This special issue is dedicated to the 10th anniversary of the University of World Economy and Diplomacy

1992-2002

The Ministry of Foreign Affairs of Republic of Uzbekistan

University of World Economy and Diplomacy

Prof. Rasulov A.S.

LECTURE ON THEORY OF PROBABILITY AND MATHEMATICAL STATISTICS

Tashkent - 2002

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.

1. 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

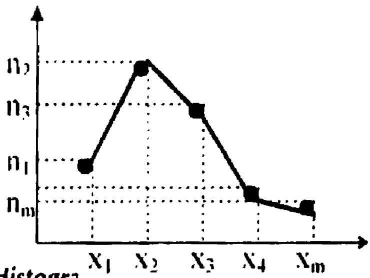
4-times 6-times 12-times 18times

Let $n_1, n_2, n_3, \ldots n_m \rightarrow$ Frequency of data, $n = \sum_{i=1}^{m} n_i$, m=6. $n_i/n \rightarrow$ Relative Frequency.

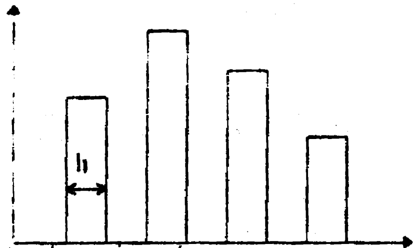
X_i	n_i	Cumulative Frequency Σ	Relative Frequency	Relative Cumulative Frequency
1	4	4	0.04	0.04
2	6	10	0.06	0.10
3	12	22	0.12	0.22
4	16	38	0.16	0.38
5	44	82	0.44	0.82
6	18	100	0.18	1.00

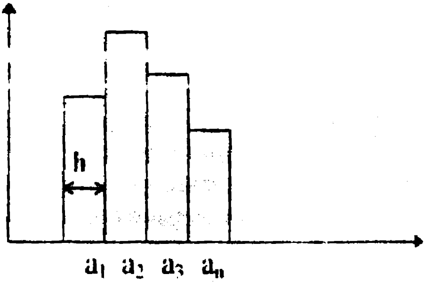
1.2 GRAPHICAL ILLUSTRATION OF DATA

1. *Polygon* *(relative polygon)*



2. *Histogra*





Sterdgess formula

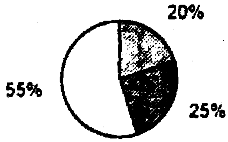
$$h=(x_{max}-x_{min})/(1+3.32\cdot lgN)$$

Where, h-length of interval

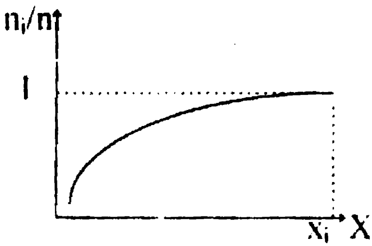
$x_{min}-h/2=a_1$; $a_2=a_1+h$; $a_3=a_2+h=a_1+2h$;... $a_n=a_1+(n-1)h$

3. ***Stem and leaf displays***

4. ***Pie chart***



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



1.3 MAIN NUMERICAL CHARACTERISTICS OF POPULATION

X : $x_1, x_2, x_3, \dots x_n$;   n - size of population

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

2) M_d = *mode* =max n_i=x_i; $1 \le i \le m$

3) *The (arithmetical) mean*

$$(1/n)\sum_{i=1}^{m} x_i n_i = \overline{X} \qquad n=\sum_{i=1}^{m} n_i$$

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

$$S^2= (\Sigma (x_i-\bar{x})^2 n_i)/(n-1)$$

where $\bar{x}$ is the *arithmetical* mean

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

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

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

$$C = 100(S/x)$$

7) *The Geometric Mean*

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

8) Harmonic mean

$$X_{harmonic} = n/(\sum_{i=1}^{n} 1/x_i)$$

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

1.4 INITIAL MOMENT OF A SAMPLE

Initial moment q order of a sample is

$$V_q = \Sigma (x_i^q n_i) / n$$

where, $v_0 = 1$, $v_1 = \bar{X}$

1.5 CENTRAL MOMENT OF A SAMPLE

Central moment q order of a sample is

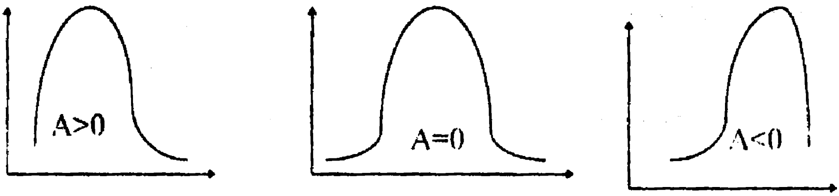
$$M_q = (\Sigma (x_i - \bar{x})^q n_i) / n$$

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

1.6 SHAPE OF DISTRIBUTION

1) *Coefficient of Asymmetry.*

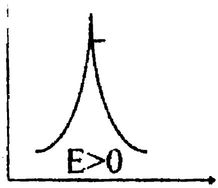
$$A = M_3 / \sigma^3$$

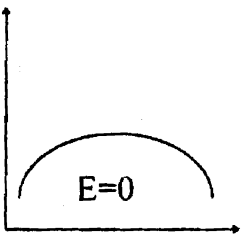


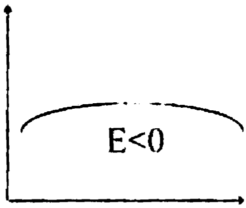
Skewedness to the left or normal or to the right

2) *Excess*

$E = (M_4 / \sigma^4) - 3$







Flatness : sharp -- normal- -- gently sloping

Example *(for using an average geometrical value).*

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

Rate→a_{i+1}/a_i , $\prod_{i=1}^{n} a_{i+1}/a_i$ →growth rate for n years. $\overline{T}$ - average $\overline{T}$ annual growth rate $\overline{T}$ -const. n=4 $T^4=\prod_{1}^{4} a_{i+1}/a_i$

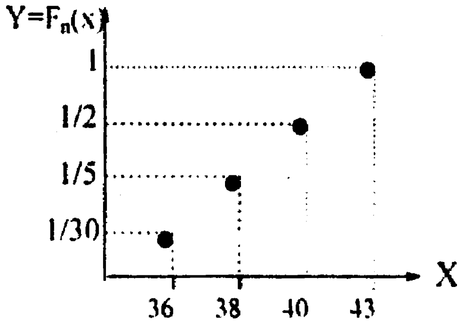
$$\overline{T} = \sqrt[4]{\prod_{i=1}^{n} a_{i+1}/a_i} = \sqrt[4]{15/5} = 1.2$$

1.7 EMPIRICAL DISTRIBUTION OF FUNCTIONS

(Relative cumulative frequency)

$F_n(x)=m(x)/n$, $m(x)$ 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.i	r.c.i



X<40 → 50%

1.8 PROPERTY OF ARITHMETICAL AVERAGE (MEAN) X:

$\begin{pmatrix} x_1 & x_2 & x_3 \dots & x_m \\ n_1 & n_2 & n_3 \dots & n_m \end{pmatrix}$ is given. $\overline{X} = (1/n) \sum_{i=1}^{m} x_i n_i$

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

if $n_i=1$ then (Arithmetical Average) $\bar{x}=(1/n)\sum x_i$

Then we have the following properties:

