition. The probability is the numerous function in σ algebra U which satisfies the axioms 1), 2), 3) above.

1.13 GEOMETRICAL DEFINITION



Lets A-Measure of black area



and C-Measure of common area

Definition: The probability of event A (the point will be in a black area) is $P(A)=\dfrac{measure(A)}{measure(C)}$, in our case what is $P(A)$=?

$\frac{ab}{2} \to measure\ (A)$

$ab \to measure\ (C)$

$P(A) = \frac{1}{2}$

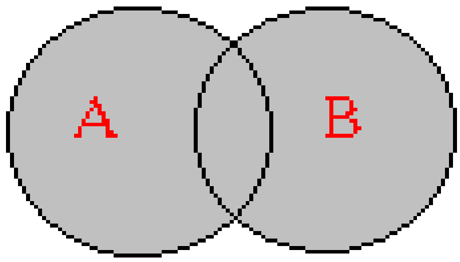
$\frac{\pi R^2}{4} \to measure\ (A)$

$R^2 \to measure\ (C)$

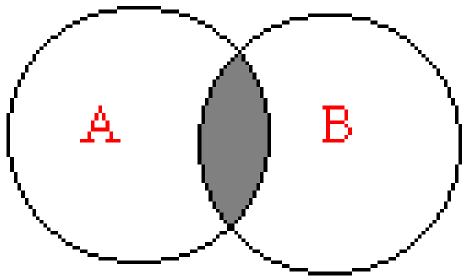
$P(A) = \frac{\pi}{4} = \frac{\pi R^2}{4R^2}$

1.14 SUM, INTERSECTION AND SUBTRACTION OF SETS (EVENTS)

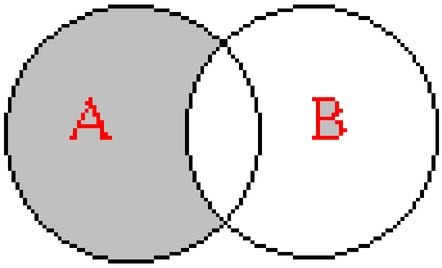
1. $C = A \cup B$, C-consists of A or B (*C is black part).*



2. C=$A \cap B$, A and B (*C is black part).*



4. $C=A \setminus B$ *(C is black part).*



1.15 ADDITIONAL RULE OF PROBABILITY

Definition. If the occurrence of one event precludes the occurrence of another, then we call the events are mutually exclusive.

I. The case of mutually exclusive :

a) If events A and B are mutually exclusive, then

$$P(A+B)=P(A)+P(B)$$

b) If $A_1, A_2, \ldots A_n$ are mutually exclusive, then

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

Example. For throwing a dice: Let $A=1$, $B=2$; $A \cup B=C$, $P(C)=1/6+1/6=1/3$

Example. For tossing a coin $P(A \cup B)=1$; Where A-head and B-tail.

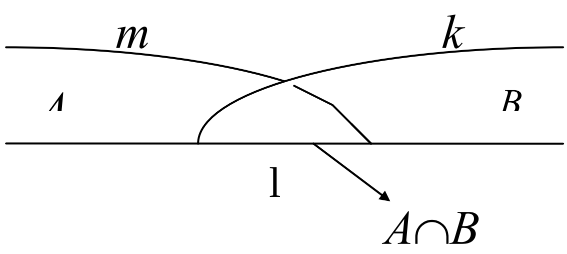
c) If $A \in B$, then $P(A) \leq P(B)$

d) If $A_1, \ldots A_n$ is a full group of events, and A_i is mutually exclusive $(A_i \cap A_j)=\varnothing$, then $P(A_1 \cup A_2 \cup \ldots \cup A_n)=P(A_1)+\ldots+P(A_n)=1$

II. The mutually not exclusive case:

If A and B are not mutually exclusive, then

$$P(A+B)=P(A)+P(B)-P(A \cap B)$$



Example 1.

$$P(A+B)=P(A)+P(B)-P(A \cap B)=\frac{m}{n}+\frac{k}{n}-\frac{1}{n}$$

Examples:

Geometrical description of this case

Example 2.

1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24

A=1,2,3,4,7,8,9,10,13,14;   *B*=3,4,5,9,10,11,15,16,17,21,22,23;

$P(A+B)=P(A)+P(B)-P(A\cap B)=(10+12-4)/24 = 18/24$

Example 3. There are 52 cards in the deck

a) *Let A*= king , *B*= queen .   What is $P(A+B)$?, Solution:
$P(A+B)=1/13+1/13=2/13$

b) *A*= spade , *B*= ace . What is $P(A+B)$? Solution:
$P(A+B)=13/52+4/52-1/52=4/13$

Example 4. There are 15 books; 5 of them are in English. Student takes any 3 books. What is the probability of taking at least one of the 3 books in English?

$B \rightarrow$ 1 English + 2 other

$C \rightarrow$ 2 English + 1 other

$D \rightarrow$ 3 English + 0 other

$A=B\cup C\cup D$ *and* *B,C,D*- are mutually exclusive, What is $P(A)$=?

$P(A)=P(B+C+D)=P(B)+P(C)+P(D)$

$P(B)=\frac{C_5^1 C_{10}^2}{C_{15}^3}$, $P(C)=\frac{C_5^2 C_{10}^1}{C_{15}^3}$, $P(D)=\frac{C_5^3}{C_{15}^3}$

***Example* 5**.There are 25 details in a box; 10 are defected among them, 15 are standard. From the box 3 details were chosen.

What is   a) P(3 standard)=?,   b) P(2 stand &1 defect)=?

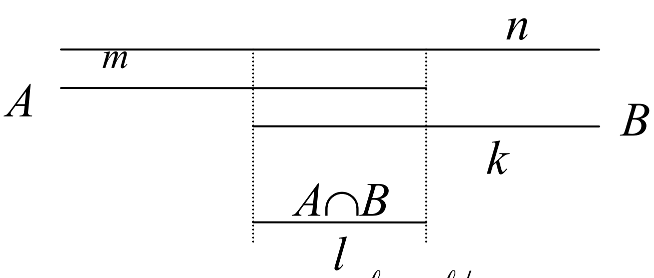
a) $P(\text{3 stand})=\frac{C_{15}^3}{C_{25}^3}$

b) $P(\text{2 stand and 1 defect})=\frac{C_{15}^2 / C_{10}^1}{C_{25}^3}$

1.16 CONDITIONAL PROBABILITY

Definition. 2 events are **independent** if the occurrence of any of them is not influenced to the probability of occurrence another. Otherwise, the events are **dependent.**

Definition. The probability is called the conditional probability of the occurrence of the event A, if it is assumed that the event B has been occurred. It is signed as ***P* (*A*/*B*) .**



$$P\,(A/B) = \frac{\ell}{k} = \frac{\ell/n}{k/n} = P(A\cap B)/P(B) \qquad (*)$$

Example. 2 dices are thrown. The sum of the numbers is less than 6. If we know that sum is even, what is $P(\text{sum}\leq 6)$?

Ω=36, B (even) = 18,

A={11,12,13,14,15,21,22,23,24,31,32,33,41,42,51}

$A\rightarrow 15$, $A\cap B\rightarrow 9$

$P(A/B)=\frac{9}{18}=\frac{9/36}{18/36}=P(A\cap B)/P(B)$,

1.17 MULTIPLICATION RULE

From **(*)** we will obtain

$\mathbf{P\,(A\cap B) = P\,(B)\,P\,(A/B)}$ **or** $\mathbf{P\,(B\cap A) = P\,(A)\,P\,(B/A)}$ **.**

Example. There are 50 details in a box. 10 of them are nonstandard (event B). One was drawn and it was not standard. What is the probability of choosing another nonstandard detail (event A) assuming that one nonstandard was already chosen?

$P(A\cap B)= \mathbf{P\,(B)\;P\,(A/B)} = \frac{10}{50}*\frac{9}{49}$

If there are n-events $A_1, A_2,...A_n$, then $P(A_1, A_2...A_n)=P(A_1) P(A_2/A_1)...P(A_n/A_1...A_{n-1})$

If the events A, B are independent, then $\mathbf{P(A) = P(A/B)}$

Theorem. If $A_1, A_2,...A_n$ are separate independent events, then the probability is their production $P(A_1,A_2...A_n)=P(A_1)...P(A_n)$. Let n=2 then $P(A\cap B)=P(A)P(B)$.

Example. What is the probability of obtaining 3 heads in 3 consecutive tosses of coin.

$$P(A_1 * A_2 * A_3) = \frac{1}{2}*\frac{1}{2}*\frac{1}{2}=\frac{1}{8}$$

1.18 THEOREM ABOUT FULL PROBABILITY

Theorem. Assume that we have $B_1, B_2, ...B_n$, n-mutually exclusive events. Assume that $B_1, B_2,...B_n$ consists of a full group of events. A-may occur only with one of B_i, i=1,…,n. For P(A) we have $\mathbf{P(A) = P(B_1)*P(A/B_1)+...+P(B_n)*P(A/B_n)}$

Proof: $A=A*U=A*(B_1+B_2+.....+B_n)=A*B_1+A*B_2+..+A*B_n$

$P(A)=P(AB_1+AB_2+..+AB_n)=P(B_1)P(A/B_1)+...+P(B_n)P(A/B_n)$

Example. There are 3 plants. They produce lamps.

1st plant produces 10% of nonstandard products

2nd plant produces 5% of nonstandard products

3rd plant produces 15% of nonstandard products

Solution:

1st plant produces 50%, 2nd - 30% and 3rd - 20% of all the lamps.

A-an event of buying nonstandard lamp.

B_1-an event of producing in the 1st plant

B_2- an event of producing in the 2nd plant

B_3- an event of producing in the 3rd plant

We should use the formula $P(A)=P(B_1)P(A/B_1)+...+P(B_n)P(A/B_n)$

$P(B_1)=0.5$, $P(A/B_1)=0.1$.

1.19 BAYES THEOREM (BAYES LAW OR FORMULA)

Theorem. Let $B_1, B_2,...B_n$ be mutually exclusive events.

A is an event for which $P(A) \neq 0$, then the conditional probability $P(B_i/A)$ for any of events B_i, given that A has already occurred is given by

$$\mathbf{P(B_i/A)=(P(B_i)*P(A/B_i))/(\Sigma P(B_i)*P(A/B_i))=(P(B_i)*P(A/B_i))/P(A)}$$

Proof.

$P(A*B_i)=P(A)*P(B_i/A)=P(B_i)*P(A/B_i)$

Calculate $P(B_i/A)$ using conditional probability.

$$P(B_i/A)=(P(B_i)*P(A/B_i))/P(A)$$

Example. The automatic machines produce similar parts. Machine A produces 40% of the total products, B-25%, C-35%. On the average, 10% of the parts turned out by A do not conform the requirements (i.e. defected), for B and C these are 5% and 1% respectively. If one part is selected randomly from the combined output and is found not to conform the requirements, what is he probability that the part is produced by machine A?

D is the event of selecting a defected part.

We need to calculate $P(A/D)=?$. From $P(A/D)=(P(A)*P(D/A)) / P(D)$ where $P(D)=P(A)P(D/A)+P(B)P(D/B)+P(C)P(D/C)$

In our case:

$P(A)=0.4$; $P(D/A)=0.1$; $P(B)=0.25$; $P(D/B)=0.05$; $P(C)=0.35$; $P(D/C)=0.01$;
$P(A/D)=0.457$, It means 45,7 %. The machine A produces 45,7% of the defected parts. B = 14, 3% and C= 40% .

1.20 FORMULA OF REPEATED TRIALS (FORMULA OF BERNOULLY)

Lets trials are independent and an event A is tossing a coin,
We will sign as p-the probability of success of an event A, q-the probability of failure of an event A.

For the 1st trial A; an events T or H will occurred.

The corresponding probability: p or q $(p+q)^1=p+q=1$

For the 2nd trial A: It could be TT or TH or HT or HH

The corresponding probability : q^2 pq qp p^2

It is true $q^2 + pq + qp + p^2 = (p+q)^2=p^2+2pq+q^2$

For the 3rd trial it could be

Events *TTT, TTH, THT, HTT, THH, HTH, HHT, HHH*
Probability p^3 p^2q p^2q p^2q pq^2 pq^2 pq^2 q^3
For this case $(p+q)^3=p^3+3p^2q+3pq^2+q^3$

Theorem. If $\boldsymbol{p}$ is the probability of success of an event, $\boldsymbol{q}$ is the probability of failure of an event in one trial, then the probability r-success of an event in n trails $\boldsymbol{P_n(r)}$ (that there are exactly r successes in n trials) is the defined term in binomial expansion of $(p+q)^n$ for which the exponent of p is r, i.e. $\boldsymbol{C_n^r p^r q^{n-r}=P_n(r)}$.

Proof. Let $\boldsymbol{r}$ is a *consecutive* success, followed by $\boldsymbol{n-r}$ consecutive failures. These $\boldsymbol{n}$ trails are independent, therefore, the desired probability is $p....pq....q=p^r q^{n-r}$

This is precisely the number of combination of **n** observations, $\boldsymbol{r}$ of them happen, and combination of occurrence equals C_n^r. The desired probability is therefore

$\boldsymbol{P_n(r)=C_n^r p^r q^{n-r}}$.

Example. Find the probability of getting exactly 5 twice as a result of 7 throws of a dice. Solution: $p=1/6$; $q=5/6$, and $P_7(2)=C_7^2 p^2 q^5$;

Example. 10% of certain articles manufactured are defective. What is the probability that in a random sample of 12 such articles at least 9 of them are defective?

$P_{12}(9 \le r)=P_{12}(9)+P_{12}(10)+P_{12}(11)+P_{12}(12)$, $p=0.1$; $q=0.9$; $n=12$.

Let $A_1, A_2,...A_n$ are mutually independent. If $P(A_i)=p$; $P(A_1 * A_2 * ... * A_n)=p^n$.

Theorem. If $A_1, A_2,...A_n$ are mutually independent. The event (A=One of A_i at least occurred $=A_1+A_2+...+A_n=A$) ; What is $P(A)$ = ?

$P(A)=P(A_1)+...+P(A_n) = 1- q_1...q_n$, here $q_i=P(\bar{A}_i)$.

Proof: $\bar{A} = \bar{A}_1 * .. * \bar{A}_n$, $P(A)=1-P(\bar{A})=1-P(\bar{A}_1 . * .. * \bar{A}_n) = 1 - q_1...q_n$.

Example. Let 3 hunters in a shoot. The probability of shooting a bird for the 1st hunter is $p_1=0.4$; for the 2nd $p_2=0.6$, for the 3rd hunter it is $p_3=0.3$.

A={One bullet is enough for hitting a bird}. What is $P(A)$?

Solution: $P(A)=1- q_1 q_2 q_3 = 1 - 0.6 \cdot 0.4 \cdot 0.3 = 0.832$.

1.21 THE LOCAL THEOREM OF MOIVRE-LAPLACE

The probability of occurrence of an event A equals to p-const. After n trails the quantity of occurrence of an event A equals to m=250. What is $P_n(m)$=?.

If p=0.4, n=600, m=250. It is very difficult to calculate $P_n(m)=C_{600}^{250}\cdot 0.4^{250}\cdot 0.6^{350}$

In this case we will give the other way of calculating $P_n(m)$=?. It is given as theorem of *Moivre-Laplace* .

Theorem. In case $n\to\infty$, $P_n(m)=\frac{f(x)}{\sqrt{npq}}$

where $f(x)=\frac{1}{\sqrt{2\pi}}\exp(-\frac{x^2}{2})$, $x=\frac{m-np}{\sqrt{npq}}$

Property of $f(x)$:

1) is even function, $f(-x)=f(x)$
2) monotonically decreasing in $x\to\infty$
3) If $x>5$, then $f(x)\approx 0$

Example. p=0.4, n=600, m=250, x=10/12≈0.83

F (0.83) ≈ 0.282 by using a table of calculation of f(x) which is given in any textbooks.

$P_{600}(250)$ = 0.282/12 = 0.0235

1.22 POISSON FORMULA (LAW)

Theorem. Let $p\to 0,\ n\to\infty,\ np=\lambda=const.$

In this case will be $P_n(m)=\frac{\lambda^m}{m!}e^{-\lambda}$.

Proof.

$$p=\frac{\lambda}{n},\ P_n(m)=C_n^m\left(\frac{\lambda}{n}\right)^m\left(1-\frac{\lambda}{n}\right)^{n-m}=\left(\left(\frac{\lambda^m}{m!}\right)\left(\frac{n(n-1)...(n-m+1)}{n^m}\right)\right.$$

$$\left.\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-m}\right)=\frac{\lambda^m}{m!}e^{-\lambda}$$

Example. The detail is not standard p=0.004, The quantity of details n=1000.

$P_{1000}(5)$ = ? Let λ=4, by formula of Poisson: $P_n(m)=\frac{\lambda^m}{m!}e^{-\lambda}$, using tables from textbook we will get $P_{1000}(5)$=0.1563, and using formula of Moivre-Laplace: $P_n(m)=\frac{f(x)}{\sqrt{npq}}$ $P_{1000}(5)$=0.1763.

1.23 INTEGRAL FORMULA OF MOIVRE-LAPLACE

Example. Let us calculate $P_{1000}(455 \le m \le 545)$ = ?

$P_{1000}(455 \le m \le 545)=P_{1000}(455)+\ldots+P_{1000}(545)$ is very difficult to calculate.

Theorem. Let the probability of event $A=p$, $n\to\infty$

then $\mathbf{P_n(a \le m \le b) = [\Phi(x_2)- \Phi(x_1)]}$,

where $x_1=\frac{a-np}{\sqrt{npq}},\ x_1=\frac{b-np}{\sqrt{npq}},\ \ \Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$

Properties of $\Phi(x)$:

1. $\Phi(x)$ is an odd function;
2. $\Phi(x)$ is monotonically increasing, so $\Phi(x_2) > \Phi(x_1)$ if $x_2>x_1$;
3. $\Phi(x)$ =1, if $x\to\infty$;
4. For $x>5$, $\Phi(x)=1$.

Example. The solution of the example above by using a table 4 from the textbook we will obtain $P_{1000}(455 \le x \le 545) = 0.9711$.

Consequence 1*.* $P_n(\,|m- n\,p\,| < r\,) \approx \Phi\left(\frac{r}{\sqrt{npq}}\right)$.

Proof: From inequality $|m- n\,p\,| < r$, we can get that np-r $\le m \le np+r$. Then

$P_n(a\le m\le b)$=$\left(\Phi\left(\frac{np+r-np}{\sqrt{npq}}\right)-\Phi\left(\frac{np-r-np}{\sqrt{npq}}\right)\right)=2\cdot\Phi\left(\frac{r}{\sqrt{npq}}\right)=$

$=2\Phi\left(\frac{r}{\sqrt{npq}}\right)$.

Consequence 2*.* Lets the frequency of event will be m/n , then

$\mathbf{P_n(\,|m/n-p\,|< \Delta) \approx 2\Phi\,(\Delta\sqrt{n/pq}\,)}$.

Proof*:* From $|m/n-p\,|< \Delta \to |m-np\,|\le n\Delta$ where $r=n\Delta$. Now using the previous results we will obtain

$P_n(|m/n-p|\leq\Delta)=2\Phi(n\Delta/\sqrt{npq})=2\Phi(\Delta\sqrt{n/pq})$.

Example 1. (for use *Consequence 1)*

Consumer needs the shoes of size 36. The probability of consumer needs size 36 is p=0.3. What is the probability among 2000 customers which were in the shop, the number of consumers who need that size varies from 570 to 630 ?

Solution: np=600, 630-600=30=600-570, r=30

$P(|m-600|\leq 30)= 2\,\Phi(1.464) = 0.8568$

Example 2. The condition from above. What is the deviation of real relative frequency from p=0.3 is less than 0.02.

$$P\left(\left|\frac{m}{2000}-0.3\right|\leq 0.02\right)=\Phi\left(0.02\frac{\sqrt{2000}}{\sqrt{0.3\cdot 0.7}}\right)=2\Phi(1.952)=0.949\,.$$

PART II. PROPERTY OF DISCRETE AND CONTINUOUS RANDOM VARIABLES AND ITS APPLICATIONS

2.1 DISCRETE RANDOM VARIABLES

In the theory of probability and statistics there are exist two types of Random Variables

- Discrete Random Variables
- Continuous Random Variables

Now we smoothly pass from Random events (about them you knows) to the Random Variables

Definition. If a random variable X takes value from a discrete set $X_1, X_2, \ldots X_n$ it is called a discrete random variable.

X_i can be the set of infinite number of variables. For full description of discrete random variable it must be given responsible probability of events X = X_i , $P\{X=X_i\}=P_i$. That function which connected the value of random variables with the probability is called distribution function. It may be presented by a table

$$X \Rightarrow \begin{pmatrix} x_1 & x_2 & x_3 \ldots x_n \\ P_1 & P_2 & P_3 \ldots P_n \end{pmatrix}$$

The following conditions must be satisfied compulsory

1. $\sum_{i=1}^{n} P_i = 1$
2. $P_i \geq 0$

The term ***Distribution function*** *is always equivalent to* ***the Law of distribution.***

Example 1. Law of distribution for tossing a coin,

when a Head = 1, a Tail = 0 , $X = \begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix}$.

Example 2. If the probability of hitting a target p=0.1, please construct the law of distribution of hitting the target in 4 time shootings.

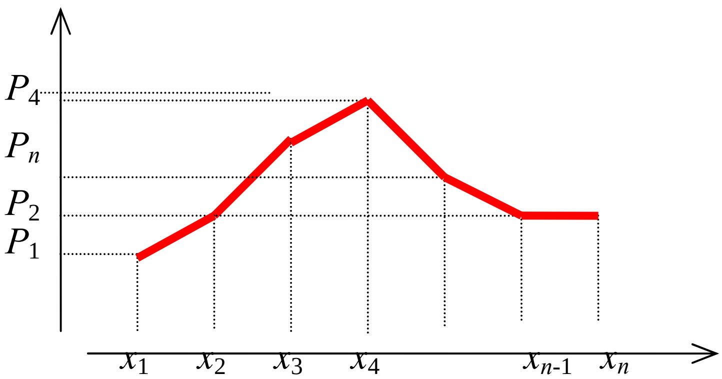
$P_4(0)=0.9^4=0.6561$, $P_4(1)=4\cdot 0.1\cdot(0.9)^3=0.2916$, $P_4(2)=0.0486$, $P_4(3)=0.0036$, $P_4(4)=0.0001$

The law of distribution is:

Quantity	0	1	2	3	4
Probability	0.656	0.2916	0.0486	0.0036	0.0001

When we use the formula of repeated trials (Bernoully formula) for computing the probability $P_n(x=m)=C_n^m p^m q^n$, the distribution function is called **Binomial law** of distribution or the ***Bernoully*** law of distribution and when we use the Poisson formula $P(x=m)=\frac{\lambda^m}{m!}e^{-\lambda}$, distribution function is called the **Poisson** distribution. A following graph shows the law of distribution

Let $X \to \begin{pmatrix} x_1 & x_2 & x_3 \ldots x_n \\ P_1 & P_2 & P_3 \ldots P_n \end{pmatrix}$



Definition. Assume that we have two discrete random variables X and Y. They are independent if the events $X=x_i$, $Y=y_i$ for any i, j=1, n are independent.

Let's calculate the probability P_{ij} where $P_{ij}=P\{X=x_i, Y=y_j\}$

If the events $X=x_i$, $Y=y_i$ are independent $P_{ij}=P\{X=x_i\}\cdot P\{Y=y_j\}$ or $P_{ij}=P_i\cdot P_j$.

In the other case, if the events are dependent and we can easily prove

$P_i=P_{ij}=P_{i1}+\ldots+P_{in}$ using the theorem of full probability (**Home work 1**).

2.2 MATHEMATICAL OPERATIONS ON RANDOM VARIABLES

If $X \to \begin{pmatrix} x_1 & x_2 & x_3 \dots x_n \\ P_1 & P_2 & P_3 \dots P_n \end{pmatrix}$ then

I. kX means $\begin{pmatrix} kx_1 & kx_2 & kx_3 \dots kx_n \\ P_1 & P_2 & P_3 \dots & P_n \end{pmatrix}$

II. $f(x)$ means $\begin{pmatrix} f(x_1) & f(x_2) & f(x_3) \dots f(x_n) \\ P_1 & P_2 & P_3 \; \dots & P_n \end{pmatrix}$

Example :

If random variable $X = \begin{pmatrix} -1 & 0 & 1 & 2 \\ 0.1 & 0.2 & 0.3 & 0.4 \end{pmatrix}$

then what is $X^2 = ?$ $X^2 = \begin{pmatrix} 0 & 1 & 4 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}$

The rules for adding, subtracting and multiplication (*X+Y, X-Y, X·Y*).

We use $P_{i,j}= P\{X=x_i , Y=y_j\} = P(X=x_i)\, P(Y=y_j)$ when X and Y not depended on each other.

Let X and Y be independent events.

Example: $X = \begin{pmatrix} 3 & 4 & 5 \\ 0.3 & 0.4 & 0.3 \end{pmatrix}$ $\qquad Y = \begin{pmatrix} 1 & 2 & 3 \\ 0.1 & 0.1 & 0.8 \end{pmatrix}$

$$X \cdot Y = \begin{pmatrix} 3 & 6 & 9 & 4 & 8 & 12 & 5 & 10 & 15 \\ 0.03 & 0.03 & 0.24 & 0.04 & 0.04 & 0.32 & 0.03 & 0.03 & 0.24 \end{pmatrix}$$

Example: $X = \begin{pmatrix} 3 & 4 & 5 \\ 0.3 & 0.4 & 0.3 \end{pmatrix}$ $\qquad Y = \begin{pmatrix} 2 & 3 \\ 0.2 & 0.8 \end{pmatrix}$

$$X - Y = \begin{pmatrix} 1 & 0 & 2 & 1 & 3 & 2 \\ 0.06 & 0.24 & 0.08 & 0.32 & 0.06 & 0.24 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 0.38 & 0.24 & 0.32 & 0.06 \end{pmatrix}$$

2.3 MATHEMATICAL EXPECTATION OR EXPECTED VALUE OF DISCRETE RANDOM VARIBLIES

If $X \to \begin{pmatrix} x_1 & x_2 & x_3 \ldots x_n \\ P_1 & P_2 & P_3 \ldots P_n \end{pmatrix}$ then the Mathematical expectation $MX = \sum_{i=1}^{n} x_i p_i$

The properties of Mathematical expectation

1) If $X=C$, $C=const$, then $MC=C$;

2) If $k \cdot X$, then $M(kX)=k\,MX$;

3) If $X+Y$ then $M(X+Y)=MX+MY$;

It is easy to prove the points 1) and 2) and we will prove the last one 3).

$$M(X+Y) = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i + y_j)P_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i P_{ij} + \sum_{i=1}^{n}\sum_{j=1}^{m} y_j P_{ij} = \sum_{i=1}^{n} x_i \sum_{j=1}^{m} P_{ij} + \sum_{j=1}^{m} y_j \sum_{i=1}^{n} P_{ij} =$$

$$\sum_{i=1}^{n} x_i P_i + \sum_{j=1}^{m} y_j P_j = MX + MY;$$

4) If $X-Y$; Using the point 2) we will get $M(X-Y)=M(X+(-1)Y)=MX-MY$;

5) If $X \cdot Y$; Then $MXY = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j P_i P_j = \sum_{i=1}^{n} x_i P_i \sum_{j=1}^{m} y_j P_j = MXMY$;

6) $M(X \pm C) = MX \pm C$; (we used the property 1)

7) $M(X-MX) = MX-MX = 0$.

2.4 VARIANCE OF DISCRETE RANDOM VARIABLES

Example: Let X is the model of die, Y is the coin

$$X = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}, \quad Y = \begin{pmatrix} 3 & 4 \\ 1/2 & 1/2 \end{pmatrix}$$

We can see that $MX=MY=3{,}5$. What is the difference X and Y ?

Definition of variance is $DX=M(X-MX)^2$

If $MX=a$ then $DX = \sum_{i=1}^{n}(x_i - a)^2 P_i$

$\sigma_x = \sqrt{DX}$ is called as the standard deviation.

Properties of variance:

1) If $C=const$, $DC=0$, $DX=M(C-MC)^2=M(C-C)^2=0$

2) $D(KX)=K^2 \cdot DX$

3) $DX=MX^2-(MX)^2$ $M(X-MX)^2=M(X^2-2XMX+(MX)^2)=MX^2-(MX)^2$

4) If X, Y **are independent**, then $D(X+Y)=DX+DY$

5) If $x_1, x_2, \dots x_n$–independent and $X = \sum_{i=1}^{n} x_i$

$\sigma_X = \sqrt{\sum_{i=1}^{n} \sigma_i^2}$, σ_i–standard deviation of x_i.

If $n=2$, than $z=x+y$ $\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2}$

6) $D(X-Y) = DX + DY$.

Example: For the die $MX=3.5$, $DX=2.917$

For the coin $MY=3.5$, $DY=0.25$

Theorem 1. Let $x_1, x_2, \dots x_n$ be identical independent random variables such that their probability distributions coincide. It means $Mx_i=a$.

Then, $M\left(\sum_{i=1}^{n} x_i\right) = na$, $M\bar{x} = a$, where $\bar{x}$ is an arithmetical average $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

Theorem 2. Let $x_1, x_2, \dots x_n$ be identical independent random variables and $Dx_i=\sigma^2$,

Then $D\left(\sum_{i=1}^{n} x_i\right) = n\sigma^2$, $D\bar{x} = \frac{\sigma^2}{n}$.

Theorem 3. Let $X = \sum_{i=1}^{n} x_i$, x_i be random variable as a result of independent trials with probability of success p.

Then the mathematical expectations $MX=np$ and $DX=npq$.

Proof: $X=x_1+x_2+\dots+x_n$, We can put $x_i = \begin{pmatrix} 1 & 0 \\ p & q \end{pmatrix}$

$Mx_i=p$, $MX=np$.

$Dx_i=(1-p)^2p+(0-p)^2q=q^2p+p^2q=pq(p+q)=pq$, $DX = \sum_{i=1}^{n} Dx_i = npq$.

Theorem 4. The probability of occurrence of event A equals to p in each independent trial. Let the relative frequency be $\frac{X}{n}$, where X is the quantity of occurred event A and n- the quantity of common trials. Then the mathematical expectation of the frequency A is equal to p, $M\left(\frac{X}{n}\right) = p$ and $D\left(\frac{X}{n}\right) = \frac{pq}{n}$.

Proof:

Relative frequency $\frac{X}{n} = \frac{\sum_{i=1}^{n} x_i}{n} \Rightarrow M\left(\frac{X}{n}\right) = p$

$$D\left(\frac{X}{n}\right) = \frac{D\left(\sum_{i=1}^{n} x_i\right)}{n^2} \Rightarrow \frac{npq}{n^2} = \frac{pq}{n} \quad , \sigma = \sqrt{\frac{pq}{n}}$$

Theorem 5. If X is distributed by law of Poisson, then $MX=\lambda$, $DX=\lambda$.

Proof: $P(X=m) = \frac{\lambda^m}{m!}e^{-\lambda}$, $MX^2=\lambda+\lambda^2$; $(MX)^2=\lambda^2$; $DX=\lambda$,m=0,1,2,..

$MX=\lambda e^{-\lambda}(1+(2\lambda)/2!+...(m\lambda^{m-1})/m!+...)=\lambda$

$M X^2=e^{-\lambda}(\lambda+2\lambda^2/1!+3\lambda^3/2!+...+m\lambda^m/(m-1)!)$

$e^{\lambda}=1+\lambda/1!+...+\lambda^{m-1}/(m-1)! +\lambda^m/(m)! + \dots$

We get $(1+\lambda)e^{\lambda}=1+2\lambda/1!+...+m\lambda^{m-1}/(m-1)!$

$(\lambda+\lambda^2)e^{\lambda}=\lambda+2\lambda^2/1!+...+m\lambda^m/(m-1)!+...$

2.5 CONTINUOUS RANDOM VARIABLES

Definition 1. If a random variable ξ takes any value out of certain interval

$[a ,b]$, we shall call a random variables ξ continuous.

Let's take $x\in R$ and we consider $\{\xi<x\}$, $\xi\in R$, $\{\xi<x\}$ is event .

Definition 2. $F(x)=P\{\xi<x\}$, $F(x)$ I s called as a distribution function (Integral function) of random variable ξ .