1. If c= const. then $c\cdot\bar{x}=(\sum c\cdot x_i n_i)/n$

2. $(\sum(x_i \pm c)n_i)/n=\bar{X}\pm c$

3. $(\sum(x_i-\bar{x})n_i)/n=\bar{X}-\bar{X}=0$

4. $(\sum x_i k n_i)/(\sum k n_i)=\bar{X}$

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	$p_1+q_1+...r_1=n_1$
...	...	...	...	...	...
x_m	p_m	q_m	...	r_m	$p_m+q_m+...r_m=n_m$
	N_1	N_2	...	N_l	$\sum N_i=N$

$\bar{X}_1=(\sum_{i=1}^{m} x_i p_i)/N_1$; $\bar{X}_l=(\sum_{i=1}^{m} x_i r_i)/N_l$

$\bar{X}=(\sum_{i=1}^{m} x_i n_i)/N=(\sum_{i=1}^{l} x_i N_i)/N$

$\sum N_i=N$

$z_i=x_i\pm y_i$; $\bar{z}=\bar{x}\pm\bar{y}$

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 is

$$\sigma^2=\left(\sum_{i=1}^{m}(x_i-\bar{x})^2 n_i\right)/n$$

Property:

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

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

3) If $k n_i$, then σ^2 stays constant

From property 2 we can obtain the following: $\sigma^2=\sum(x_i-c)^2 n_i/n-(\bar{x}-c)^2$

If $c=0$, then $\sigma^2=((\sum x_i^2 n_i)/n)-\bar{x}^2$

Simplified method of computing $\bar{X}$, σ^2

$\bar{X}=(\sum((x_i\pm c)/k)n_i)k/n\pm c$

$\sigma^2=(\sum((x_i-c)/k)^2 n_i)k^2/n-(\bar{x}-c)^2$

1.10 ELEMENTARY PROBABILITY

Classification of events

Definition 1. An event **E** is a 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, then it is called **impossible** event. **E=∅**

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

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

Definition 6. If A happens and the other $\bar{A}$ does not happen then $\bar{A}$ is 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 $\bar{A}$ are incompatible (complementary) 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 defined by A $P(A)=m/n$, 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. $P(U)=1$ $P(\Omega)=1$

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

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

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

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

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

Example.

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

b) For a die $\Omega=\{w1, w2,...w6\}$

U-is σ algebra of events

$\forall$ A, B, $\Omega \in U$, $A \cap B \in U$, $A \cup B \in U$, $A \setminus B \in U$

$U=\{\emptyset, \{w1\}, \ldots \{w6\}, \{w1, w2\}, \ldots \{w1, \ldots w6\}=\Omega\}$

Let's introduce the following *axioms*:

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

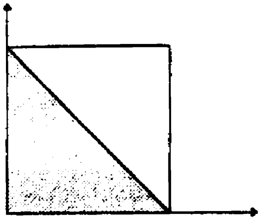
2) $P(\Omega)=1$

3) If A, B are incompatible

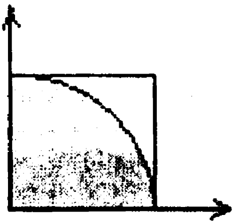
$(A \cap B)= \emptyset$, then $P(A \cup B)=P(A)+P(B)$

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

1.13 GEOMETRICAL DEFINITION



A-Measure of interest area



C-Measure of common area

Definition: The probability of event A (the point will be in black area) is P(A) = =measure (A) / measure(C)

P(A)=?

$(ab)/2 \rightarrow$ measure(A)

$ab \rightarrow$ measure (C)

$P(A) = 1/2$

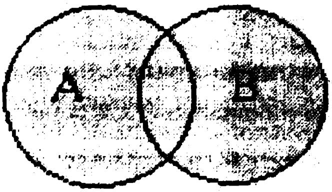
$\pi R^2/4 \rightarrow$ measure(A)

$R^2 \rightarrow$ measure(C)

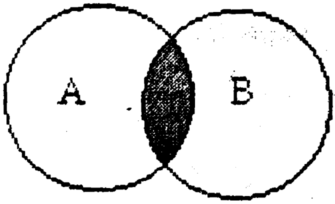
$P(A) = \pi/4 = (\pi R^2/4)/R^2$

1.14 SUM, INTERSECTION AND SUBTRACTION OF SETS

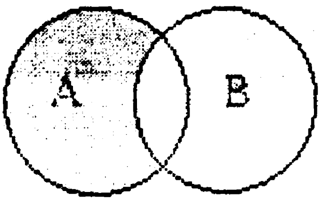
1. $C=A\cup B$, C-consists of A or B



2. $C=A\cap B$, A and B



4. $C=A\backslash B$



1.15 ADDITIONAL RULE OF PROBABILITY

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

I. The mutually exclusive case:

a) If A and B are mutually exclusive, then

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

b) If $A_1, A_2,...A_n$ 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: $A=1$, $B=2$; $A \cup B = C$ $P(C)=1/6+1/6=1/3$

Example. For tossing a coin

$P(A \cup B)=1$; A-head, B-tail

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

d) If $A_1, \ldots A_n$ is a full group of events, and A_i is mutually exclusive, 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:

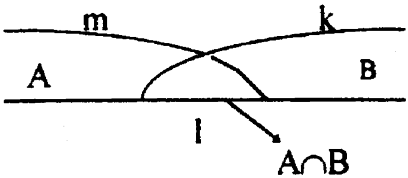
If A and B are not mutually exclusive, then

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

Examples:

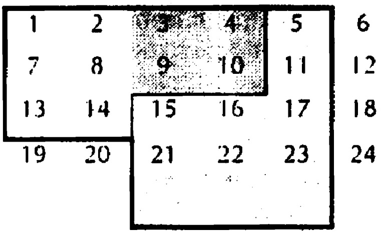
Geometrical description

Example 1.



$P(A+B) = P(A) + P(B) - P(A \cap B) = m/n + k/n - l/n$

Example 2.



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

$P(A+B) = P(A) + P(B) - P(A \cap B) = (8+9-4)/24$

Example **3. There are 52 cards in the deck**

a) 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→1 English + 2 other

C→2 English + 1 other

D→3 English + 0 other

$A=B\cup C\cup D$ B,C,D-mutually exclusive, P(A)=?

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

$P(B)=(C_5^1 C_{10}^2)/C_{15}^3$

$P(C)=(C_5^2 C_{10}^1)/C_{15}^3$, $P(D)=C_5^3/C^3_{15}$

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

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

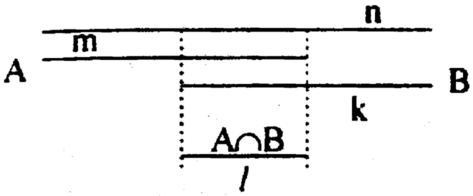
a) $P(3\ stand)=C^3_{15}/C^3_{25}$

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

1.16 CONDITIONAL PROBABILITY

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

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



$P(A/B)=l/k=(l/n)/(k/n)=P(A\cap B)/P(B)$ (*)

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

$\Omega=36$, $B(even)=18$,

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

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

$P(A/B)=9/18=(9/36)/(18/36)=P(A\cap B)/P(B)$

1.17 MULTIPLICATION RULE

From (*) we will get

$P(A\cap B)=P(B)P(A/B)$ or $P(A\cap B)=P(A)P(B/A)$

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

$P(A)=(10/50)*(9/49)$

If there are n -events $A_1, A_2,\ldots A_n$,

$P(A_1, A_2\ldots A_n)=P(A_1)P(A_2/A_1)\ldots P(A_n/A_1\ldots A_{n-1})$

If A, B are independent, then **$P(A)=P(A/B)$**

Theorem. If $A_1, A_2,\ldots A_n$ are separate independent events, then the probability is their production

$P(A_1, A_2\ldots A_n) = P(A_1)\ldots P(A_n)$

$n=2$. $P(A\cap B)=P(A)*P(B)$

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

$P=1/2*1/2*1/2=1/8$

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 $P(A)=P(B_1)P(A/B_1)+...+P(B_n)P(A/B_n)$

Proof: $A=AU=A(B_1+B_2+..... B_n)=A B_1+AB_2+..+AB_n$

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

Example. There are 3 plants. They produce lamps.

1[st] plant produces 10% of nonstandard products

2[nd] plant produces 5% of nonstandard products

3[rd] plant produces 15% of nonstandard products

Solution:

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

A-an event of buying nonstandard lamp.

B_1-an event of producing in the 1[st] plant

B_2- an event of producing in the 2[nd] plant

B_3- an event of producing in the 3[rd] plant

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

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

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

$P(B_i/A) = (P(B_i) P(A/B_i)) / (\Sigma P(B_i) P(A/B_i)) = (P(B_i)P(A/B_i)) / P(A)$

Proof.

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

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)=? P(A/D)=(P(A)P(D/A))/P(D),

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/A)=0.05; P(C)=0.35;

P(D/C)=0.01; P(A/D)=0.714 : 71.4%. Machine A produces 71.4% of the defected parts.

1.20 FORMULA OF REPEATED TRIALS (FORMULA OF BERNOULLY)

Trials are independent. An event A -tossing a coin

p-the probability of success of an event A,

q-the probability of failure of an event A

1st trial A; T or H

Corresponding prob: p, q $(p+q)^1=p+q=1$

2nd trial A: TT or TH or HT or HH

Corresponding prob: q^2 pq qp p^2

$q^2 + pq + qp + p^2 = (p+q)^2=p^2+2pr+q^2$

3rd trial

TTT, TTH, THT, HTT, THH, HTH, HHT, HHH

p^3 p^2q p^2q p^2q pq^2 pq^2 pq^2 q^3

$(p+q)^3=p^3+3p^2q+3pq^2+q^3$

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

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

Proof. Let r is a *consecutive* success, followed by n-r consecutive failures. These 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, r of them happen, and combination of occurrence equals C_n^r. The desired probability is therefore

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

Example. Find the probability of getting exactly 5 twice as a result of 7 throws of a dice.

$P_7(2)=C_7^2 p^2 q^5$; $p=1/6$; $q=5/6$

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?

$P_{12}(\leq 9) = P_{12}(9)+P_{12}(10)+P_{12}(11)+P_{12}(12)$

$P=0.1$; $q=0.9$; $n=12$

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

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)=?

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

Proof: $\bar{A}=\bar{A}_1..\bar{A}_n$, $P(A_1)=1-P(\bar{A})=1-P(\bar{A}_1...\bar{A}_n)=1-q_1...q_n$.

Example. 3 hunters 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 $p_3=0.3$. {One bullet is enough for hitting a bird} = A. What is P(A)?

Solution: $P(A)=1-q_1q_2q_3=1-0.6*0.4*0.3=0.928$

1.21 LOCAL THEOREM OF MOIVRE-LAPLACE

The probability of occurrence of an event A equals to p-const. After n trails the quantity of occurrence equals m=250. What is $P_n(m)$=?. If p=0.4, n=600, m=250. It is very difficult to calculate $P_n(m)=C^{250}_{600}*0.4^{250}*0.6^{350}$

In this case we introduce the theorem of *Moivre-Laplace*

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

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

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. $p=0.4$, $n=600$, $m=250$, $x=10/12\approx 0.83$

$f(0.83)\approx 0.282$ by using a table which is given in any textbooks.

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

1.22 THE FORMULA OF POISSON

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

In this case $P_n(m)=(\lambda^m/m!)e^{-\lambda}$

Proof. $p=\lambda/n$, $P_n(m)=C_n^m(\lambda/n)^m(1-\lambda/m)^{n-m}=((\lambda^m/m!)(n(n-1)\dots(n-m+1))/n^m)*(1-\lambda/n)^n(1-\lambda/n)^{-m}=\lambda^m/m!((1-1/n)\dots(1-(m-1)/n)(1-\lambda/n)^n(1-\lambda/n)^{-m}==(\lambda^m/m!)e^{-\lambda}$

Example. The detail is not standard. $p=0.004$, $n=1000$. $P_{1000}(5)=?$ $\Rightarrow$ $\lambda=4$.

By formula of Poisson: $P_n(m)=(\lambda^m/m!)e^{-\lambda}$ $P_{1000}(5)=0.1563$

By formila of Moivre-Laplace: $P_n(m)=f(x)/\sqrt{npq}$ $P_{1000}(5)=0.1763$

1.23 INTEGRAL FORMULA OF MOIVRE-LAPLACE

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

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

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

then $P_n(a\le m\le b)=0.5\,[\Phi(x_2)-\Phi(x_1)]$

where $x_1=(a-np)/\sqrt{npq}$, $x_2=(b-np)/\sqrt{npq}$,

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

Properties of $\Phi(x)$:

1. $\Phi(x)$ is an odd function
2. $\Phi(x)$ is monotonically increasing

so $\Phi(x_2) > \Phi(x_1)$, $x_2 > x_1$

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. $P_{1000}(455 \le x \le 545) = 0.9711$

Consequence 1. $P_n(|m-np|<r) \approx \Phi(r/\sqrt{npq})$

Proof: We can get that $np-r \le m \le np+r$. Then

$P_n(a \le m \le b) = 0.5\,(\Phi((np+r-np)/\sqrt{npq}) - \Phi((np-r-np)/\sqrt{npq})) =$

$0.5 \bullet 2 \bullet \Phi(r/\sqrt{npq}) = \Phi(r/\sqrt{npq})$

Consequence 2. Let frequency be m/n, then

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

Proof: $|m/n-p|<\Delta \to |m-np| \le n\Delta$, $r = n\Delta$

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

Example (use *Consequence 1*)

Consumer needs 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: $np=600$, $630-600=30=600-570$, $r=30$

$P(|m-600| \le 30) = \Phi(1.464) = 0.8568$

Example. Condition from above. The deviation of real relative frequency from $p=0.3$ is less than 0.02.

$P(|m/2000-0.3| \le 0.02) = \Phi(0.02\sqrt{2000/(0.3*0.7)}) = \Phi(1.952) = 0.949$

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

2.1 DISCRETE RANDOM VARIABLES

There are exist two types of Random Variables

- Discrete Random Variables
- Continuous Random Variables

Now we smoothly pass from Random events to Random Variables

Definition. If a random variable X takes value from a discrete set $X_1, X_2, \dots X_n$ 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 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 table

$$X \Rightarrow \begin{pmatrix} X_1 & X_2 & X_3 & \dots & X_n \\ P_1 & P_2 & P_3 & \dots & P_n \end{pmatrix}$$

The following conditions must be satisfied

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 Head =1, 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 shootings .