Example. *A* die is thrown:

x	$x\le1$	$1<x\le2$	$2<x\le3$	$3<x\le4$	$4<x\le5$	$5<x\le6$	$6<x$
$F(x)$	0	1/6	2/6	3/6	4/6	5/6	1

Property of distribution function (Integral function) *F*(*x*):

Property 1. $P\{x_1\le\xi<x_2\}=F(x_2)-F(x_1)$

Property 2. $F(-\infty)=0$, $F(+\infty)=1$, $0\le F(x)\le1$

Property 3. $F(x)$ not decreased; if $x_2>x_1$, $F(x_2)\ge F(x_1)$.

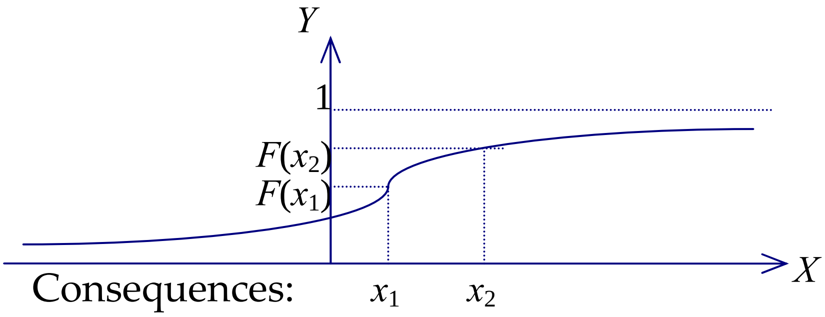
Property 4. $P(\xi=x_1)=0$

Proof. $x_1-h \le \xi < x_1+h$, $P(\xi=x_1) \le P(x_1-h \le \xi \le x_1+h)$

$P(\xi=x_1) \le F(x_1+h) - F(x_1-h)$, if $h \to 0$, then

$F(x_1+h)-F(x_1-h) \to 0$, because $F(x)$ is continuous. It means $P(\xi=x_1)=0$

Graph illustrates of $y=F(x)$



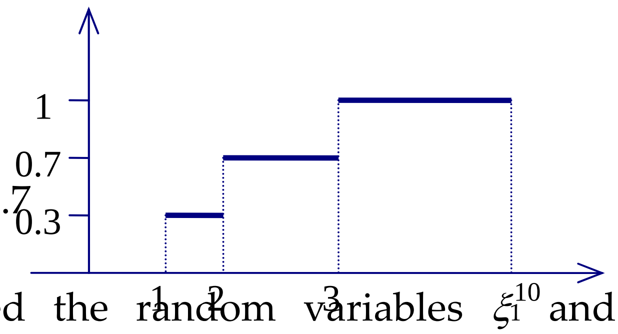
Consequences:

For continuous random variables

$P(x_1<\xi<x_2)=P(x_1\le\xi<x_2)=P(x_1<\xi\le x_2)=P(x_1\le\xi\le x_2)=P(\xi=x_1)+P(x_1<\xi\le x_2)$

Assume we are given discrete random variable

$$\xi = \begin{pmatrix} 1 & 2 & 3 \\ 0.3 & 0.4 & 0.3 \end{pmatrix}$$



If $x\le 1$, then $F(x)=0$

$1<x\le 2$, then $F(x)=0.3$

$2<x\le 3$, then $F(x)=0.3+0.4=0.7$

$3<x$, then $F(x)=0.7+0.3=1$

Definition 3. We called the random variables ξ_1 and ξ_2 are independent , if the events $\{\xi_1<x_1\}$, $\{\xi_2<x_2\}$ are independent.

Property 5. If ξ_1 and ξ_2 are independent then

$P\{\xi_1<x_1, \xi_2<x_2\}=P\{\xi_1<x_1\}P\{\xi_2<x_2\}$

Property 6. $P\{\xi_1<x_1, \xi_2<x_2\}=F_1(x_1)\cdot F_2(x_2)$

Definition 4. The random variables $\xi_1, \xi_2, \ldots \xi_n$ are independent in aggregate if for any $k \le n$, the events$\{\xi_1<x_1\},\ldots\{\xi_n<x_n\}$ are independent.

For independent random variables in aggregate $F(x_1,x_2,\ldots x_n)=P\{\xi_1<x_1,\xi_2<x_2,\xi_n<x_n\}=F_1(x_1)\cdot F_2(x_2)\cdot\ldots\cdot F_n(x_n)$

Definition 5. We shall call the random variable ξ continuous, if its distribution function is continuous, except for some break points.

Definition 6. The derivative of the distribution function $F(x)$,

$F'(x) = f(x)$,

(F-prime) is called the probability density.

Property 7. If $\xi\in[a,b]$, then $P\{\xi<a\}=0$; $P\{b<\xi\}=1$

Property 8. According to the definition $P\{\xi<x\} = P\{-\infty<\xi<x\} = F(x)$

The density functions (differential functions) $f(x)$ must satisfy two conditions:

a) $f(x) \geq 0$;

b) $\int_{-\infty}^{+\infty} f(x)dx = 1, \quad -\infty < x < +\infty$.

Property 9.

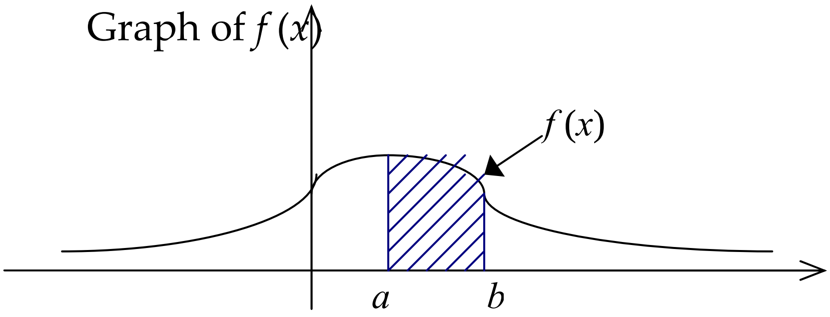
$F(x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{x} f(t)dt = \mathrm{F(x)} - F(-\infty) = \mathrm{F(x)}$.

Theorem 1. $P(a<\xi<b) = \int_{a}^{b} f(x)dx$

Proof.

$$P(a<\xi<b) = F(b) - F(a) = (\text{formula of Newton-Leibnitz}) = \int_a^b F'(x)dx = \int_a^b f(x)dx$$

Graph of $f(x)$



The probability $P(a<\xi<b)$=square under curve

Definition 7. The random variable ξ is called uniformly distributed in [a, b], if the probability density will be the following expression (function)

$$f(x) = \begin{cases} 0, & -\infty < x \leq a; \\ \dfrac{1}{b-a}, & a \leq x \leq b; \\ 0, & b < x < +\infty. \end{cases}$$

Definition 8. If $\lambda>0$, $0 \leq X < +\infty$, and the density $f(x)=\lambda e^{-\lambda x}$, then random variable ξ called as exponential law.

Definition 9. If

$$f(x) = \begin{cases} 0, & x \leq 0; \\ \dfrac{x^{k/2-1}e^{-x/2}}{\Gamma(k/2)2^{k/2}}, & x > 0. \end{cases}$$

k–number of freedom.

$\Gamma(k/2)$–gamma function, $\Gamma(m)=\int\limits_{-\infty}^{+\infty} x^{m-1}e^{-x}dx$ the random variable is called as χ^2 (Chi squared distribution).

Definition 10. If the density function f (x) is $f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$, $x\in(-\infty;\infty)$,

where a, σ - any parameters , the random variable is called normal (Gaussian) random variable.

Example. If density function $f(x)=\frac{A}{1+x^2}$, $-\infty<x<+\infty$.

a) Calculate A?

b) Calculate $P(0<\xi<+\infty)$=?

Solution:

$$\int\limits_{-\infty}^{+\infty}\frac{A}{1+x^2}dx=1=Aarctgx\Big|_{-\infty}^{+\infty}=A\pi$$

$$A=\frac{1}{\pi}\qquad P(0<\xi<+\infty)=\int\limits_{0}^{+\infty}\frac{1}{\pi(1+x^2)}dx=\frac{1}{2}$$

c) Find $F(x)$=? . *Solution:*

$$F(x)=\int\limits_{-\infty}^{x}\frac{dx}{\pi(1+x^2)}=\frac{1}{\pi}arctgx\Big|_{-\infty}^{x}=\frac{1}{2}+\frac{1}{\pi}arctgx$$

Example. For normal random variable find $F(x)$=?

Solution: From definition of distribution function

$$F(x)=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{x}e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\left\{t=\frac{x-a}{\sigma},\ x=a+t\sigma,\ dx=\sigma dz\right\}=$$

$$=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{(x-a)/\sigma}e^{-\frac{t^2}{2}}dt=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{0}e^{-\frac{t^2}{2}}dt+\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{(x-a)/\sigma}e^{-\frac{t^2}{2}}dt=\frac{1}{2}+\Phi\left(\frac{x-a}{\sigma}\right).$$

2.6 MATHEMATICAL EXPECTATION AND VARIANCE OF CONTINUOUS RANDOM VARIABLES.

Definition 1. Mathematical expectation is $M\xi=\int\limits_{-\infty}^{+\infty}xf(x)dx$, if it converges absolutely.

Definition 2. Variance is $D\xi=\int\limits_{-\infty}^{+\infty}(x-a)^2f(x)dx,\quad a=M\xi.$

2.7 NORMAL RANDOM VARIABLES

A random variable X defined on the axis (-∞, +∞) and characterized by the density $f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$, where a, σ>0 are numerical parameters, is said to be normal (or Gaussian) random variable. Here *Mx=a, Dx=*σ^2

Proof:

$$MX=a=\int_{-\infty}^{+\infty}xf(x)dx=\left\{z=\frac{x-a}{\sigma},\ x=\sigma z+a,\ dx=\sigma dz\right\}=\int_{-\infty}^{+\infty}\frac{\sigma}{\sigma\sqrt{2\pi}}(\sigma z+a)e^{-\frac{z^2}{2}}dz=$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\sigma ze^{-\frac{z^2}{2}}dz+\frac{a}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{z^2}{2}}dz=a$$

$$DX=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}(x-a)^2e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\{x-a=\sigma z;\ dx=\sigma dz\}=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}z\cdot z\cdot e^{-\frac{z^2}{2}}dz=$$

$$=\left\{u=z; dv=ze^{-\frac{z^2}{2}}dz\right\}=\sigma^2$$

If *a*=0, σ=1 then the random variable is called **standard normal random variable**. For standard normal random variable density function $f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$.

Distribution function will be $F(x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x}-e^{-\frac{(z-a)^2}{2\sigma^2}}dz$.

If standard normal random variable $F_0(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{z^2}{2}}dz$ then $F(x)=F_0\left(\frac{x-a}{\sigma}\right)$.

Consequence 1. $P(0<X<x)=\int_{0}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz=\Phi(x)$

Consequence 2. $F_0(x)=P(-\infty<X<0)+P(0<X<x)=0.5+\Phi(x)$

2.8. THE GRAPH OF DENSITY FUNCTION OF NORMAL RANDOM VARIABLE

It is known that the density function of normal random variable is

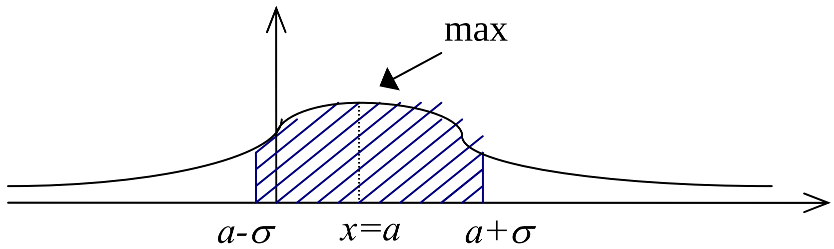
$$f(x) = y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

1. *Def* $(y) \Rightarrow (-\infty, +\infty)$
2. $y \geq 0$
3. $\lim\limits_{|x|\to\infty} y = 0$
4. Point of max and min (extremum)

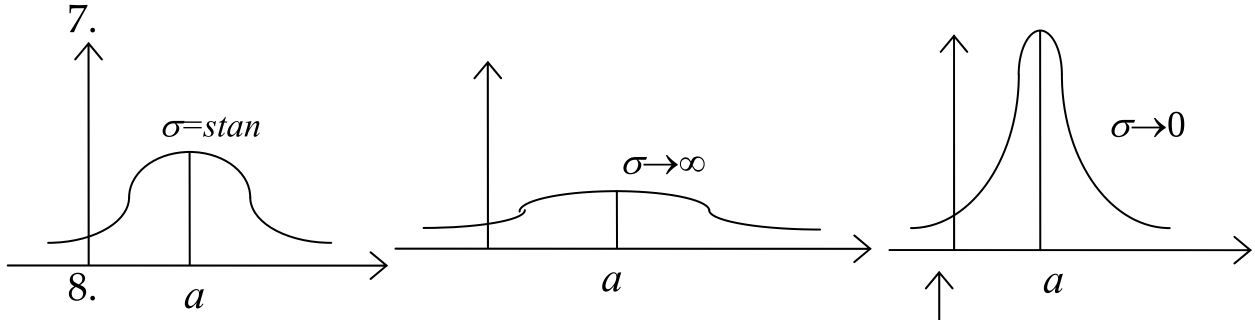
$y' = -\frac{(x-a)}{\sigma^3\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$, $y' = 0$ if $x=a$; $y' > 0$ if $x<a$ and $y' < 0$ if $x>a$ in the point, where $x=a$, $y_{\max} = \frac{1}{\sigma\sqrt{2\pi}}$

5. y is a symmetric function on line $x=a$
6. $y'' = -\frac{1}{\sigma^3\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} \left[1 - \frac{(x-a)^2}{\sigma^2}\right]$

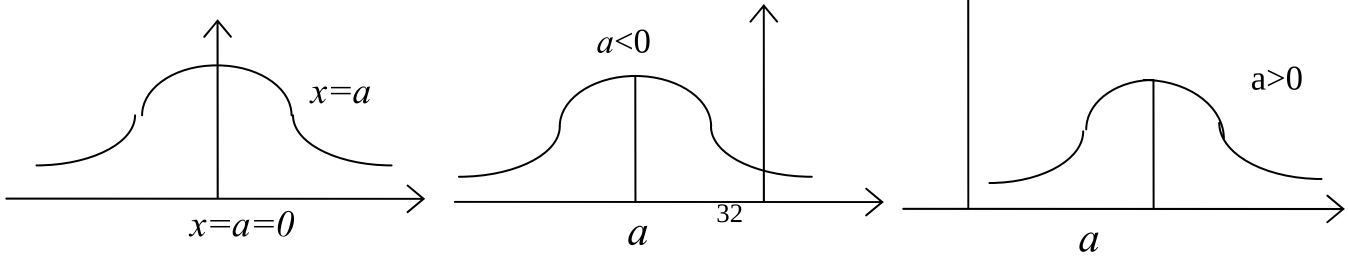
$y'' = 0$, $x=a\pm\sigma$, the points $\left(a-\sigma, \frac{1}{\sigma\sqrt{2\pi e}}\right), \left(a+\sigma, \frac{1}{\sigma\sqrt{2\pi e}}\right)$ are points of inflection.



7.



8.



Formula for computing $P(\alpha<X<\beta)$=?

If X is normal random variable with parameter (a, σ^2)

$$P(\alpha < \mathrm{X} < \beta) = \int_{\alpha}^{\beta} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \{z = \frac{x-a}{\sigma}; x = \sigma z + a;$$

$$dx = \sigma dz\} = \frac{1}{\sqrt{2\pi}} \int_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}} e^{-\frac{z^2}{2}} dz = \Phi(\frac{\beta-a}{\sigma}) - \Phi(\frac{\alpha-a}{\sigma})$$

2.8 MOMENTS OF RANDOM VARIABLES IN CONTINOUS CASE

Initial moment and central moment:

Deviation from normal distributions defined by initial or central moments. If X is a random variable $\nu_k = MX^k = \int_{-\infty}^{+\infty} x^k f(x)dx,$ where $f(x)$ is the density function.

ν_k - initial moment of k-order and μ_k - central moments of k-order

Skewed to the left Skewed to the right

μ_k = ... Positive skewedness ...$)dx$, $\mu_0=1$, $\nu_0=1$, ν_1-Negative skewedness

$-\infty$

$\mu_2=M(X-a)^2=DX=MX^2-(MX)^2=\nu_2-\nu_1^2$

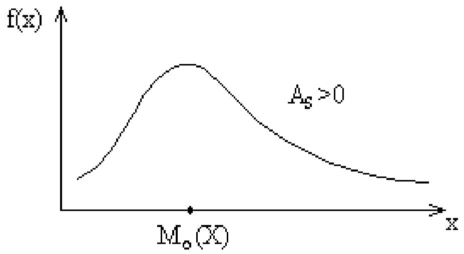
$$\mu_n = \sum_{k=0}^{n-1} (-1)^k C_n^k \nu_1^k \nu_{n-k} + (-1)^{n-1}(n-1)\nu_1^n$$

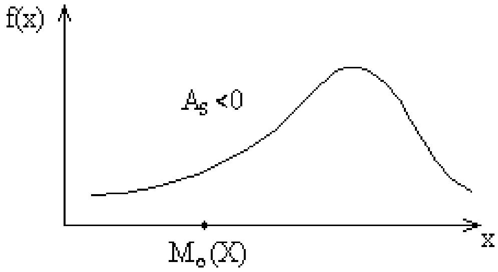
$$\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^2 \qquad \mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 + 3\nu_1^4$$

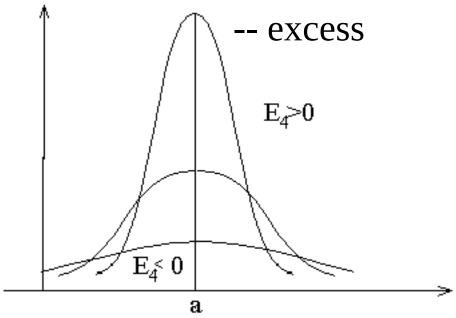
Theorem: If the moment of k-order exists, then every moment less than k exists.

Application of moments. Asymmetry $A_s = \frac{\mu_3}{\sigma^3}$ and excess $E_4 = \frac{\mu_4}{\sigma^4} - 3$

For normal random variable $\mu_4=3\sigma^4$ and $E_4=0$







2.9 CALCULATING PROBABILITY WITH GIVEN DEVIATION

Let Δ be deviation, $|x\text{-}a|<\Delta$. $P(|x\text{-}a|<\Delta)=P(a\text{-}\Delta<x<a+\Delta)=\Phi((a+\Delta\text{-}a)/\sigma)\text{-}\Phi((a\text{-}\Delta\text{-}a)/\sigma)=2\Phi(\Delta/\sigma)$, if a=0, $P(|x|<\Delta)=2\Phi(\Delta/\sigma)$

2.10 THE «THREE SIGMA» RULE

Let's$\Delta=\sigma t$, than $\Phi(\Delta/\sigma)=\Phi(t)$ t=3 $P(|x\text{-}a|<3\sigma)=2\Phi(3)=2\cdot 0.498=0.9973$. It is practically impossible to obtain in a single trial a value of X deviating from $MX=a$ by more then 3σ

$$\int_{a-3\sigma}^{a+3\sigma} f(x)dx = 0.997$$

The probability 0.997 is, thus, close to 1.

2.11 THE GENERAL SCHEME OF THE MONTE CARLO METHOD

Let's assume that we need to calculate some unknown value a=?. We shall try to find a random variable X such that $MX=a$. Let's assume $DX=\sigma^2$. We consider n random independent variable x_1, x_2,...x_n with distribution identical to x. Let's consider $X=\rho_n=\sum_{i=1}^{n} x_i$.

Central limit theorem of probability theory states that sum ρ_n of a large number of identical random variable is approximately normal.

$M\rho_n$=$M(\sum_1^n x_i)$=$n\cdot a$=$a*$

$D\rho_n$=$D(\sum_1^a x_i)$=$n\sigma^2$

$\sigma*=\sqrt{Dx}=\sigma\sqrt{n}$,

Using «three sigma» rule we obtain $P(|x-a*|<3\sigma*)=$ $P\{a*-3\sigma*<X<a*+3\sigma*\}=P\{n\cdot a-3\sigma\sqrt{n}<\rho_n<n\cdot a+3\sigma\sqrt{n}\}=P\{a-\frac{3\sigma}{\sqrt{n}}<\frac{\rho_n}{n}<a+\frac{3\sigma}{\sqrt{n}}\}=P\{|\frac{\rho_n}{n}-a|<\frac{3\sigma}{\sqrt{n}}\}=0.997$

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}x_i-a\right|<\frac{3\sigma}{\sqrt{n}}\right\}=0.997$$

$a\approx\frac{1}{n}\sum_{i=1}^{n}x_i$. Error=$\frac{3\sigma}{\sqrt{n}}$, when $n\to\infty$, error$\to 0$. By modeling with computer we can calculate the integral $a=\int_D xf(x)dx,$ where x_i-random variable with density function $f(x)$. Calculation of some common given integral. If $y=g(x)$, where x is random variable with density $f(x)$, then $MY=\int_D g(x)f(x)dx,\ MY=\frac{1}{n}\sum_{i=1}^{n}g(x_i),$ where x_i - random variable with density function $f(x)$.

2.12. THE FUNCTION OF CONTINUOUS RANDOM VARIABLE

Consider that X- continuous random variable with density function $f(x)$. If $Y=\varphi(X)$, what is the distribution function of Y ?

If $\varphi(X)$ is monotonically increasing or decreasing function, then the inverse of $\varphi(X)$ exists, $X=\psi(Y)$. The density function of random variable Y is signed by $g(y)$ and the density function $g(y)$ equals $g(y)=f(\psi(y))\cdot|\psi'(y)|$

Example 1. X is standard normal random variable. If $Y=e^X$, inverse of e^X is

$X=LnY=\psi(y)$, $\psi'(y)=\frac{1}{y},\ \ g(y)=\frac{1}{y\sqrt{2\pi}}e^{-\frac{(\ln Y)^2}{2}},\ y>0,\ g(y)$ is named as the density of logarithmic normal (log normal) random variable.

Example 2. $Y=aX+b$, $X=\frac{Y-b}{a}$, $g(y)=f\left(\frac{y-b}{a}\right)\frac{1}{|a|}$

For distribution function $F_Y(x)=P(\psi(X)<x)=$
$P\{X<\psi^{1}(x)\}=F_X(\psi^{1}(x))=F_X(\psi(x))$.

In the example 2 , $F_y(x)=F_x\left(\frac{x-b}{a}\right)$

$$F_y(x)=P(aX+b<x)=P\left(X<\frac{x-b}{a}\right)=F_X\left(\frac{x-b}{a}\right)$$

Example 3. $X\to N(a,\sigma)$, $Y=bX+c=\varphi(x)$

$$g(y)=f\left(\frac{y-c}{b}\right)\frac{1}{|b|}=\frac{1}{\sigma\sqrt{2\pi}\,|b|}\exp\left(-\frac{1}{2}\frac{(x-(ab+c))^2}{\sigma^2b^2}\right).$$

Example 4. X is uniformly distributed random variable in $\left[-\frac{\pi}{2};\frac{\pi}{2}\right]$.

Let $y=\sin X$. What is $g(y)$ = ?

Density function $f(x)=\frac{1}{\pi}$, $\sin x$ in $\left[-\frac{\pi}{2};\frac{\pi}{2}\right]$ is increasing, the inverse function is $\psi(y)=\arcsin y$,

$$\psi'(y)=\frac{1}{\sqrt{1-y^2}},\quad f(\psi(y))=\frac{1}{\pi},\quad y\in[-1;1],$$

$$g(y)=\frac{1}{\pi\sqrt{1-y^2}}$$

, *Example* 5. $X\sim N(a,\sigma)$, $Y=|X-a|$, $F_y(x)$=? $p_y(x)$=?

$F_y(x)=P\{|X-a|<x\}=P\{a-x<X<a+x\}=F_x(a+x)-F_x(a-x)$

$F'_y(x)=p_y(x)F'_x(a+x)=p_X(a+x)$; $F'_x(a-x)=-p_X(a-x)$.

Then $p_y(x)=p_X(a+x)+p_X(a-x)==\frac{1}{\sigma\sqrt{2\pi}}\left(e^{-\frac{(a+x-a)^2}{2\sigma^2}}+e^{-\frac{(a-x-a)^2}{2\sigma^2}}\right)=\frac{2}{\sigma\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}}$

In case $Y=\varphi(X)$, where φ is not a monotone function. $Y=X^2$, if $x\le 0$, $P(Y<x)=0$. $F_Y(x)=P(Y<x)$, $p_Y(x)=F'_Y(x)=0$. If $x>0$, $F_Y(X)=P(X^2<x)=$

$$=P(-\sqrt{x}<X<\sqrt{x})=F_X(\sqrt{x})-F_X(-\sqrt{x})$$

$$f_Y(x)=(F_X(\sqrt{x})-F_X(-\sqrt{x}))'=\frac{1}{2\sqrt{x}}(f_X(\sqrt{x})+f_X(-\sqrt{x}))$$

Example 6. When $X\sim N(0,1)$, $Y=X^2$, in $x>0$

$$f_Y(x)=\frac{1}{2\sqrt{x}}(f_X(\sqrt{x})+f_X(-\sqrt{x}))=\frac{1}{2\sqrt{x}}\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{x}{2}}+\frac{1}{\sqrt{2\pi}}e^{-\frac{x}{2}}\right)=\frac{1}{\sqrt{2\pi x}}e^{-\frac{x}{2}}=$$

$=f_Y(x)=f_{X^2}(x)$ and $p_X(x)=0$ if $x\le 0$

2.13 COMPOSITION OF RANDOM VARIABLE

Let $Y=\varphi(x)=x_1+x_2$, What is the law of distribution of $Y=x_1+x_2$. $F_Y(x)$=? and density function $p_Y(x)$=?

$F_Y(x)=P\{x_1+x_2<x\}=\iint\limits_A p_x(x_1,x_2)dx_1dx_2$

Where $A=\{(x_1, x_2): x_1+x_2<x\}$ $-\infty<x_1<+\infty$, $-\infty<x_2<x-x_1$

$$F_Y(x)=\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{x-x_1}p_x(x_1,x_2)dx_2\right)dx_1=\{\text{change the variable: } x_2=t-x_1\}=$$

$$=\int_{-\infty}^{+\infty}(p_x(x_1,t-x_1)dt)dx_1=\int_{-\infty}^{x}\left(\int_{-\infty}^{\infty}p_x(x_1,t-x_1)dx_1\right)dt$$

$f_Y(t)=\int_{-\infty}^{+\infty}p_x(x_1,t-x_1)dx_1-$ density of Y. If x_1 and x_2 are independent

$p_x(x_1,x_2)=p_{x_1}(x_1)p_{x_2}(x_2)$, $p_{x_1+x_2}(t)=\int_{-\infty}^{+\infty}p_{x_1}(x_1)p_{x_2}(t-x_1)dx_1$ (the formula of composition).

Example 1. Let x_1 and x_2 be independent normal random variables with parameters (0, 1); a=0, σ=1. What is the law of distribution of $Y=x_1+x_2$? Using the formula of composition we get

$$p_Y(t)=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\frac{1}{\sqrt{2\pi}}e^{-\frac{(t-x)^2}{2}}dx=?=\frac{1}{\sqrt{2\pi}\sqrt{2}}e^{-\frac{1}{2}\left(\frac{t}{\sqrt{2}}\right)^2}\Rightarrow N(0,\sqrt{2})$$

Example 2. If $x_1\to N(a_1,\sigma_1)$, $x_2\to N(a_2,\sigma_2)$, $Y=x_1+x_2$, $Y\sim N(a_1+a_2, \sqrt{\sigma_1^2+\sigma_2^2})$

Example 3. Two collective firms provide water supply system of a town. The time of expectation is random variable with exponential law of distribution with parameter λ .

1st collective firm's expectation time is x_1, 2nd collective firm's expectation time is x_2; Max expectation time of $x_1+x_2=X$. Find $P_{x_1+x_2}(t)=?$

The density function is

$$p(t)=\begin{cases}0, & \text{if } t<0;\\ \lambda e^{-\lambda t}, & \text{if } t\geq 0.\end{cases}$$

Solution:

$$p_{x_1+x_2}(t)=\int_0^t \lambda e^{-\lambda x}\lambda e^{-\lambda(t-x)}dx=\lambda^2 e^{-\lambda t}\int_0^t dt=\lambda^2 te^{-\lambda t}.$$

2.14 χ^2—DISTRIBUTION (CHI SQUARE DISTRIBUTION)

Let $x_1, x_2,..., x_\nu$ be independent normal random variable with parameters a=0, σ=1. Compose $\chi_\nu^2 = x_1^2 + x_2^2 + ... + x_\nu^2$ where ν - degree of freedom,

$$\nu=1,\quad X_1^2=x_1^2,\quad p_1(x)=\begin{cases}0, & \text{if } x\le 0;\\ \dfrac{1}{\sqrt{2\pi x}}e^{-\frac{x}{2}}, & \text{if } x>0.\end{cases}$$

$$\nu=2,\quad X_2^2=x_1^2+x_2^2,\quad p_2(x)=\begin{cases}\dfrac{1}{2}e^{-\frac{x}{2}}, & \text{if } x>0;\\ 0, & \text{if } x\le 0.\end{cases}$$

$\forall\,\nu$, $p_\nu(x)=k_\nu e^{-x/2}x^{\nu/2-1}$, where k_ν can be found from condition $\int_{-\infty}^{+\infty} p_\nu(x)dx=1$. Home task: : Find k_ν .

Let $x_1, x_2,..., x_\nu$ be independent normal random variable and $x_k \sim N(a, \sigma)$, $k=1, \nu$.

Changing the variables we can obtain a standard normal random variable

$$Y_i=\frac{x_i-a}{\sigma} \sim N(0,1)$$

$$\chi_\nu^2=\sum_{i=1}^{\nu}\frac{(x_i-a)^2}{\sigma^2}\to\chi^2\text{distrubition.}$$

In the table of χ^2 you may find distribution $P\{\chi_\nu^2>\chi_{\alpha,\nu}^2\}=\alpha$.

For ν>30 the value $\chi_{\alpha,\nu}^2$. From the normal distribution table #4 we can find it out.

2.15 STUDENT (OR T-) DISTRIBUTION

Let $x_0, x_1,..., x_\nu$ be independent normal random variables $x_i \sim N(0,\sigma)$, i=0, 1, 2,..., ν

The random variable $t_\nu = \dfrac{X_0}{\sqrt{\dfrac{1}{\nu}\sum_{i=1}^{\nu} x_i^2}}$, is called *t*-distribution or Student Distribution random variable with degree of freedom *ν*.

The Density function is $p_{t_\nu}(x) = b_\nu\,(1+x^2/2)^{-(\nu+1)/2}$

Let $F(x,\sigma)$ be distribution function. If x_i is independent, then $Y_i = x_i/\sigma$ is independent as well and $Y_i \sim N(0,1)$

$$F(x,\sigma) = P\left\{\frac{x_0}{\sqrt{\dfrac{1}{\nu}\sum_1^{\nu} x_i^2}} < x\right\} = P\left\{\frac{\sigma Y_0}{\sqrt{\dfrac{1}{\nu}\sum_1^{\nu} (\sigma Y_i)^2}} < x\right\} = P\left\{\frac{Y_0}{\sqrt{\dfrac{1}{\nu}\sum_1^{\nu} Y_i^2}} < x\right\} = F(x,1).$$

The distribution is not dependent on σ. If $x_0, x_1, x_2, \ldots, x_\nu$ is independent and $x_i \sim N(a, \sigma)$ $t_\nu = \dfrac{x_0 - a}{\sqrt{\dfrac{1}{\nu}\sum_1^{\nu}(x_i - a)^2}}$

If $x_i \sim N(0,1)$, $t_\nu = \dfrac{x_0}{\sqrt{\dfrac{1}{\nu}X_\nu^2}}$, then X_ν^2 has χ^2 distribution.