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

The law of distribution is:

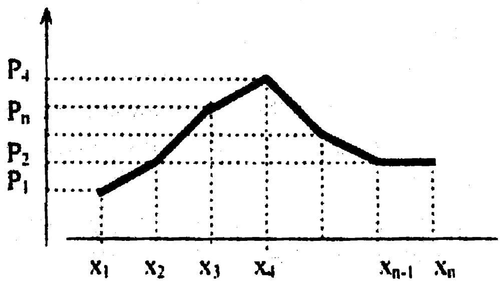
Quantity	0	1	2	3	4
Probability	0.656	0.2916	0.0486	0.0036	0.0001

When we use the formula of repeated trials for computing the probability $P_n(x=m)=C_n^m p^m q^n$, Distribution function is called **Binomial law** of distribution or the ***Bernoully*** law of distribution

When we use the formula $P(x=m)=(\lambda^m e^{-\lambda})/m!$, distribution function is called the **Poisson** distribution

Graph shows the law of distribution

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



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

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

If the events $X=x_i$, $Y=y_j$ 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 $P_i = P_{i1} = P_{i1}+\dots P_{in}$ using the theorem of full probability.

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

27. The rules for X+Y, X-Y, X·Y

We use $P_{i,j} = P(X=x_i) \bullet P_{X=x_i}(Y=y_j)$ when X and Y depend on each other. Let X and Y be independent events.

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

$$X\bullet Y=\begin{pmatrix} 3\bullet 1 & 6 & 9 & 4 & 8 & 12 & 5 & 10 & 15 \\ 0.03 & 0.03 & 0.24 & 0.04 & 0.04 & 0.32 & 0.03 & 0.03 & 0.24 \end{pmatrix}$$

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

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

2.3 MATHEMATICAL EXPECTATION OR EXPECTED VALUE OF DISCRETE RANDOM VARIBLIES

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

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·X, then M(K·X) =K ·MX

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

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

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

$$=\sum_{i=1}^{n}x_i\sum_{j=1}^{m}P_{ij}+\sum_{j=1}^{m}y_j\sum_{i=1}^{n}P_{ij}=\sum_{i=1}^{n}x_iP_i+\sum_{j=1}^{m}y_jP_j=MX+MY$$

4) If X-Y; $M(X-Y)=M(X+(-1)Y)=MX-MY$

5) If X·Y; $MX\cdot Y=\sum_{i=1}^{n}x_i\sum_{j=1}^{m}y_jP_iP_j=MX\cdot MY$

6) $M(X\pm C)=MX\pm C$

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

2.4 VARIANCE OF DISCRETE RANDOM VARIABLES

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

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

We can see that MX = MY. What is the difference X and Y ?

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

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

$\sigma_X = \sqrt{DX}$ 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 $x_1, x_2, \ldots x_n$ – independent, $X=\sum_{i=1} x_i$

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

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

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

Then, $M(\sum_{i=1}^{n} x_i)=n \cdot a$, and $M \bar{x} = a$, Where $\bar{x}$ is an arithmetical average $\bar{x}= (1/n) \cdot \sum x_i$

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

Then $D(\sum_{i=1}^{n} x_i)=n \cdot \sigma^2$, $D\bar{x}=\sigma^2/n$

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 $MX=n \cdot p$ and $DX=n \cdot p \cdot q$

Proof: $X=x_1+x_2+\ldots+x_n$ $\qquad x_i= \begin{pmatrix} 0 & 1 \\ p & q \end{pmatrix}$

$Mx_i=p \quad MX=n \cdot p$

$Dx_i=(0-p)^2q+(1-p)^2p=p^2q+q^2p=pq\,(p+q)=p \cdot q$

$$DX=\sum_{i=1}^{n} Dx_i= n\cdot p\cdot q$$

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

Proof:

Relative frequency $\frac{x}{n}=\frac{\sum x_i}{n}\Rightarrow M(\frac{x}{n})=p$

$$D(\frac{x}{n})=\frac{D(\sum x_i)}{n^2}\Rightarrow \frac{npq}{n^2})=\frac{pq}{n}$$

$$\sigma=\sqrt{\frac{pq}{n}}$$

Theorem 5. If x is distributed by law of Poisson, then $Mx=\lambda$, $Dx=\lambda$.

Proof: $P(x=m)=(\lambda^m e^{-\lambda})/m!$ $Mx^2=\lambda+\lambda^2$; $(Mx)^2=\lambda^2$; $Dx=\lambda$,m=0,1,2,..

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

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

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

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

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

2.3 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) is called as a distribution function (Integral function) of random variable ξ .

Example. A die is thrown:

x	$x \le 1$	$1 < x \le 2$	$2 < x \le 3$	$3 < x \le 4$	$4 < x \le 5$	$5 < x \le 6$	$6 < x$
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. $F(-\infty)=0$, $F(+\infty)=1$, $0 \le F(x) \le 1$

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

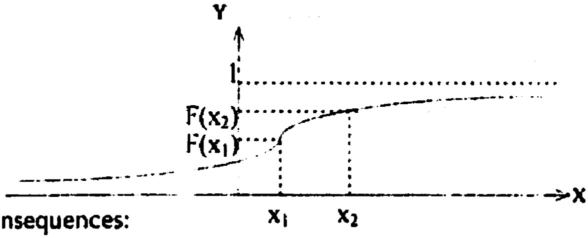
Property 4. $P(\xi = x_1) = 0$

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 y = F(x)



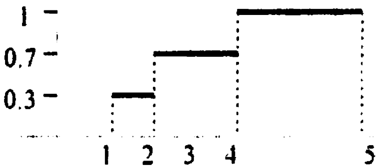
Consequences:

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 $F(x) = 0$

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

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

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

Definition 3. 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

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

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 $\xi_1, \xi_2,...\xi_n$ are independent in aggregate if for any $k\leq n$, the events $\{\xi_1<x_1\},...\{\xi_n<x_n\}$ are independent.

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

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

Definition 6. The derivative of distribution function $F(x)$, $F'(x)=f(x)$, (F-prime) is called the probability density.

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

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

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

a) $f(x)\leq 0$

b) $\int f(x)dx = 1, -\infty < x < +\infty$

Property 9.

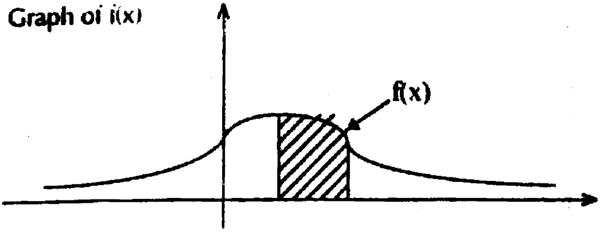
$$F(x)=\int f(x)dx=\int f(t)dt$$

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

Proof.

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

Graph of f(x)



The probability P(a<ξ<b) = square under curve

Definition 7. The random variable ξ is called uniformly distributed in [a, b], if the probability density

$$f(x)=\begin{cases} 0, & -\infty<x\le a \\ 1/(b-a), & x\in[a, b] \\ 0, & b<x<+\infty \end{cases}$$

Definition 8. If λ>0, 0≤X<+∞, and the density

$f(x)=\lambda e^{-\lambda x}$, then random variable ξ called that it has exponential law.

Definition 9. If

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

K – number of freedom.