In the table of *T* distribution for $t_{\nu,\alpha}$ is given

$P\{|t_\nu| < t_{\nu,\alpha}\} = 1-\alpha$

2.16 F—DISTRIBUTION (DISTRIBUTION OF FISHER)

Let $x_1, x_2, \ldots, x_{n1}, x_{n1+1}, \ldots, x_{n1+n2}$ be independent normal random variable. $x_i \sim N(0, \sigma)$, $i=1,\ldots, n_1+n_2$

The random variable $F_{n_1 n_2} = \dfrac{\dfrac{1}{n_1}\sum_{i=1}^{n_1} x_i^2}{\dfrac{1}{n_2}\sum_{j=n_1+1}^{n_1+n_2} x_j^2}$ is called

F **distribution random variable with degrees of freedom (*n*1, *n*2).**

If $x_i \sim N(a,\sigma)$ $i=1,\ldots, n_1+n_2$, then

$$F_{n_1n_2}=\frac{\frac{1}{n_1}\sum_{i=1}^{n_1}(x_i-a)^2}{\frac{1}{n_2}\sum_{j=n_1+1}^{n_1+n_2}(x_j-a)^2}\quad \text{have } F\text{—distribution}$$

If σ=1, $F_{n_1n_2}=\dfrac{\frac{1}{n_1}X_{n_1}^2}{\frac{1}{n_2}X_{n_2}^2}$, where $X_{n_1}^2$ and $X_{n_2}^2$, are χ^2 distributed variables.

In the table of Fisher distribution it is given that $P\{F_{n1,n2}>F_{n1,n2,\alpha}\}=\alpha$

The density function

$$P_{n_1n_2}(x)=\begin{cases}0, & \text{when } x\le 0;\\ C_0\dfrac{x^{(n_1-2)/2}}{(n_1x+n_2)^{(n_1+n_2)/2}}, & \text{when } x>0.\end{cases}$$

Here, $C_0=\dfrac{\Gamma(\frac{n_1+n_2}{2})n_1^{n_1/2}n_2^{n_2/2}}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})}$, $\Gamma(x)=\int_0^\infty t^{x-1}e^{-t}dt$

2.17 MULTIVARIABLE RANDOM VARIABLES

Let X_1, X_2,..., $X_n\in\Omega$, $X=(X_1,X_2,...,X_n)$ be random vector with n-dimension. We shall define the distribution function in following way

$F_X(x_1, x_2,..., x_n) = P\{X_1<x_1,...,X_n<x_n\}$

Here, $F_X(x_1,x_2,...,x_n)$ – not decreasing function.

If positive function $p_X(x_1,x_2,...,x_n)$ exists,

$$F_X(x_1,x_2,...,x_n)=\int_{-\infty}^{x_1}dt_1...\int_{-\infty}^{x_n}p_X(t_1,t_2,...,t_n)dt_n$$

$p_X(x_1,x_2,...,x_n)$ is called as density function.

$$p_X(x_1,x_2,...,x_n)=\frac{\partial^n F_X(x_1,x_2,...,x_n)}{\partial x_1...\partial x_n}$$

Let A be subset of E_n, $X\in A$. We will use the following formula in order to compute $P(X\in A)$

$$P(X\in A)=\int_A p_x(x_1,x_2,...,x_n)dx_1...dx_n.$$

We know that random variables $x_1, x_2,..., x_n$ are independent if $\forall x_i$, $i=1,n$; the events $\{X_1<x_1\},...,\{X_n<x_n\}$ are independent.

For independent random variables $X_1, X_2,..., X_n$, if $X=(X_1, X_2,..., X_n)$ the distribution function $F_X(x_1,x_2,...,x_n)=F_{X_1}(x_1)...F_{X_n}(x_n)$ and

$p_X(x_1,x_2,...,x_n)=p_{X_1}(x_1)...p_{X_n}(x_n)$.

If $X_1\in A_1$, $X_2\in A_2$,... $X_n\in A_n$ and is independent, then $P(X_1\in A_1,...X_n\in A_n)=P(X_1\in A_1)...P(X_n\in A_n)$

Properties: Let $n=2$, $X=(X_1, X_2)$.

1. If $x_i\to-\infty$ $F_X(x_1, x_2)=F_X(x_1, -\infty)=\lim F(x_1, x_2)=$
$=\lim P(\{X_1<x_⊚\}\cap\{X_2<x_2\})=P(\varnothing)=0$ $\underline{F_X}(x_1, -\infty)=0$

2. If $F_X(x_1, +\infty)=F_{X_1}(x_1)$, $F_X(x_1, +\infty)=\lim F_X(x_1, x_2)=$
$=\lim P(\{X_1<x_1\}\cap\{X_2<x_2\})=P\{X_1<x_1\}=F_{X_1}(x_1)$.

$$F_{X_1}(x_1)=F_X(x_1, +\infty)=\int_{-\infty}^{x_1}\left(\int_{-\infty}^{+\infty}p_X(y_1,y_2)dy_2\right)dy_1$$

$\int_{-\infty}^{+\infty}p_X(y_1,y_2)dy_2=p_{X_1}(y_1)$ -density function of X_1

$$F_{X_1}(x_1)=\int_{-\infty}^{+\infty}p_X(x_1,x_2)dx_1.$$

3. $p_{X_2}(x_2)=\int_{-\infty}^{+\infty}p_X(x_1,x_2)dx_1$

4. Let Σ be a symmetric matrix $(n\times n)$ and

$$x=\begin{pmatrix}x_1\\ .\\ x_n\end{pmatrix}\qquad a=\begin{pmatrix}a_1\\ .\\ a_n\end{pmatrix}$$

Where a_{ij} is element of Σ^{-1}, $x'=(x_1, ..., x_n)$, $a'=(a_1, ..., a_n)$;

Then n dimension normal random variable x will have density function

$$p_X(x_1,x_2,...,x_n)=\frac{1}{(2\pi)^{n/2}|\Sigma|}\exp\{\frac{1}{2}(x'-a')\Sigma^{-1}(x-a)\}=$$

$$=\frac{1}{(2\pi)^{n/2}\sqrt{|\Sigma|}}\exp\{-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}(x_i-a_i)(x_j-a_j)\}\ .$$

If $n=2$, $p_X(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\exp\{-\frac{1}{2(1-r^2)}(z_1^2-2rz_1z_2+z_2^2)\}$.

Where $z_i=(x_i-a_i)/\sigma_i$, $i=1,2$

σ_1, σ_2, a_1, a_2 and r some parameters of normal random variable, $\sigma_1>0$, $\sigma_2>0$, $|r|<1$

Homework. 1) $X=(X_1, X_2)$, X is 2 dimensional normal random variable with $a_1=a_2=0$, $\sigma_1=\sigma_2=1$. Find $p_{X_1}(x_1)=?,\ p_{X_2}(x_2)=?$

2) $X=(X_1, X_2)$ and X is 2 dimension random variable with parameters

$a_1, a_2, \sigma_1, \sigma_2, r$

Find $p_{X_1}(x_1)=?,\ p_{X_2}(x_2)=?$

2.18 CONDITIONAL DENSITY FUNCTION

Let $Z=(X, Y)$ be a random variable with density function of $Z=p_z(x, y)$.

Let $p_Y(y)$ is density function of Y.

$p_Y(y_0)\neq 0$, if $Y=y_0$.

Definition: Conditional density function of random variable X under condition $y=y_0$ is $p_X(x/y)=p_z(x, y)/p_Y(y_0)$.

Conditional density function of Y in $x=x_0$ is $p_Y(y/x_0)=p_z(x, y)/p_X(x_0)$.

Properties:

1) $\int\limits_{-\infty}^{+\infty} p_X(x/y)dx=1$

2) $\int\limits_{-\infty}^{+\infty} p_Y(x/y)dy=1$

3) $p_X(x_1,x_2..x_n)=p_1(x_1)p_2(x_2/x_1)p_3(x_3/x_1,x_2)\ldots p_n(x_n/x_1,..x_{n-1})$

$$p_1(x_1)=\int(n-1)\int\limits_{-\infty}^{+\infty} p_X(x_1,...,x_n)dx_2...dx_n$$

$$p_2(x_2/x_1)=\int\int\limits_{-\infty}^{+\infty} p_X(x_1,...,x_n)dx_3...dx_n[p_1(x_1)]^{-1}$$

$$p_{n-1}(x_{n-1}/x_1,...,x_{n-2}) = \int_{-\infty}^{+\infty} p_X(x_1,...,x_n)dx_n[p_1(x_1)]^{-1}$$

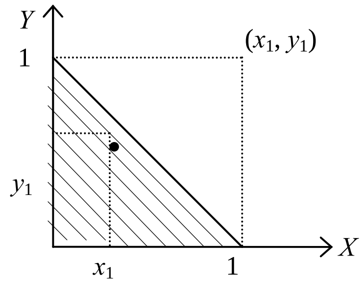
$$p_{n-1}(x_{n-1}/x_1,...x_{n-2}) = \int_{-\infty}^{+\infty} p_X(x_1,...x_n)dx_n[p_1(x_1)p_2(\cdots)[p_{n-2}(x_{n-2}/x_1,...x_{n-3})]^{-1}$$

$$p_n(x_n/x_1,..x_{n-1})=(p_X(x_1, x_2\dots x_n))/(p_1(x_1)\dots p_{n-1}(x_{n-1}/x_1\dots x_{n-2}))$$

Example. Area $x+y<1$, $x>0$, $y>0$

Joint density is $p_z(x, y)= 6x$. What is $p_X(x)$ and $p_Y(y/x)$?

$$p_X(x) = \int_0^{1-x} p_z(x,y)dy = 6x(1-x),\ \ 0<x<1;$$



$$p_Y(y/x) = \frac{p_z(x,y)}{p_x(x)} = \frac{1}{1-x},\ \ 0<y<1-x;$$

$$F_X(x)=\int_0^x p_z(t)dt = 3x^2-2x^3,\ \ 0<x<1;$$

$$F_Y(y/x)=\int_0^y p_y(y/x)dy = \frac{y}{1-x},\ \ 0<y<1-x.$$

Homework. Let $z=(x, y)$ 2-dimension normal random variable with

$a_1=a_2=0$, $\sigma_1=\sigma_2=1$ find $p_z(x/y)$=?

Homework. In the example $p_z(x, y)=6x$; find $p_Y(y)$, $p_Z(x/y)$=?

2.19 CONDITIONAL MATHEMATICAL EXPECTATION

Mathematical expectation of random variable

X, when $Y=y$, is defined as

$$M(X/y) = M(X/Y=y) = \int_{-\infty}^{+\infty} xp_X(x/y)dx = \frac{1}{p_Y(y)}\int_{-\infty}^{+\infty} xp_Z(x,y)dx$$

or $M(Y/X=x) = M(Y/x) = \frac{1}{p_X(x)}\int_{-\infty}^{+\infty} yp_Z(x,y)dy$.

The function $f_X(y)=M(X/y)$ depends on y. Conditional mean of X depends on y.

We shall call $f_X(y)$ as the regression function of X on Y .

$M(X/Y) = f_X(y)$ function of random variable Y.

$$M(M(X/y)) = \int_{-\infty}^{+\infty} M(X/y)p_Y(y)dy = \int_{-\infty}^{+\infty}(\int_{-\infty}^{+\infty} x\frac{p_Z(x,y)}{p_Y(y)}dx)p_Y(y)dy =$$

$$\int_{-\infty}^{+\infty} x(\int_{-\infty}^{+\infty} p_Z(x,y)dy)dx = \int_{-\infty}^{+\infty} xp_X(x)dx = MX$$

2.20 THE LAW OF LARGE NUMBERS

If the probability of an event A is small, the occurrence of A is practically impossible.

The probability which we can not accept for practical using is called as the level of significance and is denoted by α.

If the level of significance α=0.05, we shall use it for initial study and when α=0.01 for conclusion.

Definition. The law of large numbers is the set of conditions when the probability that the deviation of arithmetical mean from arithmetical mean of mathematical expectation is not less than given ε>0, tends to

2.21 CHEBISHEV'S INEQUALITY

Lemma. If among the values of random variables X is not negative, then the probability of $X > A$, $A > 0$, is not less than MX / A, $P\{ X > A \}<MX / A$.

Proof. Let $P(X = x_i) = p_i$, $X=x_i$ we shall put the values of X in order to increase and A will divide the sequence in two parts, and $x_i > 0$.

According to definition $x_1 \cdot p_1+x_2 \cdot p_2+ \dots x_n \cdot p_n = MX$

Let the first k members $x_i < A$, excluding the 1st k members $x_i \cdot p_i$ we can obtain:

$x_{k+1}p_{k+1}+\dots+x_np_n \le MX$, $x_k \approx A$ $A(p_{k+1}+\dots+p_n) \le MX$

$p_{k+1}+\dots+p_n \le MX/A \rightarrow P(X>A) \le MX/A$

Theorem. Let X be random variable then $P(|X-a|>\varepsilon)< DX/\varepsilon^2$; $a=MX$

Proof. Let's assume that $(x-a)^2 \ge 0$ the random variable

$(x-a)^2 \sim X$, $\varepsilon^2 \sim A$

$P((X-a)^2>\varepsilon^2)<M(X-a)^2/\varepsilon^2$ $(X-a)^2>\varepsilon^2 \rightarrow |X-a|>\varepsilon$

$M(X-a)^2=DX$ $P(|X-a|>\varepsilon)<DX/\varepsilon^2$

For $|X-a| \le \varepsilon$ $P(|X-a| \le \varepsilon)=1-P(|X-a|>\varepsilon) \rightarrow P(|X-a| \le \varepsilon) \ge 1-DX/\varepsilon^2$

Example 1. 80% of seed corn is sprout. X is the quantity of sprouting real seed corn. 10000 of seeds were overlooked (sow). $X/10000$ is the share of sprouting. What is the $P(|(x/10000)-p|<0.01)=?$

Solution:

$p=0.8$, $q=0.2$, $n=10000$, $\varepsilon=0.01$, $DX = npq$, $D(X/n)=pq/n$,

$P(|(x/10000)-0.8|\le 0.01) \ge 1-(0.8\cdot 0.2)/(0.01^2 \cdot 10000)= 0.84$

Example 2. The probability that the detail is not standard is 0.1. Why can't we use Chebishev's inequality in calculating the probability that among 10000 details the quantity of not standard details is in interval [950, 1030]?

Solution: X is the quantity of not standards, $n=10000$, $p=0.1$, $q=0.9$. $MX=np=1000$, $DX=10000\cdot 0.1\cdot 0.9=900$ $MX-950=1000-950=50$, 1030-1000=30

instead of 1030 we should take 1050

$950 \le x \le 1050 \rightarrow |x-1000| \le 50$

$P(|x-1000| \le 50) \ge 1-900/50^2 = 0.64$

2.22 THEOREM OF CHEBISHEV

Theorem. Let $X_1, X_2, \ldots X_n$ be independent random variables and $DX_i<C$, C=const. $MX_i=a_i$, then

$$P\left(\left|\frac{1}{n}\sum X_i - \frac{1}{n}\sum a_i\right| \le \varepsilon\right) > 1-\sigma, \text{ where } \sigma = \frac{c}{n\varepsilon^2}.$$

Proof: $X = \frac{1}{n}\sum_{i=1}^{n} X_i$,

$$MX = \frac{1}{n}\sum_{i=1}^{n} a_i \quad DX = D\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}D\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum DX_i \le \frac{c}{n}$$

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i\right) - \frac{1}{n}\sum_{i=1}^{n} a_i\right| \le \varepsilon) > 1 - \frac{c}{n\varepsilon^2} = 1-\sigma$$

Consequence 1. If $MX_i = a$, $DX_i < c$, then $P\left(\left|\frac{1}{n}\sum X_i - a\right| \le \varepsilon\right) > 1-\sigma.$

Because, $\frac{1}{n}\sum a_i = \sum \frac{Mx_i}{n} = \frac{na}{n} = a.$

Consequence 2. (Theorem of Poisson). If the probability of event A in i trials equals p_i and X is the quantity of occurrence of event A.

$\frac{X}{n}$ - relative frequency. $P\left(\left|\frac{x}{n}-\frac{p_1+...+p_n}{n}\right|\le\varepsilon\right) > 1-\sigma.$

Proof: $\frac{X}{n}=\frac{x_1+...+x_n}{n}$, $MX_i=p_i$, $DXi=p_iq_i \le 0.25$.

The condition of Chebishev theorem is satisfied.

$$P\left(\left|\frac{x}{n}-\frac{p_1+...+p_n}{n}\right|\le\varepsilon\right) > 1-\frac{0.25}{ne^2}.$$

Consequence 3. (Theorem of Bernoully) If the probability of event *A* is equal to *p*, then $P\left(\left|\frac{X}{n}-p\right|\le\varepsilon\right) > 1-\sigma.$ $(\frac{X}{n}\approx p)$

Example 3. $X_1, X_2,...,X_n$ is the sequence of independent random variables,

$$X_k=\begin{pmatrix}-k\alpha & 0 & k\alpha\\ \frac{1}{2k^2} & 1-\frac{1}{k^2} & \frac{1}{2k^2}\end{pmatrix}$$

Is the law of large numbers (Theorem of Chebishev) applied or not?