Γ(K/2) – gamma function, $\Gamma(m)=\int_{-\infty}^{+\infty} x^{m-1}e^{-x}dx$

the random variable is called as χ^2 (Chi squared distribution)

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

$x\in[-\infty, +\infty]$

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

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

a) Calculate A?

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

Solution:

$$\int_{-\infty}^{+\infty}\frac{A}{1+x^2}dx = 1 = A\,arctgx\Big|_{-\infty}^{+\infty} = A\pi$$

$A=1/\pi$ $\quad P(0<\xi<+\infty)=\int_0^{+\infty}\frac{1}{\pi(1+x^2)}dx = 1/2$

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

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

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

Solution: From definition of distribution function

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

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{(x-a)/\sigma}e^{-\frac{t^2}{2}}dt=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0}e^{-\frac{t^2}{2}}dt+\frac{1}{\sqrt{2\pi}}\int_{0}^{(x-a)/\sigma}e^{-\frac{t^2}{2}}dt=1/2+$$

$$\Phi\left(\frac{x-a}{\sigma}\right).$$

2.6 MATHEMATICAL EXPECTATION AND VARIANCE OF CONTINUOUS RANDOM VARIABLES.

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

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

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

Proof:

1) MX = a =

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

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\sigma z e^{-\frac{z^2}{2}}dz + \frac{a}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{z^2}{2}}dz = 0$$

2) DX=

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

$$\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} z * z * e^{-\frac{z^2}{2}}dz = \{u = z; dv = ze^{-\frac{z^2}{2}}dz; (u\int dv = uv - v\int dv)\} = \sigma^2$$

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

For standard normal random variable $F_u(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{z^2}{2}}dz$

$F(x) = F_u((x-a)/\sigma)$

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

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

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. $\lim\limits_{|x| \to \infty} y = 0$

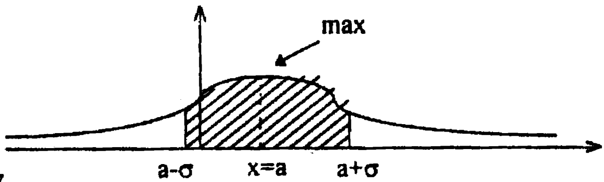
4. Point of max and min (extremum)

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

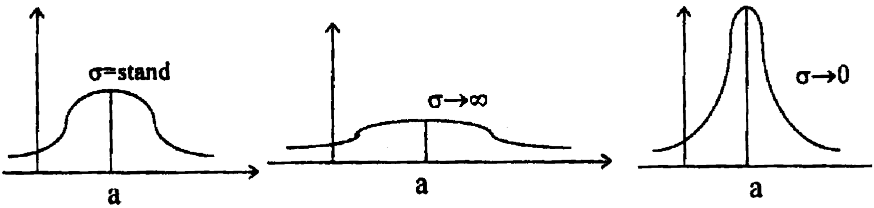
5. y is a symmetric function on line x=a

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

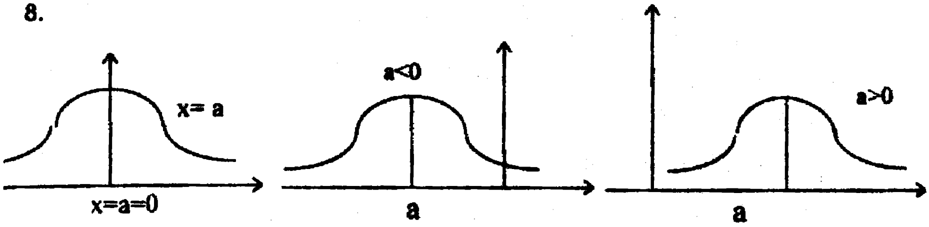
$y'' = 0$, $x = a \pm \sigma$, the points $\left(a-\sigma, \frac{1}{\sigma\sqrt{2\pi e}}\right), \left(a+\sigma, \frac{1}{\sigma\sqrt{2\pi e}}\right)$ are points of inflection.



7.



8.



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

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

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

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

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 $v_k=MX^k=\int_{-\infty}^{+\infty} x^k f(x)dx$, where f(x) is the density function.

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

$\mu_k = M(X-a)^k = \int_{-\infty}^{+\infty}(x-a)^k f(x)dx$ $\mu_0=1$, $v_0=1$, $v_1=MX$, $\mu_1=0$

$\mu_2=M(X-a)^2=DX=MX^2-(MX)^2=v_2-v_1^2$

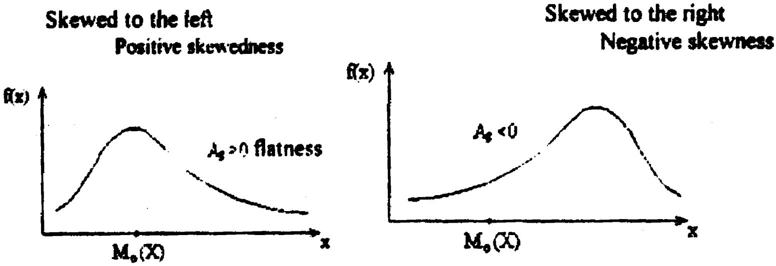
$$\mu_n = \sum_{k=0}^{n-2}(-1)^k C_n^k v_1^k v_{n-k} + (-1)^{n-1}(n-1)v_1^n$$

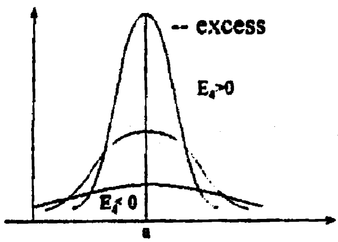
$\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3$ $\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 + 3\nu_1^4$

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

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

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





2.9 CALCULATING PROBABILITY WITH GIVEN DEVIATION

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

2.10 THE «THREE SIGMA» RULE

Let's $\Delta=\sigma t$, than $\Phi(\Delta/\sigma)=\Phi(t)$ $t=3$ $P(|x-a|<3\sigma)=2\Phi(3)=2\cdot 0.498=0.9973$. 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 $x_1, x_2, \ldots x_n$ with distribution

identical to x. Let's consider $X = \quad \rho_n = \sum_{i=1}^{n} x_i$

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

$$M\rho_n = M(\sum_{1}^{n} x_i) = n \cdot a = a*$$

$$D\rho_n = D(\sum_{1}^{n} x_i) = n\sigma^2$$

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

Using «three sigma» rule we obtain $P(|x-a*|<3\sigma*)=$

$P\{a*-3\sigma*<X<a*+3\sigma*\}= P\{n\cdot a-3\sigma\sqrt{n} <\rho_n<n\cdot a+3\sigma\sqrt{n}\}=P\{a-\frac{3\sigma}{\sqrt{n}} < \rho_n/n <$

$a+\frac{3\sigma}{\sqrt{n}}\}= P\{|(\rho_n/n)-a|<\frac{3\sigma}{\sqrt{n}}\}=0.997$

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

$a \approx \frac{1}{n}\sum_{i=1}^{n} x_i$. Error = $\frac{3\sigma}{\sqrt{n}}$, when $n\to\infty$, error$\to 0$. By modeling with

computer we can calculate the integral $a = \int_D x f(x)dx$, 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 $X=\ln Y=\psi(y)$, $\psi'(y)=1/y$, $g(x)=\frac{1}{y\sqrt{2\pi}}e^{-\frac{(\ln Y)^2}{2}}$, $y>0$, g(y) is named as the density of logarithmic normal (log normal) random variable.