Solution:

$MX_k=0$, $DX_k=MX_k^2-(MX_k)^2=(-k\alpha)^2/2k^2+0^2(1-1/k^2)+(k\alpha)^2/2k^2=\alpha^2=C$

The conditions of Chebishev theorem are satisfied.

Home task:

Example 1. The quantity of water is a random variable with $MX=125m^3$. Evaluate the probability that next day some firm needs more than $500m^3$ of water?

Example 2. The probability of the passengers' being late for the train is equal to 0.007. The number of passengers (out of 20000) who are late varies in interval [100, 180]. What is the probability of that event?

Example 3. The probability of that a consumption will be made in the shop is equal to 0.65. Why is not it possible to apply Chebishev's inequality for evaluation of the probability that among 2000 consumers the number of consumers who purchased something is in interval [1260, 1360]. Solve the problem with changing the left border.

Example 4. Variance of each independent random variable out of 2500 does not exceed 5. Evaluate the probability that an absolute value of deviation of average arithmetical of their math. Expectation does not exceed 0.4.

Example 5. The probability of making a defected detail is equal to 0.8. Why cannot Chebishev's inequality be used for evaluating that share of defected detail out of 4000 can be in interval [0.78; 0.89].

2.23 THE MEASURE OF RELATIONSHIP OF RANDOM VARIABLES

The main character of stochastically relationship gives a co-variation. Let's have random variables x_i, x_j.

We shall call the value $\sigma_{ij}=cov(x_i, x_j)=M(x_i-Mx_i)(x_j-Mx_j)$ as covariance.

$\sigma_{ij}=M(x_ix_i)-Mx_i\,Mx_j$, because $\sigma_{ij}=M(x_ix_j-x_jMx_i-x_iMx_j+Mx_iMx_j)=M(x_ix_j)-Mx_iMx_j$

Property 1. If x_i and x_j are independent then $cov(x_i, x_j)=0$

Property 2. If x_i, x_j are dependent then $cov(x_i, x_j)\neq 0$

Property 3. Let $i=1$, $j=2$. $cov(x_1, x_2)=cov(x_2, x_1)$

Property 4. $cov(x_1, x_1)=Dx_1$

Property 5. $cov(x_1+c_1, x_2+c_2)=cov(x_1, x_2)$

Property 6. $cov(c_1x_1+c_2x_2, x_3)=c_1\,cov(x_1, x_3)+c_2\,cov(x_2, x_3)$

Let $X=(x_1, x_2,...,x_n)$ be a random vector.

$$\Sigma=\begin{pmatrix}\sigma_{11} & \sigma_{12}\ldots\sigma_{1n}\\ \ldots & \ldots\;\ldots\;\ldots\\ \sigma_{n1} & \sigma_{n2}\quad\sigma_{nm}\end{pmatrix}$$

The matrix Σ is called the matrix of covariation.

Property 7. The diagonal element of Σ , $\sigma_{ii}=Dx_i$

Definition. $|\Sigma|$, the determinant of Σ is called as common variance.

Let $Y_i=a_{i1}x_1+a_{i2}x_2+...+a_{in}x_n$; $i=1,n$

$A=(a_{ij})$ is the matrix of coefficient.

$$X=\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}\quad Y=\begin{pmatrix}y_1\\ y_2\\ y_3\end{pmatrix}\quad A=\begin{pmatrix}a_{11} & a_{12}\ldots & a_{1n}\\ a_{21} & a_{22}\ldots & a_{2n}\\ a_{n1} & a_{n2}\ldots & a_{nn}\end{pmatrix}$$

then $Y=A\cdot X$, $MY=A\cdot MX$

Let $H=(H_{ij})$ be the matrix of covariation of Y, then $H=A\Sigma A'$

Property 8. $D(\sum_{k=1}^{n} c_k x_k) = \sum_{k=1}^{n}\sum_{i=1}^{n} c_k c_i \sigma_{ki} \geq 0,$

Let $n=2$, $c_1=c_2=1$

Property 9. $D(c_1x_1)=c_1c_1\sigma_{11}=c_1^2 Dx_1$

Property 10. $D(x_1+x_2)=1\cdot1\sigma_{11}+1\cdot1\sigma_{12}+1\cdot1\sigma_{21}+1\cdot1\sigma_{22}=Dx_1+Dx_2+2\sigma_{12}$.

Property 11. $D(x_1-x_2)=\sigma_{11}-2\sigma_{12}+\sigma_{22}=Dx_1+Dx_2-2\sigma_{12}$

$D(x_1\pm x_2) = Dx_1+Dx_2\pm2\sigma_{12}$

Property 12. If x_1, x_2 are independent then $D(x_1\pm x_2)=Dx_1+Dx_2$ because $\sigma_{12}=0$

Let $c_k=c_l=1$; $k=1,2,...,n$. Then $D(\sum_{k=1}^{n} x_k) = \sum_{k=1}^{n}\sum_{l=1}^{n} \sigma_{kl}$ equals the sum of elements of the matrix of co-variation.

Property 13. If the components of random vector X $x_1, x_2,..., x_n$ are independent then the elements of Σ, which $i\neq j$ equal zero,

$$\Sigma = \begin{pmatrix} \sigma_{11} & 0 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & & \sigma_{nm} \end{pmatrix} \qquad \sigma_{ij} = \sigma_{x_j}^2 .$$

$D(\sum_{k=1}^{n} x_k) = \sum_{k=1}^{n} \sigma_{x_k}^2 = \sum_{k=1}^{n} Dx_k$ for independent x_i

Example. The quality of some detail defined by two parameters X, Y with the given law of distribution $Z=(X,Y)$

$x_i \setminus y_j$	0	0.1	0.2	0.3	P_i
5	0.2	0.1	0.05	0.05	0.4
6	0	0.15	0.15	0.1	0.4
7	0	0	0.1	0.1	0.2
P_j	0.2	0.25	0.3	0.25	$\sum_{i=1}^{3}\sum_{j=1}^{4} p_{ij} = 1$

Let $Z_1 = X - Y,$

$Z_2 = 3X - 2Y$

Find $DZ_1 = ?, DZ_2 = ?, Z = (Z_1, Z_2), H - ?$ Where H is the matrix of co-variation.

Solution: $MX = 5.8,\ MY = 0.16,\ M(XY) = \sum_{i=1}^{3}\sum_{j=1}^{4} x_i y_j p_{ij}$ =5·0·0.2 + 5·0.1·0.1+ 5·0.2·0.05 + 5·0.3·0.05+6·0·0+ ... +7·0.3·0.1=0.975

Let's calculate σ_{ij} the elements of $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$

$\sigma_{12}=\sigma_{21}=cov(X,Y)=M(XY)-MX\cdot MY=0.975-5\cdot 8\cdot 0.16=0.047$

$\sigma_{11}=DX=MX^2-(MX)^2=5^2\cdot 0.4+6^2\cdot 0.4+7^2\cdot 0.2-5\cdot 8^2=0.56,$

$\sigma_{22}=DY=MY^2-(MY)^2=0.037$

$$\Sigma = \begin{pmatrix} 0.56 & 0.047 \\ 0.047 & 0.037 \end{pmatrix}$$

$$DZ_1 = D(X-Y) = DX + DY - 2\sigma_{12} = \sum_{k=1}^{2}\sum_{l=1}^{2}\sigma_{kl}c_k c_l = 0.56 + 0.037 - 2\cdot 0.047 = 0.503;$$

$$DZ_2 = D(3X-2Y) =$$

$$\sum_{k=1}^{2}\sum_{l=1}^{2}\sigma_{kl}c_k c_l = 3\cdot(-3)\cdot 0.56 + 3\cdot(-2)\cdot 0.047 + (-2)\cdot 3\cdot 0.047 + (-2)\cdot(-3)\cdot 0.037 =$$

$$= 5.04 - 0.282 - 0.282 + 0.148 = 4.624.$$

For $H = A\Sigma A' = \begin{pmatrix} 1 & -1 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} 0.56 & 0.047 \\ 0.047 & 0.037 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ -1 & -2 \end{pmatrix}.$

Let $X=(X_1, X_2)$, where X_i is random variable with parameters $MX_i = a_i; \sigma_{x_i} = \sqrt{DX_i}$. We will transform as $Y_i = \frac{X_i - MX_i}{\sigma_{xi}},\quad MY_i = 0,\ DY_i = 1$

For $i = 1,2;\ Y = (Y_1, Y_2)$,

Let's calculate

$$cov(Y_1,Y_2) = cov\left(\frac{X_1 - MX_1}{\sigma_{x_i}}, \frac{X_2 - MX_2}{\sigma_{x_i}}\right) = cov\left(\frac{X_1 - MX_1, X_2 - MX_2}{\sigma_{x_1}\sigma_{x_2}}\right) =$$

$$= \frac{\text{cov}(X_1, X_2)}{\sigma_{x_1}\sigma_{x_2}}$$

We shall define $\rho_{x_1,x_2} = \frac{cov(X_1,X_2)}{\sigma_{x_1}\sigma_{x_2}}$ as a measure of stochastically dependents X_1, X_2 and this is called as a coefficient of correlation.

$$\rho_{x_1,x_2} = cov\left(\frac{x_1 - Mx_1}{\sigma_{x_1}}, \frac{x_2 - Mx_2}{\sigma_{x_2}}\right) = cov(Y_1, Y_2)$$

$$cov(x_1, x_2) = \rho_{x_1,x_2}\sigma_{x_1}\sigma_{x_2}$$

For independent random variable $\rho_{x_1,x_2}=0$, the inverse conclusion is not right.

Property 1. $\left|\rho_{x_1,x_2}\right|\le 1$

Proof: $Y_i=\dfrac{X_i-MX_i}{\sigma_{x_i}},\ i=1,2$

$D(Y_1\pm Y_2)=DY_1+DY_2\pm 2cov\left(Y_1,Y_2\right)=2\pm 2\rho_{x_1,x_2}\ge 0$

$1\pm\rho_{x_1,x_2}\ge 0,\ 1+\rho_{x_1,x_2}\ge 0,\ \ 1-\rho_{x_1,x_2}\ge 0,\ \left|\rho_{x_1,x_2}\right|\le 1$

Property 2. If x_1 and x_2 are linearly dependent: $x_2=ax_1+b$,

$\left|\rho_{x_1,x_2}\right|=1$necessary and enough condition, where $a\neq 0$, $\rho_{x_1,x_2}=1$, if $a>0$ and if $\rho_{x_1,x_2}=-1$, $a<0$.

Proof: (Necessary) Let $\rho_{x_1,x_2}=1,\ D\left(Y_1-Y_2\right)=2(1-\rho_{x_1,x_2})=0;$

It means that $Y_1-Y_2=C$

$M\left(Y_1-Y_2\right)=MC=C=MY_1-MY_2=0-0=0$

$Y_1-Y_2=0,\ Y_1=Y_2,\ \dfrac{X_1-MX_1}{\sigma_{x_1}}=\dfrac{X_2-MX_2}{\sigma_{x_2}}$

$x_2=ax_1+b$ where $a=\dfrac{\sigma_{x_1}}{\sigma_{x_2}},\ b=MX_2-\dfrac{\sigma_{x_1}}{\sigma_{x_2}}MX_1$ and $x_2=\dfrac{\sigma_{x_2}}{\sigma_{x_1}}x_1+(MX_2-\dfrac{\sigma_{x_2}}{\sigma_{x_1}}MX_1)$

For $\rho_{x_1,x_2}=-1,\quad D\left(Y_1,Y_2\right)=0;\ x_2=\dfrac{\sigma_{x_2}}{\sigma_{x_1}}x_1+MX_2+\dfrac{\sigma_{x_2}}{\sigma_{x_1}}MX_1$

It is enough that: $\rho_{x_1,ax_1+b}=\begin{cases}1, & a>0;\\ -1, & a<0.\end{cases}$

$$\rho_{x_1,x_2}=cov\left(Y_1,Y_2\right)=cov\left(Y_1,Y_1\frac{a}{|a|}\right)=cov\left(Y_1,\ Y_1\right)\frac{a}{|a|}=\frac{a}{|a|}=\begin{cases}1, & a>0;\\ -1, & a<0.\end{cases}$$

2.24 REGRESSION FUNCTION

Let Y and $ax+b$ be dependent. Consider $Y-\left(ax+b\right)$ in which *a,b, will take value min* $D\left(Y-ax-b\right)=?$

If $a_0=\dfrac{\rho_{xy}\sigma_y}{\sigma_x},\quad b_0=MY-\rho_{xy}\dfrac{\sigma_y}{\sigma_x}MX;$ Then $minD\left(Y-a_0x-b_0\right)=\sigma_Y^2(1-\rho_{xy}^2),$

If $\rho_{xy}\to 1$ *then* $D\left(Y-a_0x-b_0\right)\to 0$ and $Y=\rho_{xy}\dfrac{\sigma_y}{\sigma_x}X+MY-\rho_{xy}\dfrac{\sigma_y}{\sigma_x}MX$.

Let $X=(x_1,x_2,...,x_n)$ be a random vector.

$$\rho_{ij}=\rho_{ji}=\frac{cov(x_j,x_i)}{\sigma_{x_j}/\sigma_{x_i}}=\frac{cov(x_i,x_j)}{\sigma_{x_i}/\sigma_{x_j}}\quad,\ \rho_{ij}=\frac{\sigma_{ij}}{\sigma_{x_i}/\sigma_{x_j}}$$

$$R=\begin{pmatrix}\rho_{11}\ \rho_{12}\ \cdots\ \rho_{1n}\\ \ldots\ \ldots\ \ldots\ \ldots\\ \rho_{n1}\ \rho_{n2}\ \cdots\ \rho_{nn}\end{pmatrix}$$ WE shall call R as matrix of correlation.

Example.

Let $\Sigma=\begin{pmatrix}0.56 & 0.047\\ 0.047 & 0.037\end{pmatrix},\ R=?$

$$\rho_{11}=\rho_{22}=1,\ \rho_{12}=\frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}}=\frac{0.047}{\sqrt{0.56}\sqrt{0.037}}=0.33$$

$y=MY+\rho_{xy}\dfrac{\sigma_y}{\sigma_x}(x-MX)=0.84x-0.32.$ (regression function)

For normal random variable $X\sim N(a_1,\sigma_1),\ Y\sim N(a_2,\sigma_2)$

$$p_z(x,y)=\frac{1}{2\pi\sigma_1\sigma_2}\exp\left\{-\frac{1}{2(1-r^2)}\left[\left(\frac{x-a_1}{\sigma_1}\right)^2-2r\frac{x-a_1}{\sigma_1}\frac{y-a_2}{\sigma_2}+\left(\frac{y-a_2}{\sigma_2}\right)^2\right]\right\}$$

$(X,Y)=Z,\ X\sim N(a_1,\sigma_1),\ Y\sim N(a_2,\sigma_2)\ ,\ \rho_{xy}=1$

If $r=0$, X and Y are independent,

$$p(x,y)=\frac{1}{2\pi\sigma_1\sigma_2}e^{-\frac{1}{2}\left[\left(\frac{x-a_1}{\sigma_1}\right)^2+\left(\frac{y-a_2}{\sigma_2}\right)^2\right]}=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{1}{2}\left(\frac{x-a_1}{\sigma_1}\right)^2}\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2}\left(\frac{y-a_2}{\sigma_2}\right)^2}=p_X(x)p_Y(y)$$

2.25 MARKOV CHAIN

Assume that we have $S_1,S_2,...,S_n$ - position.

Let P_0—initial distribution and P—matrix of position (transition matrix).

$$P_0=\begin{pmatrix}p_1\\ \ldots\\ p_n\end{pmatrix}\quad P=\begin{pmatrix}p_{11}\ p_{12}\ \cdots\ p_{1n}\\ \ldots\ \ldots\ \ldots\ \ldots\\ p_{n1}\ p_{n2}\ \cdots\ p_{nn}\end{pmatrix}$$

$\sum_j P_{ij}=1,\ P\{X=x_i,Y=y_j\}.$

Marcov chains will use when we study random processes.

PART III. STATISTICAL ESTIMATION OF PARAMETERS θ IN F (X, θ) DISTRIBUTIONS

3.1 POINT ESTIMATION

Definition 1. *A* numerical measure of population is called a population parameter.