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

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 example 2 $F_y(x)=F_x((x-b)/a)$

$F_y(x)=P(aX+b<x)=P(X<(x-b)/a)=F_X((x-b)/a)$

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

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

Example 4. X is uniformly distributed random variable in $[-\pi/2, \pi/2]$ Let $y=\sin X$. What is $g(y)=$?

Density function $f(x)=1/\pi$, sinx in $[-\pi/2, \pi/2]$ is increasing, inverse function

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

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

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

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

$$\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{x} p_x(x_1,t-x_2)dt\right)dx_1 = \int_{-\infty}^{x}\left(\int_{-\infty}^{\infty} p_x(x_1,t-x_1)dx_1\right)dt$$

$f_y(t)=\int_{-\infty}^{+\infty} p_x(x_1, t-x_1)dx_1$ – 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 $\underline{x_1+x_2=X}$. Find $P_{x1+x2}(t)=?$

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 $x_1, x_2,\dots, x_\nu$ be independent normal random variable with a=0, σ=1. Compose $\chi_\nu^2=x_1^2+x_2^2+\dots+x_\nu^2$

v - degree of freedom

$$v=1,\ X_1^2=x_1^2,\ p_1(x)=\begin{cases}0, if \rightarrow x\le 0\\ \frac{1}{\sqrt{2\pi x}}e^{-\frac{x}{2}}, if \rightarrow x>0\end{cases}$$

$$v=2,\ X_2^2=x_1^2+x_2^2,\quad p_2(x)=\begin{cases}\frac{1}{2}e^{-\frac{x}{2}}, if \rightarrow x>0\\ 0, if \rightarrow x\le 0\end{cases}$$

$\forall v$, $p_v(x)=k_v e^{-x/2}x^{v/2-1}$, where k_v can be found from condition $\int_{-\infty}^{+\infty} p_v(x)dx = 1$.

Home task: : Find k_v .

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

Changing the variables we can get standard normal random variable $Y_i = (x_i-a)/\sigma \sim N(0,1)$

$$\chi_v^2=\sum_{i=1}^{v}\frac{(x_i-a)^2}{\sigma^2}\rightarrow \chi^2\text{distribution}$$

In the table of χ^2 you may find distribution $P(\chi_v^2>\chi_{\alpha,v}^2)=\alpha$

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

2.15 STUDENT (OR T-) DISTRIBUTION

Let $x_0, x_1, \ldots, x_v$ be independent normal random variables

$x_i \sim N(0,\sigma)$, i=0, 1, 2,..., v

The random variable $t_v = \frac{Y_0}{\sqrt{\frac{1}{v}\sum_{i=1}^{v} x_i^2}}$, is called t—distribution or Student

Distribution random variable with degree of freedom v.

The Density function is $p_{t_v}(x)=b_v(1+x^2/2)^{-(v+1)/2}$

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

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

The distribution is not dependent on σ. If $x_0, x_1, x_2, \ldots, x_\nu$ is independent and $x_i \sim N(a, \sigma)$

$$t_\nu=\frac{x_0-a}{\sqrt{\frac{1}{\nu}\sum_1^{\nu}(x_i-a)^2}}$$

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

In the table of χ^2 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 $F_{n_1 n_2}=\frac{\frac{1}{n_1}\sum_{i=1}^{n_1}x_i^2}{\frac{1}{n_2}\sum_{i=n_1+1}^{n_1+n_2}x_i^2}$ 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

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

If σ=1, $F_{n_1 n_2} = \dfrac{\frac{1}{n_1} X^2_{n_1}}{\frac{1}{n_2} X^2_{n_2}}$, where X^2_{n1} and X^2_{n2}, 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

$$P_{n_1,n_2}(x) = \begin{cases} 0, when \rightarrow x \le 0 \\ C_0 \dfrac{x^{(n_1-2)/2}}{(n_1 x + n_2)^{(n_1+n_2)/2}}, when \rightarrow x > 0 \end{cases}$$

Here, $C_0 = \dfrac{\Gamma(\frac{n_1+n_2}{2}) n_1^{n_1/2} n_2^{n_2/2}}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})}$, $\Gamma(x) = \int_1^{\infty} t^{x-1} e^{-t} dt$

2.17 MULTIVARIABLE RANDOM VARIABLES

Let $x_1, x_2, \ldots, x_n \in \Omega$. $X=(x_1, x_2, \ldots, x_n)$ be random vector with n-dimension. We shall define the distribution function in following way $F_X(x_1, x_2, \ldots, x_n) = P\{x_1 < x_1, \ldots, x_n < x_n\}$

Here, $F_x(x_1, x_2, \ldots, x_n)$ – not decreasing function.

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

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

$p_x(x_1, x_2, \ldots, x_n)$ is called as density function.

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

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

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

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

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

If $X_1 \in A_1, X_2 \in A_2, \ldots X_n \in A_n$ and is independent, then

$P(X_1 \in A_1, \ldots X_n \in A_n)=P(X_1 \in A_1)\ldots P(X_n \in A_n)$

Properties: Let $n=2$, $X=(x_1, x_2)$.

1. If $x_i \to -\infty$ $F_X(x_1, x_2)=F_X(x_1, -\infty)=\lim F(x_1, x_2)=\lim$

$P(\{x_1<x_1\}\cap\{x_2<x_2\})=P(\varnothing)=0$ $F_x(x_1, -\infty)=0$

2. If $F_x(x_1, +\infty)=F_{x_1}(x_1)$ $F_x(x_1, +\infty)=\lim F_x(x_1, x_2)=\lim$

$P(\{x_1<x_1\}\cap\{x_2<x_2\})=P\{x_1<x_1\}=F_{X_1}(x_1)$.

$F_{x_1}(x_1)=F_x(x_1, +\infty)=$

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

$$\int_{-\infty}^{+\infty} p_X(y_1,y_2)dy_2 = p_{x_1}(y_1)$$ --density function of x_1

$$F_{x_1}(x_1) = \int_{-\infty}^{x_1} p_{x_1}(y_1)dy_1$$

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

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

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

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

Then n dimension normal random variable x will have density function

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

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

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 are 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 $a_1=a_2=0$, $\sigma_1=\sigma_2=1$ $p_{X1}(x_1)=?$, $p_{X2}(x_2)=?$

2) $X=(x_1, x_2)$ and X is 2 dimension random variable with parameters a_1, a_2, σ_1, σ_2, r

$p_{X1}(x_1)=?$, $p_{X2}(x_2)=?$

2.18 CONDITIONAL DENSITY FUNCTION

Let Z=(X, Y) be a random variable with density function of Z = $p_z(x, y)$. $p_Y(y)$ 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) $\int_{-\infty}^{+\infty} p_X(x / y)dx = 1$

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

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

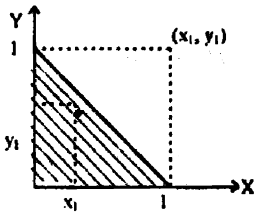
$$p_1(x_1)=\int\int_{-\infty}^{+\infty} p_X(x_1,\dots x_n)dx_2\dots dx_n$$

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

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

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

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



$p_z(x, y)= 6x$

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

$p_Y(y/x)=p_z(x, y)/p_X(x)=1/(1-x);\ 0<y<1-x$

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

$$F_Y(y/x)=\int_0^y p_y(y/x)dv = y(1-x)^{-1},\quad 0<y<1-x$$

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

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

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

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 Y on X.