Definition 2. An estimator of a population parameter is a sample statistics used to estimate the parameter.

An estimate of the parameter is a particular numerical value of the estimator obtained by sampling.

Definition 3. When a single value is used as an estimator, the estimator is called a point estimation of population parameter.

Example. $\overline{X}$ a sample mean, is estimator of population μ.

Definition 4. An interval estimate exists and an estimator constitutes an interval of numbers rather than a single number.

An interval estimate is an interval which belongs to the unknown population parameter.

Properties of estimators $\theta_n = f(x_1, \ldots, x_n)$

Definition 5. An estimator is said to be **unbiased**, if its expected value is equal to the population parameter which estimates. That is $M\theta_n = \theta$.

Definition 6. Any systematic deviation of estimator from the parameter of interest is called a bias.

Definition 7. An estimator is called an **efficient** if it has a relatively small variance than the other estimators. Let θ_n and T_n be estimators . If $D\theta_n < DT_n$ then θ_n is an effective estimator.

Definition 8. An estimator is said to be **consistent** if the sample size increases its probability of being close to 1 and the parameter θ_n tends to parameter θ.

That is $\lim\limits_{n\to\infty} P\{|\theta_n - \theta| < \varepsilon\} = 1$.

Theorem 1. Assume that we are given $x_1, \ldots, x_n$, the sample size equals *n*. If $Mx_i = \mu$, then $\overline{X}$-arithmetical mean is unbiased estimator for μ.

Theorem 2. $\overline{X}$ is consistent estimator of μ. $P\{|\overline{x}-\mu|<\varepsilon\}\to 1$ when $n\to\infty$

Proof. From the law of large numbers $P(|\overline{x}-\mu|<\varepsilon)>1-D\overline{x}/\varepsilon^2$, $D\overline{x}=\sigma^2/n$.

Theorem 3. If $x_i \in N(\mu, \sigma^2)$, $\overline{X}$ is unbiased , efficient and consistent estimator for μ .

Theorem 4. If $Mx_i=\mu$, $Dx_i=\sigma^2$, then $S^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2$ is not unbiased estimator for σ^2 and $MS^2=\frac{n-1}{n}\sigma^2$

Proof:.

$$S^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu+\mu-\overline{x})^2=\frac{1}{n}\sum_{i=1}^{n}((x_i-\mu)-(\overline{x}-\mu))^2=$$

$$\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2-\frac{2}{n}(\overline{x}-\mu)\sum_{i=1}^{n}(x_i-\mu)+\frac{n}{n}(\overline{x}-\mu)^2;-\frac{2}{n}(\overline{x}-\mu)\sum_{i=1}^{n}(x_i-\mu)\Rightarrow$$

$$\Rightarrow n\overline{x}=\sum_{i=1}^{n}x_i\Rightarrow 2(\overline{x}-\mu)^2$$

$$S^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2-(\overline{x}-\mu)^2, MS^2=\sigma^2-\frac{\sigma^2}{n}$$

$$M(\overline{x}-\mu)^2=D\overline{X}=\frac{\sigma^2}{n}, MS^2=\frac{n-1}{n}\sigma^2$$

Theorem 6. $\hat{S}^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2$ is unbiased estimator for σ^2, where $\hat{S}^2=\frac{n}{n-1}S^2$;

Theorem 7. $S_*^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$ is unbiased, consistent, efficient estimator for σ^2.

Theorem 8. If X is a random variable with parameters (μ, σ^2) and $x_1, x_2, ..., x_n$ are independent trials of X and $Mx_i=\mu$, $Dx_i=\sigma^2$, then the arithmetical mean $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}x_i \in N(\mu, \frac{\sigma^2}{n})$.

Consequence 1. $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \in N(0,1)$

$$M\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right)=\frac{\sqrt{n}}{\sigma}M(\overline{X}-\mu)=\frac{\sqrt{n}}{\sigma}(M\overline{X}-\mu)=0$$

$$D\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right)=\frac{n}{\sigma^2}(D(\overline{X})-D\mu)=\frac{n}{\sigma^2}D\overline{X}=\frac{n}{\sigma^2}\frac{\sigma^2}{n}=1$$

$$z=\frac{\overline{X}-\mu}{\sigma}\sqrt{n}\in N(0,1)$$

If σ^2 is unknown, instead of σ we use

$$\hat{S}^2=\frac{1}{n-1}\sum(x_i-\overline{X})^2$$

Remark:

I. $Z=\frac{\overline{X}-\mu}{\hat{S}}\sqrt{n}$ – Student distribution or *T*- distribution

For example:

$t_{n,\alpha}$=?, if n=13, α=0.05, $t_{13,\,0.95}$=2.18

II. The random variable $\frac{n\hat{S}^2}{\sigma^2}$ *is* χ^2 distribution with degree of freedom $k=(n-1)$.

3.2 INTERVAL ESTIMATION FOR μ

A) (σ is known) . Let $P[\theta_n^{(1)}<\theta<\theta_n^{(2)}]=1-\alpha$

Here, $[\theta_n^{(1)},\theta_n^{(2)}]$ – is called an interval of confidence.

$P=1-\alpha$, α is the level of confidence.

Let $X\in(\mu,\sigma^2)$.

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}x_i;\ \overline{X}\in N(\mu,\sigma^2);\ \frac{\overline{X}-\mu}{\sigma}\sqrt{n}\in N(0,1)$$

$$P\left(\left|\frac{\overline{X}-\mu}{\sigma}\sqrt{n}\right|<z_n\right)=\Phi(z)$$

$$P\left(\overline{X}-z_p\frac{\sigma}{\sqrt{n}}<\mu<\overline{X}+z_p\frac{\sigma}{\sqrt{n}}\right)=\Phi(z)=1-\alpha$$

$$\mu\in\left[\overline{X}-z_p\frac{\sigma}{\sqrt{n}};\ \overline{X}+z_p\frac{\sigma}{\sqrt{n}}\right]=[\theta_n^{(1)},\theta_n^{(2)}]$$

Example. $X\in(\mu,\sigma)$, $\sigma=2$, $n=16$, $p=1-\alpha=0.95$

$z_{0.95}$=1.96 from table of $\Phi(x)$.

$$\Delta=\left(z_p\frac{\sigma}{\sqrt{n}}\right)=1.96\frac{2}{\sqrt{16}}=0.98$$

$\mu \in [\overline{X} - 0.98, \overline{X} + 0.98]$ with probability of 0.95 if $\overline{X}$ =4.1 in our example

$\mu \in [3.12,\ 5.08];\quad 3.12 < \mu < 5.08.$

B) (σ is unknown). If σ is unknown then $\sigma = \hat{S};\ t = \frac{\overline{X} - \mu}{\hat{S}}\sqrt{n}$

$$P\left(\frac{|\overline{X} - \mu|}{\hat{S}}\sqrt{n} < t_{n,p} \right) = 1 - \alpha$$

$$P\left(|\overline{X} - \mu| < t_{n,p}\frac{\hat{S}}{\sqrt{n}} \right) = 1 - \alpha$$

$$P\left(\overline{X} - t_{n,p}\frac{\hat{S}}{\sqrt{n}} < \mu < \overline{X} + t_{n,p}\frac{\hat{S}}{\sqrt{n}} \right) = 1 - \alpha$$

$$\mu \in \left[\overline{X} - t_{n,p}\frac{\hat{S}}{\sqrt{n}};\ \overline{X} + t_{n,p}\frac{\hat{S}}{\sqrt{n}} \right] = [\theta_n^{(1)}, \theta_n^{(2)}]$$

Example. If *n*=9, *p*=0.95, *S*=3, $\overline{X}$ =6 than the interval will be

$6 - \frac{3}{\sqrt{9}}1.96 < \mu < 6 + \frac{3}{\sqrt{9}}1.96 \Rightarrow$ 4.04<μ<7.96.

From Student distribution, when *p*=0.95, *n*=9, $t_{8,0.95} = 2.31$.

$\overline{X} - t_{n,p}\frac{\hat{S}}{\sqrt{n}} < \mu < \overline{X} + t_{n,p}\frac{\hat{S}}{\sqrt{n}},\ 6 - \frac{3}{\sqrt{9}}2{,}31 < \mu < 6 + \frac{3}{\sqrt{9}}2.31$ the interval will be 3.69 <μ<8.31 .

3.3 INTERVAL ESTIMATION FOR σ^2

If $\frac{n\hat{S}^2}{\sigma^2}$ has χ^2 distribution with degree of freedom *k*=*n*-1.

$$P\left(\frac{\sqrt{n}\hat{S}}{X_2} < \sigma < \frac{\sqrt{n}\hat{S}}{X_1} \right) = 1 - \alpha$$

$$P(\chi^2 < \chi_1^2) = P(\chi^2 > \chi_2^2) = \frac{a}{2}$$

$$S^2 = 10, p = 0.96, \text{ If } n = 20,\ p_2 = \frac{a}{2},$$

$$p_1 = 1 - \frac{a}{2},\ a = 0.04.$$

$$p_2 = 0.02,\ k = n - 1 = 19 \rightarrow \chi_2^2 = 33.7$$

$p_1 = 0.98,\ k = 19,\ \chi_1^2 = 8,6$

$\sqrt{5.935} < \sigma < \sqrt{23.256} \rightarrow 2.43 < \sigma < 4.82$

Appendix

TABLE № 1. THE VALUE OF FUNCTION e^{-x}

x	*exp(-x)*	*x*	*exp(-x)*	*x*	*exp(-x)*	*x*	*exp(-x)*
0,00	1,000	0,40	0,670	0,80	0,449	3,0	0,0498
0,02	0,980	0,42	0,657	0,82	0,440	3,2	0,0408
0,04	0,961	0,44	0,644	0,84	0,432	3,4	0,0334
0,06	0,942	0,46	0,631	0,86	0,423	3,6	0,0273
0,08	0,923	0,48	0,619	0,88	0,415	3,8	0,0224
0,10	0,905	0,50	0,607	0,90	0,407	4,0	0,0183
0,12	0,887	0,52	0,595	0,92	0,399	4,2	0,0150
0,14	0,869	0,54	0,583	0,94	0,391	4,4	0,0123
0,16	0,852	0,56	0,571	0,96	0,383	4,6	0,0101
0,18	0,835	0,58	0,560	0,98	0,375	4,8	0,0082
0,20	0,819	0,60	0,549	1,00	0,368	5,0	0,0067
0,22	0,803	0,62	0,538	1,20	0,301	5,2	0,0055
0,24	0,787	0,64	0,527	1,40	0,247	5,4	0,0045
0,26	0,771	0,66	0,517	1,60	0,202	5,6	0,0037
0,28	0,756	0,68	0,507	1,80	0,165	5,8	0,0030
0,30	0,741	0,70	0,497	2,00	0,135	6,0	0,0025
0,32	0,726	0,72	0,487	2,20	0,111	6,2	0,0020
0,34	0,712	0,74	0,477	2,40	0,091	6,4	0,0017
0,36	0,698	0,76	0,468	2,60	0,074	6,6	0,0014
0,38	0,684	0,78	0,458	2,80	0,061	6,8	0,0011
0,40	0,670	0,80	0,449	3,00	0,050	7,0	0,0009

TABLE № 2. THE VALUE OF FUNCTION $\frac{\lambda^m e^{-\lambda}}{m!}$ (POISSON)

m	*λ=0.1*	*λ=0.2*	*λ=0.3*	*λ=0.4*	*λ=0.5*	*λ=0.6*
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488
1	0.0905	0.1638	0.2222	0.2681	0.3033	0.3293

2	0.0045	0.0164	0.0333	0.0536	0.0758	0.0988
3	0.0002	0.0011	0.0033	0.0072	0.0126	0.0198
4		0.0001	0.0002	0.0007	0.0016	0.0030
5				0.0001	0.0002	0.0004

m	*λ=0.7*	*λ=0.8*	*λ=0.9*	*λ=1.0*	*λ=2.0*	*λ=3.0*
0	0.4966	0.4493	0.4066	0.3679	0.1353	0.0498
1	0.3476	0.3595	0.3659	0.3679	0.2707	0.1494
2	0.1217	0.1438	0.1647	0.1879	0.2707	0.2240
3	0.0284	0.0383	0.0494	0.0613	0.1804	0.2240
4	0.0050	0.0077	0.0111	0.0153	0.0902	0.1680
5	0.0007	0.0012	0.0020	0.0031	0.0361	0.1008
6	0.0001	0.0002	0.0003	0.0005	0.0120	0.0504
7				0.0001	0.0034	0.0216
8					0.0009	0.0081
9					0.0002	0.0027
10						0.0008
11						0.0002
12						0.0001

m	*λ=4.0*	*λ=5.0*	*λ=6.0*	*λ=7.0*	*λ=8.0*	*λ=9.0*
0	0.0183	0.0067	0.0025	0.0009	0.0003	0.0001
1	0.0733	0.0337	0.0149	0.0064	0.0027	0.0011
2	0.1465	0.0842	0.0446	0.0223	0.0107	0.0050
3	0.1954	0.1404	0.0892	0.0521	0.0286	0.0150
4	0.1954	0.1755	0.1339	0.0912	0.0572	0.0337
5	0.1563	0.1755	0.1606	0.1277	0.0916	0.0607
6	0.1042	0.1462	0.1606	0.1490	0.1221	0.0911
7	0.0595	0.1044	0.1377	0.1490	0.1396	0.1171
8	0.0298	0.0653	0.1033	0.1304	0.13.96	0.1318
9	0.0132	0.0363	0.0688	0.1014	0.1241	0.1318
10	0.0053	0.0181	0.0413	0.0710	0.0993	0.1186

11	0.0019	0.0082	0.0225	0.0452	0.0722	0.0970
12	0.0006	0.0034	0.0113	0.0264	0.0481	0.0728
13	0.0002	0.0013	0.0052	0.0142	0.0296	0.0504
14	0.0001	0.0005	0.0022	0.0071	0.0169	0.0324
15		0.0002	0.0009	0.0033	0.0090	0.0194
16		0.0001	0.0003	0.0015	0.0045	0.0109
17			0.0001	0.0006	0.0021	0.0058

TABLE №3. THE VALUE OF LAPLACE FUNCTION

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

x	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	0,3970	0,3965	0,3961	0,3956	0,3951	0,3945	0,3939	0,3932	0,3925	0,3918
0,2	0,3910	0,3902	0,3894	0,3885	0,3876	0,3867	0,3857	0,3847	0,3836	0,3825
0,3	0,3814	0,3802	0,3790	0,3778	0,3765	0,3752	0,3739	0,3725	0,3712	0,3697
0,4	0,3683	0,3668	0,3653	0,3637	0,3621	0,3605	0,3589	0,3572	0,3555	0,3538
0,5	0,3521	0,3503	0,3485	0,3467	0,3448	0,3429	0,3410	0,3391	0,3372	0,3352
0,6	0,3332	0,3312	0,3292	0,3271	0,3251	0,3230	0,3209	0,3187	0,3166	0,3144
0,7	0,3123	0,3101	0,3079	0,3056	0,3034	0,3011	0,2989	0,2966	0,2943	0,2920
0,8	0,2897	0,2874	0,2850	0,2827	0,2803	0,2780	0,2756	0,2732	0,2709	0,2685
0,9	0,2661	0,2637	0,2613	0,2589	0,2565	0,2541	0,2516	0,2492	0,2468	0,2444
1,0	0,2420	0,2396	0,2371	0,2347	0,2323	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	0,2179	0,2155	0,2131	0,2107	0,2083	0,2059	0,2036	0,2012	0,1989	0,1965
1,2	0,1942	0,1919	0,1895	0,1872	0,1849	0,1826	0,1804	0,1781	0,1758	0,1736
1,3	0,1714	0,1691	0,1669	0,1647	0,1626	0,1604	0,1582	0,1561	0,1539	0,1518
1,4	0,1497	0,1476	0,1456	0,1435	0,1415	0,1394	0,1374	0,1354	0,1334	0,1315
1,5	0,1295	0,1276	0,1257	0,1238	0,1219	0,1200	0,1182	0,1163	0,1145	0,1127
1,6	0,1109	0,1092	0,1074	0,1057	0,1040	0,1023	0,1006	0,0989	0,0973	0,0957
1,7	0,0940	0,0925	0,0909	0,0893	0,0878	0,0863	0,0848	0,0833	0,0818	0,0804
1,8	0,0790	0,0775	0,0761	0,0748	0,0734	0,0721	0,0707	0,0694	0,0681	0,0669
1,9	0,0656	0,0644	0,0632	0,0620	0,0608	0,0596	0,0584	0,0573	0,0562	0,0551

x	0	1	2	3	4	5	6	7	8	9
2,0	0,0540	0,0529	0,0519	0,0508	0,0498	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0,0440	0,0431	0,0422	0,0413	0,0404	0,0395	0,0387	0,0379	0,0371	0,0363
2,2	0,0355	0,0347	0,0339	0,0332	0,0325	0,0317	0,0310	0,0303	0,0297	0,0290
2,3	0,0283	0,0277	0,0270	0,0264	0,0258	0,0252	0,0246	0,0241	0,0235	0,0229
2,4	0,0224	0,0219	0,0213	0,0208	0,0203	0,0198	0,0194	0,0189	0,0184	0,0180
2,5	0,0175	0,0171	0,0167	0,0163	0,0158	0,0154	0,0151	0,0147	0,0143	0,0139
2,6	0,0136	0,0132	0,0129	0,0126	0,0122	0,0119	0,0116	0,0113	0,0110	0,0107
2,7	0,0104	0,0101	0,0099	0,0096	0,0093	0,0091	0,0088	0,0086	0,0084	0,0081
2,8	0,0079	0,0077	0,0075	0,0073	0,0071	0,0069	0,0067	0,0065	0,0063	0,0061
2,9	0,0060	0,0058	0,0056	0,0055	0,0053	0,0051	0,0050	0,0048	0,0047	0,0046
3,0	0,0044	0,0043	0,0042	0,0040	0,0039	0,0038	0,0037	0,0036	0,0035	0,0034
3,1	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026	0,0025	0,0025
3,2	0,0024	0,0023	0,0022	0,0022	0,0021	0,0020	0,0020	0,0019	0,0018	0,0018
3,3	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014	0,0013	0,0013
3,4	0,0012	0,0012	0,0012	0,0011	0,0011	0,0010	0,0010	0,0010	0,0009	0,0009
3,5	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007	0,0007	0,0007	0,0006
3,6	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005	0,0005	0,0005	0,0005	0,0004
3,7	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003	0,0003	0,0003	0,0003
3,8	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002	0,0002	0,0002	0,0002	0,0002
3,9	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0001	0,0001

$\varphi(x) = \varphi(-x)$;

for $x \geq 4$: $\varphi(x) = 0$.