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

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

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

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 is called as the level of significance and is denoted by α.

If the level of significance $\alpha=0.05$, we shall use it for initial study and when $\alpha=0.01$ 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 $\varepsilon>0$, tends to 1.

2.21 CHEBISHEV'S INEQUALITY

Lemma. If among 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 $x_i>0$. Firstly, $x_1 \cdot p_1+x_2 \cdot p_2+\ldots x_n \cdot p_n = MX$

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

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

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

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

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

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

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

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:

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

$P(|(x/10000)-0.8| \le 0.01) \ge 1-(0.8\cdot0.2)/(0.01^2\cdot10000) = 0.84$

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, $n=10000$, $p=0.1$, $q=0.9$. $MX = np = 1000$, $DX=10000\cdot0.1\cdot0.9=900$ $\quad MX-950=1000-950=50$, $1030-1000=30$

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

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

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

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

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

Consequence 1. If $MX_i=a$, $DX_i<c$, then

$$P(|\frac{1}{n}\sum_{i=1}^{n} X_i - a| \le \varepsilon) > 1 - \sigma .$$

Because, $(1/n)\Sigma a_i = \Sigma Mx_i/n = na/n = a$

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

X/n-relative frequency. $P(|x/n-(p_1+...+p_n)/n| \le \varepsilon) > 1-\sigma$

Proof. $X/n=(x_1+...+x_n)/n$, $MX_i=p_i$, $DXi=p_iq_i \le 0.25$

The condition of Chebishev theorem is satisfied.

$P(|x/n-(p_1+...+p_n)/n| \le \varepsilon) > 1-0.25/n\varepsilon^2$

Consequence 3. (Theorem of Bernoully) If the probability of event A is equal to p, then $P(|x/n - p| \le \varepsilon) > 1-\sigma$ $(x/n \sim p)$

Example 3. X1, X2,...,Xn is the sequence of independent random variables.

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

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

Solution:

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

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 relation gives a covariation. 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.

$\sigma_{ij} = M(x_ix_j)-Mx_iMx_j$, 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 $cov(x_i, x_j)\neq0$

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. $cov(c_1x_1+c_2x_2, x_3)=c_1cov(x_1, x_3)+c_2cov(x_2, x_3)$

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

$$\Sigma=\begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \dots & \dots & \dots & \dots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} \end{pmatrix}$$

The matrix Σ is called the matrix of covariation.

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

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

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

$c=(c_{ij})$ is the matrix of coefficient.

$$X=\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \quad Y=\begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} \qquad C=\begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mn} \end{pmatrix}$$

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

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

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

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

$D(x_1\pm x_2) = Dx_1+Dx_2\pm 2\sigma_{12}$

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 $c_k=c_l=1$; $k=1,2,\dots,n$. 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 covariation. ·

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

$$\Sigma=\begin{pmatrix} \sigma_{11} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_{nn} \end{pmatrix} \qquad \sigma_{ii} = \sigma^2_{xi}$$

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

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

$x_i \backslash y_i$	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_i	0.2	0.25	0.3	0.25	$\sum_{i=1}^{3}\sum_{j=1}^{4} p_{ij} = 1$

Let $u_1=x-y$, $u_2 = 3x-2y$

Find $Du_1=?$, $Du_2=?$, $u=(u_1, u_2)$, H-? H is the matrix of covariation.

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

$5\cdot 0.2\cdot 0.05 + 5\cdot 0.3\cdot 0.05+6\cdot 0\cdot 0+\ldots+7\cdot 0.3\cdot 0.1=0.975$

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

$\sigma_{12}=\sigma_{21}=cov(x,y)=M(xy)-MX\cdot MY=0.975-5\cdot 8\cdot 0.16=0.047$

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

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

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

$Du_1=D(X-Y)=DX+DY-2\sigma_{12}=\sum_{k=1}^{2}\sum_{l=1}^{2}\sigma_{kl}c_k c_l=0.56+0.037-2\cdot 0.047=0.503$

$Du_2=D(3X-2Y)=\sum_{k=1}^{2}\sum_{k=1}^{2}c_k c_l \sigma_{kl}=3\cdot(-3)\cdot 0.56+3\cdot(-2)\cdot 0.047+(-2)$

$\cdot 3\cdot 0.047+(-2)\cdot(-3)\cdot 0.037=5.04-0.282-0.282+0.148=4.624$

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

Let $X=(x_1, x_2)$, where x_i is random variable with parameter

$Mx_i = a_i; \sigma_{x_i} = \sqrt{Dx_i}$. We transform $Y_i=(x_i-Mx_i)/\sigma_{xi}$, $MY_i=0$, $DY_i=1$

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

Let's calculate $cov(Y_1, Y_2)=cov((x_1-Mx_1)/\sigma_{x1}; (x_2-Mx_2)/\sigma_{x2}) = cov(x_1-Mx_1;$ $x_2-Mx_2)/(\sigma_{x1}\sigma_{x2}) = cov(x_1, x_2)/(\sigma_{x1}\sigma_{x2})$

We shall define $\rho_{x1,x2}=cov(x_1,x_2)/(\sigma_{x1}\sigma_{x2})$ as a measure of stochastically dependents x_1, x_2 and this is called as a coefficient of correlation.

$\rho_{x1,x2}=cov((x_1-Mx_1)/\sigma_{x1},\ (x_2-Mx_2)/\sigma_{x2}) = cov(Y_1, Y_2)$

$cov(x_1, x_2) = \rho_{x1,x2}\sigma_{x1}\sigma_{x2}$

For independent random variable $\rho_{x1,x2} = 0$, the inverse conclusion is not right.

Property 1. $|\rho_{x1,x2}| \leq 1$

Proof: $Y_i = (x_i-Mx_i)/\sigma_{xi}$, $i = 1,2$

$D(Y_1 \pm Y_2) = DY_1+DY_2 \pm 2cov(Y_1, Y_2) = 2 \pm 2\rho_{x1,x2} \geq 0$

$1 \pm \rho_{x1,x2} \geq 0, \quad 1+\rho_{x1,x2} \geq 0, \quad 1-\rho_{x1,x2} \geq 0, \quad |\rho_{x1,x2}| \leq 1$

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

$|\rho_{x1,x2}| = 1$ necessary and enough condition, where $a\neq 0$, $\rho_{x1,x2} = 1$, if $a>0$ and $\rho_{x1,x2} = -1$, if $a<0$.

Proof: (Necessary) Let $\rho_{x1,x2} = 1$, $D(Y_1-Y_2) = 2(1-\rho_{x1,x2}) = 0$;

It means that $Y_1-Y_2 = C$

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

$Y_1-Y_2=0 \quad Y_1 = Y_2,\ (x_1-M_{x1})/\sigma_{x1} = (x_2-Mx_2)/\sigma_{x2}$

$x_2=ax_1+b$ where $a = \sigma_{x2}/\sigma_{x1}$, $b=Mx_2-(\sigma_{x2}/\sigma_{x1})Mx_1$

$x_2 = (\sigma_{x2}/\sigma_{x1})x_1+(Mx_2-(\sigma_{x2}/\sigma_{x1})Mx_1)$

For $\rho_{x1,x2} = -1$, $D(Y_1, Y_2) = 0$; $x_2 = (\sigma_{x2}/\sigma_{x1})x_1+(Mx_2+(\sigma_{x2}/\sigma_{x1})Mx_1)$

It is enough that: $\rho_{x1,ax1+b} = \begin{cases} 1, & a>0 \\ -1, & a<0 \end{cases}$

If $x_2=ax_1+b$, $Y_2=(x_2-Mx_2)/\sigma$, $x_2 = (ax_1+b-M(ax_1+b))/(\sigma ax_1+b) = (a(x_1-Mx_1)/(|a|\sigma x_1)= Y_1a/|a|$; $Y_2 = Y_1a/|a|$

$\rho_{x1,x2} = cov(Y_1, Y_2) = cov(Y_1, Y_1a/|a|) = cov(Y_1, Y_1)a/|a| =$

$a/|a| = \begin{cases} 1, & a>0 \\ -1, & a<0 \end{cases}$

2.24 REGRESSION FUNCTION

Let Y and ax+b be dependent. Consider Y-(ax+b) in which a, b $\min D(Y-ax-b)=?$

If $a_0=\rho_{xy}\sigma_y/\sigma_x$, $b_0 = MY-\rho_{xy}(\sigma_y/\sigma_x)MX$; Then $\min D(Y-a_0x-b_0) = \sigma_Y^2(1-\rho^2_{xy})$

$\rho_{xy}\to 1 \qquad D(Y-a_0x-b_0)\to 0$

$Y= \rho_{xy}(\sigma_y/\sigma_x)X+MY-\rho_{xy}(\sigma_y/\sigma_x)MX$

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

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

$\rho_{ij} = \sigma_{ij}/(\sigma_{xi}/\sigma_{xj})$

$$R=\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ p_{n1} & p_{n2} & \cdots & p_{nn}\end{pmatrix}$$ Matrix of correlation

Example.

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

$$\rho_{11} = \rho_{22} = 1 \quad \rho_{12} = \frac{\sigma_{12}}{\sqrt{\sigma_{12}}\sqrt{\sigma_{22}}} = \frac{0{,}047}{\sqrt{0{,}56}\sqrt{0{,}037}} = 0{,}33$$

$y = MY+\rho_{xy}(\sigma_y/\sigma_x)(x-MX) = 0.84x-0.32$

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

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

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

If r = 0, X and Y are independent,

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

2.25 MARKOV CHAIN

Assume that we have $S_1, S_2, \dots, S_n$- position

P_0—initial distribution, P—matrix of position

$$P_0 = \begin{pmatrix} p_1 \\ \dots \\ p_n \end{pmatrix} \qquad P = \begin{pmatrix} p_{11} & p_{22} & \dots & p_{1n} \\ \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix}$$

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

Marcov chains will use when we study random processes.

III. STATISTICAL ESTIMATION OF PARAMETER 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. 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 $\theta_n = f(x_1, \dots, x_n)$

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

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

Definition 7. An estimator is 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.

$$\lim_{n\to\infty} P\{|\theta_n - \theta| < \varepsilon\} = 1$$

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

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

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

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

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

Proof

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

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

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

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

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

Theorem 6. $\hat{S}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ 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 $x_1, x_2, \ldots, x_n$ are independent trials of X $Mx_i=\mu$, $Dx_i=\sigma^2$, then the arithmetical mean $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i \in N(\mu, \frac{\sigma^2}{n})$.

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

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

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

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

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

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

Remark:

I. $z = \frac{\bar{x} - \mu}{\hat{S}}\sqrt{n}$ – Student distribution or T- distribution

For example:

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

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

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

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

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

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

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

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

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

$$\Delta = (z_p\frac{\sigma}{\sqrt{n}}) = 1{,}96\frac{2}{\sqrt{16}} = 0.98$$

$\mu \in [\overline{x} - 0.98, \overline{x} + 0.98]$ with probability of 0.95 if $x = 4.1$ $\mu \in$ $[3.12, 5.08]$ $9.12<\mu<5.08$

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

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

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

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

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

Example. If n=9, p=0.95, S=3, x=6

$$6-\frac{3}{\sqrt{9}}1.96<\mu<6+\frac{3}{\sqrt{9}}1.96 \rightarrow \quad 4.04<\mu<7.96$$

From Student distribution, when p = 0.95, n = 9,

$t_{8,0.95} = 2.31$

$$\bar{x}-t_{n,p}\frac{\hat{S}}{\sqrt{n}}<\mu<\bar{x}+t_{n,p}\frac{\hat{S}}{\sqrt{n}}, \qquad 6-\frac{3}{\sqrt{9}}2,31<\mu<6+\frac{3}{\sqrt{9}}2.31$$

$3.69 < \mu < 8.31$

3.3 INTERVAL ESTIMATION FOR σ^2

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

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

$P(\chi^2<\chi_1^2) = P(\chi^2>\chi_2^2) = \alpha/2$

$S^2 = 10$, $p = 0.96$, If $n = 20$, $p_2 = \alpha/2$,

$p_1 = 1-\alpha/2$, $\alpha = 0.04$.

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

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

$$\sqrt{5,.935}<\sigma<\sqrt{23,256}\rightarrow 2.43<\sigma<4.82$$

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 $\frac{\lambda^m e^{-\lambda}}{m!}$

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

m	λ=0.7	λ=0.8	λ=0.9	λ=1.0	λ=2.0	λ=3.0
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 ²3. 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

x	0	1	2	3	4	5	6	7	8	9
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
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 º4. 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
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

x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
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

x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
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
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)

The level of confidience α =0,01

k2	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

The level of confidience α =0,05

k2	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 $P\{\chi^2_k > \chi^2_{\alpha,k}\} = \alpha$, 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. Gmurman V.E. Theory of probability and mathematical statistics. Moscow, «Visshaya shkola» 1982.
2. Kolemaev V.E. Theory of probability and mathematical statistics. Moscow, «Visshaya shkola» 1994.
3. Chistiyakov B.B. Theory of probability and mathematical statistics. Moscow, «Visshaya shkola» 1996.
4. Sirajiddinov S.X., Mamatov M.M Theory of probability and mathematical statistics. Tashkent, «Ukituvchi» 1980.
5. Amir D. Aczel ., Complete Business Statistics. IRWIN. Homewood, IL 60430, Boston, MA 02116, 1989.
6. John E. Hank, Arthur G. Reitsch. Understanding Business Statistics. Boston, 1994.
7. Masson Lind. Statistical techniques in business and economics. IRWIN. Homewood, IL 60430, Boston, MA 02116, 1990.
8. Brawn D. Applied Statistics. Springers-Kerlog. 1987.
9. Groebner D., Essential of Business Statistics. London, 1987.
10. Anderson D., ,D.Swrcney,T. Williams., Quantitative methods for business., West Publ. Company.,1992.
11. Zamkov O., Mathematical mehtods in economy., M.,Moskow., 1999.
12. Zaysev I.A. Higer Mathemetics., M., Visshay skola., 1998.

CONTENTS

THE UNIVERSITY OF WORLD ECONOMY AND DIPLOMACY

LECTURE ON THEORY OF PROBABILITY AND MATHEMATICAL STATISTICS

Author Prof. A. Rasulov

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

Duplicated by the printing house of the University of World Economy and Diplomacy.
Tashkent – 700113, Buyuk Ipak Yuli str.,54