TABLE №4. THE VALUE OF LAPLACE INTEGRAL FUNCTION

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt .$$

x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
0,00	0,0000	0,32	0,1255	0,64	0,2389	0,96	0,3315
0,01	0,0040	0,33	0,1293	0,65	0,2422	0,97	0,3340

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,02	0,0080	0,34	0,1331	0,66	0,2454	0,98	0,3365
0,03	0,0120	0,35	0,1368	0,67	0,2486	0,99	0,3389
0,04	0,0160	0,36	0,1406	0,68	0,2517	1,00	0,3413
0,05	0,0199	0,37	0,1443	0,69	0,2549	1,01	0,3438
0,06	0,0239	0,38	0,1480	0,70	0,2580	1,02	0,3461
0,07	0,0279	0,39	0,1517	0,71	0,2611	1,03	0,3485
0,08	0,0319	0,40	0,1554	0,72	0,2642	1,04	0,3508
0,09	0,0359	0,41	0,1591	0,73	0,2673	1,05	0,3531
0,10	0,0398	0,42	0,1628	0,74	0,2704	1,06	0,3554
0,11	0,0438	0,43	0,1664	0,75	0,2734	1,07	0,3577
0,12	0,0478	0,44	0,1700	0,76	0,2764	1,08	0,3599
0,13	0,0517	0,45	0,1736	0,77	0,2794	1,09	0,3621
0,14	0,0557	0,46	0,1772	0,78	0,2823	1,10	0,3643
0,15	0,0596	0,47	0,1808	0,79	0,2852	1,11	0,3665
0,16	0,0636	0,48	0,1844	0,80	0,2881	1,12	0,3686
0,17	0,0675	0,49	0,1879	0,81	0,2910	1,13	0,3708
0,18	0,0714	0,50	0,1915	0,82	0,2939	1,14	0,3729
0,19	0,0753	0,51	0,1950	0,83	0,2967	1,15	0,3749
0,20	0,0793	0,52	0,1985	0,84	0,2995	1,16	0,3770
0,21	0,0832	0,53	0,2019	0,85	0,3023	1,17	0,3790
0,22	0,0871	0,54	0,2054	0,86	0,3051	1,18	0,3810
0,23	0,0910	0,55	0,2088	0,87	0,3078	1,19	0,3830
0,24	0,0948	0,56	0,2123	0,88	0,3106	1,20	0,3849
0,25	0,0987	0,57	0,2157	0,89	0,3133	1,21	0,3869
0,26	0,1026	0,58	0,2190	0,90	0,3159	1,22	0,3888
0,27	0,1064	0,59	0,2224	0,91	0,3186	1,23	0,3907
0,28	0,1103	0,60	0,2257	0,92	0,3212	1,24	0,3925
0,29	0,1141	0,61	0,2291	0,93	0,3238	1,25	0,3944
0,30	0,1179	0,62	0,2324	0,94	0,3264	1,26	0,3962
0,31	0,1217	0,63	0,2357	0,95	0,3289	1,27	0,3980

$\Phi(-x) = -\Phi(x)$;

for $x > 5$: $\Phi(x) = 0{,}5$.

x	*Φ*(*x*)	*x*	*Φ*(*x*)	*x*	*Φ*(*x*)	*x*	*Φ*(*x*)
1,28	0,3997	1,61	0,4463	1,94	0,4738	2,54	0,4945
1,29	0,4015	1,62	0,4474	1,95	0,4744	2,56	0,4948
1,30	0,4032	1,63	0,4484	1,96	0,4750	2,58	0,4951
1,31	0,4049	1,64	0,4495	1,97	0,4756	2,60	0,4953
1,32	0,4066	1,65	0,4505	1,98	0,4761	2,62	0,4956
1,33	0,4082	1,66	0,4515	1,99	0,4767	2,64	0,4959
1,34	0,4099	1,67	0,4525	2,00	0,4772	2,66	0,4961
1,35	0,4115	1,68	0,4535	2,02	0,4783	2,68	0,4963
1,36	0,4131	1,69	0,4545	2,04	0,4793	2,70	0,4965
1,37	0,4147	1,70	0,4554	2,06	0,4803	2,72	0,4967
1,38	0,4162	1,71	0,4564	2,08	0,4812	2,74	0,4969
1,39	0,4177	1,72	0,4573	2,10	0,4821	2,76	0,4971
1,40	0,4192	1,73	0,4582	2,12	0,4830	2,78	0,4973
1,41	0,4207	1,74	0,4591	2,14	0,4838	2,80	0,4974
1,42	0,4222	1,75	0,4599	2,16	0,4846	2,82	0,4976
1,43	0,4236	1,76	0,4608	2,18	0,4854	2,84	0,4977
1,44	0,4251	1,77	0,4616	2,20	0,4861	2,86	0,4979
1,45	0,4265	1,78	0,4625	2,22	0,4868	2,88	0,4980
1,46	0,4279	1,79	0,4633	2,24	0,4875	2,90	0,4981
1,47	0,4292	1,80	0,4641	2,26	0,4881	2,92	0,4982
1,48	0,4306	1,81	0,4649	2,28	0,4887	2,94	0,4984
1,49	0,4319	1,82	0,4656	2,30	0,4893	2,96	0,4985
1,50	0,4332	1,83	0,4664	2,32	0,4898	2,98	0,4986
1,51	0,4345	1,84	0,4671	2,34	0,4904	3,00	0,49865
1,52	0,4357	1,85	0,4678	2,36	0,4909	3,20	0,49931
1,53	0,4370	1,86	0,4686	2,38	0,4913	3,40	0,49966
1,54	0,4382	1,87	0,4693	2,40	0,4918	3,60	0,499841
1,55	0,4394	1,88	0,4699	2,42	0,4922	3,80	0,499928
1,56	0,4406	1,89	0,4706	2,44	0,4927	4,00	0,499968
1,57	0,4418	1,90	0,4713	2,46	0,4931	4,25	0,499989
1,58	0,4429	1,91	0,4719	2,48	0,4934	4,50	0,499997

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,59	0,4441	1,92	0,4726	2,50	0,4938	4,75	0,499999
1,60	0,4452	1,93	0,4732	2,52	0,4941	5,00	0,500000

for $x > 5$: $\Phi(x) = 0,5$.

TABLE №5. STUDENT'S T- DISTRIBUTION $t_\gamma = t(\gamma, n)$

n \ γ	0,90	0,95	0,99	0,999	n \ γ	0,90	0,95	0,99	0,999
5	2,131	2,776	4,604	8,61	20	1,729	2,093	2,861	3,883
6	2,015	2,570	4,032	6,86	25	1,711	2,064	2,797	3,745
7	1,943	2,446	3,707	5,96	30	1,699	0,045	2,756	3,659
8	1,894	2,364	3,499	5,41	35	1,688	2,032	2,729	3,600
9	1,859	2,306	3,355	5,04	40	1,683	2,023	2,708	4,558
10	1,833	2,262	3,249	4,78	45	1,679	2,016	2,692	3,527
11	1,812	2,228	3,169	4,59	50	1,675	2,009	2,679	3,502
12	1,795	2,201	3,106	4,44	60	1,671	2,001	2,662	3,464
13	1,782	2,178	3,054	4,32	70	1,666	1,996	2,649	3,439
14	1,770	2,160	3,012	4,22	80	1,664	1,991	2,640	3,418
15	1,761	2,144	2,976	4,14	90	1,662	1,987	2,633	3,403
16	1,753	2,131	2,946	4,07	100	1,660	1,984	2,627	3,392
17	1,745	2,119	2,921	4,02	120	1,657	1,980	2,617	3,374
17	1,739	2,109	2,898	3,97	∞	1,645	1,960	2,576	3,291
19	1,734	2,101	2,878	3,92					

TABLE № 6. STUDENT DISTRIBUTION

The number of freeedom k	The level of confedience α					
	0,1	0,05	0,02	0,01	0,002	0,001
1	6,31	12,71	31,82	63,66	318,29	636,58

2	2,92	4,30	6,96	9,92	22,33	31,60
3	2,35	3,18	4,54	5,84	10,21	12,92
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,02	2,57	3,36	4,03	5,89	6,87
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,41
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,02	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,97
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,72	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,50	3,79
23	1,71	2,07	2,50	2,81	3,48	3,77
24	1,71	2,06	2,49	2,80	3,47	3,75
25	1,71	2,06	2,49	2,79	3,45	3,73
26	1,71	2,06	2,48	2,78	3,43	3,71
27	1,70	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,47	2,76	3,41	3,67
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,16	3,37
∞	1,64	1,96	2,33	2,58	3,09	3,29
	0,05	0,025	0,01	0,005	0,001	0,0005
	The level of confidience α					

Table №7. F- Distribution (Fisher)

(k_1-degree of freedom of numerator,
k_2 - degree of freedom of denominator)

k2	The level of confidience α =0,01											
	k1											
	1	2	3	4	5	6	7	8	9	10	11	12
1	4052	4999	5403	5625	5764	5859	5928	5981	6022	6056	6083	6107
2	98,50	99,00	99,16	99,25	99,30	99,33	99,36	99,38	99,39	99,40	99,41	99,42
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,13	27,05
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,45	14,37
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,96	9,89
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,54	6,47
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,73	5,67
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,18	5,11
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,77	4,71
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,46	4,40
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,22	4,16
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	4,02	3,96
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,86	3,80
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73	3,67
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,62	3,55
17	8,40	6,11	5,19	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,52	3,46

k2	The level of confidience α =0,05											
	k1											
	1	2	3	4	5	6	7	8	9	10	11	12
1	161	200	216	225	230	234	237	239	241	242	243	244
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,40	19,41
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76	8,74
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,94	5,91

5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,70	4,68
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60	3,57
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31	3,28
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10	3,07
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82	2,79
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,72	2,69
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,63	2,60
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,57	2,53
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,51	2,48
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,46	2,42
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,41	2,38

Table №8. χ^2-Chi Square Distribution

The value $\chi^2_{\alpha,k}$ defined from condition $P\left\{\chi^2_k > \chi^2_{\alpha.k}\right\} = \alpha$, where χ^2_k Chi Square Distribution with nuber of freedom k

The number of freedom k	The level of confidience a					
	0,01	0,025	0,05	0,95	0,975	0,99
1	6,635	5,024	3,841	0,00393	0,00098	0,00016
2	9,210	7,378	5,991	0,10259	0,05064	0,02010
3	11,345	9,348	7,815	0,35185	0,21579	0,11483
4	13,277	11,143	9,488	0,71072	0,48442	0,29711
5	15,086	12,832	11,070	1,145	0,831	0,554
6	16,812	14,449	12,592	1,635	1,237	0,872
7	18,475	16,013	14,067	2,167	1,690	1,239
8	20,090	17,535	15,507	2,733	2,180	1,647
9	21,666	19,023	16,919	3,325	2,700	2,088

10	23,209	20,483	18,307	3,940	3,247	2,558
11	24,725	21,920	19,675	4,575	3,816	3,053
12	26,217	23,337	21,026	5,226	4,404	3,571
13	27,688	24,736	22,362	5,892	5,009	4,107
14	29,141	26,119	23,685	6,571	5,629	4,660
15	30,578	27,488	24,996	7,261	6,262	5,229
16	32,000	28,845	26,296	7,962	6,908	5,812
17	33,409	30,191	27,587	8,672	7,564	6,408
18	34,805	31,526	28,869	9,390	8,231	7,015
19	36,191	32,852	30,144	10,117	8,907	7,633
20	37,566	34,170	31,410	10,851	9,591	8,260
21	38,932	35,479	32,671	11,591	10,283	8,897
22	40,289	36,781	33,924	12,338	10,982	9,542
23	41,638	38,076	35,172	13,091	11,689	10,196
24	42,980	39,364	36,415	13,848	12,401	10,856
25	44,314	40,646	37,652	14,611	13,120	11,524
26	45,642	41,923	38,885	15,379	13,844	12,198
27	46,963	43,195	40,113	16,151	14,573	12,878
28	48,278	44,461	41,337	16,928	15,308	13,565
29	49,588	45,722	42,557	17,708	16,047	14,256
30	50,892	46,979	43,773	18,493	16,791	14,953

Selected literature

1. Rasulov A.S., Theory of probability and mathematical statistics, UWED press, 2002.
2. Gmurman V.E. Theory of probability and mathematical statistics. Moscow, «Visshaya shkola» , 2008.
3. Sirajiddinov S.X., Mamàtov M.M Theory of probability and mathematical statistics. Tashkent, «Ukituvchi» 1980.
4. Amir D. Aczel ., Complete Business Statistics. IRWIN. Homewood, IL 60430, Boston, MA 02116, 1989.
5. John E. Hank, Arthur G. Reitsch. Understanding Business Statistics.
   Boston, 1994.
6. Masson Lind. Statistical techniques in business and economics. IRWIN. Homewood, IL 60430, Boston, MA 02116, 1990.
7. Brawn D. Applied Statistics. Springers-Kerlog. 1987.
8. Groebner D., Essential of Business Statistics. London, 1987.
9. Anderson D., ,D.Swreney,T. Williams., Quantitative methods for business., West Publ. Company.,1992.
10. Zamkov O., Mathematical mehtods in economy., M.,Moskîw., 1999.
11. Zaysev I.A. Higer Mathemetics., M., Visshay skola., 1998.

C O N T E N T S

THE UNIVERSITY OF WORLD ECONOMY AND DIPLOMACY

INTRODUCTION TO THE THEORY OF PROBABILITY AND MATHEMATICAL STATISTICS

Author Prof. Rasulov A.S.

Reviewer: Associate prof. Bakoev M.T., Muxamedov A.K.

Responsibility is upon the author of this aid for the quality of preparing publishing materials, brought facts and known terminology as well as other given allowance that not forbidden for the open publishing.

Approved by teaching and methodological council UWED, protocol of meeting № 8, 5 of December 2013.

Format 84x108 1/32. Volume 4.5
Amount of copies 100.

Duplicated by the printing house of the University of World Economy and Diplomacy.
Tashkent – 100077, Mustaqillik avenue.,54